A perfect random number generator

René Blacher

A Perfect Random Number Generator

René BLACHER

Laboratory LJK
Université Joseph Fourier
Grenoble
France

Summary : In this report one obtains a method to generate random numbers whose the randomness is proved. In this aim, one transforms data resulting from electronic files. One can also use data provided by machines. The randomness is proved by mathematical theorems and logical reasoning. This method can be applied directly in computers in the same way that the function "random". In this case, it required neither chips nor machines: one transforms the data provided by the electronic files. It can be also applied with the machines and the chips. In this case one transforms the random noises which they provide. Of course, one can also transform the majority of the noises which exists. Then, one obtains really independent sequences contrary to the current methods for machines which often are satisfied to remove the linear correlation. Moreover that can put a stop to certain malfunctions of the machines or of the chips.

Résumé : Dans ce rapport on donne une méthode pour générer des nombres aléatoires dont on soit sûr de la qualité. Pour cela, on transforme des données issues de fichiers informatiques ou fournies par des machines. L'aléarité est prouvée par des théorèmes mathématiques et par des raisonnements logiques. Cette méthode peut être appliquée directement sur ordinateur de la même manière que les fonction "random". Dans ce cas, elle ne nécessite ni puces ni machines : on transforme les données fournies par les fichiers informatiques. Elle peut être aussi appliquée avec les machines, les puces et la plupart des bruits. Dans ce cas, on transforme les bruits aléatoires qu'elles fournissent. On obtiendra alors des suites réellement indépendantes contrairement aux méthodes actuelles qui souvent se contentent de supprimer la corrélation linéaire des machines. De plus cela peut mettre fin à certains dysfonctionnements des machines ou des puces.

Key Words : Central limit theorem, Or exclusive, Fibonacci sequence, Random numbers, Random noise, Higher order correlation coefficients.

NOTICE

This report represents the result of many years of work. It should be read in relation to a secund report which we go published on this subject "A perfect random generator II" (for any information, ask us in rene.blacher@imag.fr).

This new report will be mainly a summary of the results of this report with a simpler presentation. That is all the more necessary as certain results evolved here constitute already a complete subject of study. All the results of this report thus have not inevitably to be read to have a clear comprehension of the construction of an random sequence. But they are not less essential to prove than the obtained sequences are well random.

The publication of this report was delays owing to the fact that each time we find a result, we find immediately another best than this one.

Thus, we obtained much more numerical results than those which we described in this report. It is the same for the theoretical results. Some are even not described here because they are not less useful. For example, we described only a little the possibility that, starting from a sample $x_n \in \{0/m, ..., m/m\}$, n=1,2,...,N, we can almost always describe it by a continuous model if m is rather large compared to m. However, this result can be useful when one employs the Central Limit Theorem for understanding well that the conditional probabilities have a continuous curve which is not concentrated close to a few points.

There are thus many results which we cannot give for lack of place. However, all these results lead to a certainty: the numbers obtained are well random. For those which would need more than precise details, do not hesitate to contact us on rene.blacher@imag.fr.

Anyway this research being of an entirely new type, it is possible that many papers where useful points will be developed follow this report. Indeed, many improvements can be made and the many leads of study exist: they are sometimes quoted in this report.

On the other hand, because this report is very long, it is possible that in spite of our efforts, the errors have slipped through to us during the correction. Thank you for excuse us and to point them to us.

Contents

Notations and Abbreviations

We remind notations and abbreviation used in this report.

CLT : CLT for Central Limit Theorem
XORLT : XORLT for XOR Limit Theorem : cf section 5.2
$F^*(m)$: $F^*(m) = \{0, 1, ..., m-1\}$ where $m \in \mathbb{N}^*$.
F(m) : $F(m) = \{0/m, 1/m, ..., (m-1)/m\}$ where $m \in \mathbb{N}^*$.
μ_m : μ_m is the uniform measure over F(m).
μ_m^* : μ_m^* is the uniform measure over $F^*(m)$.
μ : μ is the Lebesgue measure over $\mathbb{R}$.
L : L is also the Lebesgue measure over $\mathbb{R}$ or over $\mathbb{R}^p$.
E_2 : $E_2 = \left\{\ell, \overline{T(\ell)} | \ell \in \{0, 1, ..., m-1\}\right\}$ where T is a congruence.
Ob(1) : Ob(1) is the classical "O" with the condition $|Ob(1)| \leq 1$.
j_s : j_s, s=1,2,...,p, is an injective sequence such that $j_s \in \mathbb{Z}$, $j_1 = 0$.
j'_s : j'_s, s=1,2,...,p is a sequence such that $j'_s \in \mathbb{N}$, $j'_1 = 0 < j'_2 < ... < j'_p$.
Bo : Bo means a Borel set of $\mathbb{R}^p$.
I : I or I_s mean intervals of $\mathbb{R}$.
$N(Bo)$: $N(Bo)$ is associated to a sample x_n . $N(Bo)$ is the number of x_n which belongs to Bo.
ϵ : ϵ is associated to a sequence of random variables X_n which satisfies $P\{X_n \in Bo\} = L(Bo) + Ob(1)\epsilon$ for all n and for all Bo.
$\lfloor x \rfloor$: $\lfloor x \rfloor$ means the integer part of $x \in \mathbb{R}_+$.
T : T means a congruence $T(x) \equiv ax$ modulo m.
T_q^d : T_q^d is the function of Fibonacci defined in definition 1.3.5 .
T_q : $T_q = T_q^2$.
T^* : T^* is the function of fundamental : cf theorem 1
fi_n : fi_n is the sequence of Fibonacci : $fi_{n+2} = fi_{n+1} + fi_n$, $fi_1 = fi_2 = 1$.
$\overline{h}^m$: $\overline{h}^m \equiv h$ modulo m and $0 \leq \overline{h} < m$.
$\overline{h}^1$: $\overline{h}^1 = \overline{mh}^m/m$ when $h \in F(m)$.
$\overline{h}$: $\overline{h} = \overline{h}^m$ when m is given unambiguous.
$\overline{T}$: $\overline{T}(k) = \overline{T(k)}$, $k \in F^*(m)$.
$\widehat{T}$: $\widehat{T}(k/m) = \overline{T}(k)/m$.
$x \approx y$: $x \approx y$, where x,y $\in \mathbb{R}$, if, numerically X is approximately equal to y.

$N(0,\sigma^2)$: $N(0,\sigma^2)$ means the normal distribution with mean 0 and variance σ^2.

$Y \sim N(0,1)$: $Y \sim N(0,1)$ if Y is a random variable which has the distribution N(0,1).

X_G : $X_G \sim N(0,1)$.

X_G^* : $X_G^* = X_G$.

$x << y$: $x << y$, where $0 \leq x < y$, if and only if $x/y \approx 0$.

$\Gamma(b)$: $\Gamma(b) = P\{|X_G| \geq b\}$ where $X_G \sim N(0,1)$.

$P\big\{X_n \in I | x_2, ..., x_p\big\}$: $P\big\{X_n \in I | x_2, ..., x_p\big\} = P\big\{X_n \in I | X_{n+j_2} = x_2, ..., X_{n+j_p} = x_p\big\}$ where X_n is a sequence of random variables.

$\overline{\overline{0, d_1 d_2 ...}}$: $\overline{\overline{0, d_1 d_2 ...}}$ means the writing base 2 (or base d) of a number $z = \overline{\overline{0, d_1 d_2 ....}}$

A,a,B,b,C,c,.... : When one has a sequence of real numbers a_n which can be regarded as a realization of a sequence of random variables, one will always note datas a_n with small letters and the random variables A_n with CAPITAL LETTERS : $a_n = A_n(\omega)$ where the A_n's are defined on a probability space $(\Omega, \mathcal{A}, P)$.

$b^1(n')$ and $b^2(n)$: $b^1(n')$ and $b^2(n)$ are the sequences of random bits concretely obtained.

$X_n \overset{D}{\to} X$: $X_n \overset{D}{\to} X$ means that the sequence of random variables X_n converges en distribution to X.

$X_n \overset{P}{\to} X$: $X_n \overset{P}{\to} X$ means that the sequence of random variables X_n converges in probability to X.

C_n^p : $C_n^p = \frac{n!}{p!(n-p)!}$

H(n) : $H(n) = \big\{m = 1, 2, ..., N \big| \exists \mu : m = n + c(\mu)\big\}$: cf Notation 9.2.2

$H^*(n)$: $H^*(n) = \big\{m = 1, 2, ..., N \big| \exists \mu : m = n + c(\mu), m \neq n\big\}$: cf Notation 9.2.2

$H(n,q)$: $H(n,q) = \big\{m = 1, 2, ..., N \big| \exists\ \mu : |n + c(\mu) - m| \leq q\big\}$: cf Notation 9.2.3

$H^*(n,q)$: $H^*(n,q) = H(n,q) \setminus H(n)$: cf Notation 9.2.3

$o_P(1)$: $X_n = o_P(1)$ if $X_m \overset{P}{\to} 0$: cf notation 9.1.1.

$O_P(1)$: $X_n = O_P(1)$ if the sequence of random variables X_n is bounded in probability.

$E\{X\}$: $\mathbb{E}\{X\}$ means the expectation of the random variables X

σ_s^2 : σ_s^2 means always variances.

Card(E) : Card(E) is the cardinality of the set E.

Chapter 1

Introduction

1.1 General presentation of the matter

In this report, we present a new method to obtain sequences x_n of random numbers. This method can be used as well with machines as directly on a computer alone.

Let us announce immediately that, by abuse of language, we will call also "IID sequence" (Independent Identically Distributed) the sequences of random numbers. Indeed, let x_n be a sequence of random numbers. Which one wishes, it is that x_n can be regarded like a sample of a sequence of IID random variables X_n defined on a probability space (Ω, A, P) such that, for all $n \in \mathbb{N}^*$, $x_n = X_n(\omega)$ where $\omega \in \Omega$.

One imposes that X_n has the uniform distribution. If it is not the case, we shall precize it.

1.1.1 Presentation of the result

It is wellknown that numbers which are chosen randomly are useful in many different kinds of applications. To have such number two methods exists :

1) Use of pseudo-random generators

2) Use of random noise.

These two methods have different defects.

1) For the best of them, the pseudo-random generators seem nondeterminist only during a certain time. This can be long enough for the cryptographic generators, but it is with the current means of calculations. Moreover in simulation, the pseudo-random generators must be tested for each application : cf [2] page 151.

2) If random noises are used, bias and dependences can appear : cf [3]. One tries to remove them by mathematical transformations. But these methods have defects. They remove bias and the linear correlation, but not necessarily the

dependence. Indeed, they are satisfied to remove the linear correlation: it does not remove the correlation of higher order : cf [10].
On the other hand, these random noises can be produced by machines or chips. In this case, that thus require additional material which can suffer from malfunctions extremely difficult to detect : cf [1] page 3.

Now, for some applications, a maximum quality is essential (Nuclear power, medical, cryptography). It is thus necessary to have generators without defects. But, up to now *no completely reliable solution had been proposed* .

To set straight this situation, Marsaglia has created a Cd-Rom of random numbers by using sequences of numbers provided by Rap music, by a machine and by a pseudo-random generator. However, it does not have proved that the sequence obtained is really random.

However, there exists simple means of obtaining random sequences whose the quality is sure.

One can obtain perfect generators by using random noises, for example those produced by the machines. In this case, one transforms these noises in a more effective way. Indeed, one uses assumptions much weaker than those of the current methods

One can also obtain perfect generators usable directly on computer (without the use of machines). In this case, one uses the electronic files as random noises (like Marsaglia uses Rap music). Then they are transformed by the same method that we use for the machines.

One can thus obtain sequences of real numbers which are **proved** random, which is a completely new result. To obtain this proof, one uses mathematical properties and simple logical reasoning.

1.1.2 Summary of the method

When one uses random noise, bias and linear correlations are removed by current methods. In this aim, one supposes that theses noises check some assumptions. But, generally, those are not checked. Moreover, for each samples x_n there exists many possibles models X_n such that $x_n = X_n(\omega)$. That can be problematic.

Our method consists to transform random noises under very weak hypotheses: we assume only that theses noises are not completely deterministic. It is really a very weak assumption.

Moreover our results are true for *all* logical models possible. That suppress the problems. That allows also to satisfy the mathematical definitions of the random numbers. Though these definitions are very difficult to etablish.

Then, the obtained sequences will be always IID.

Now one can apply this method to many noises. So texts can be regarded as noises which satisfy these assumptions. It is also the case for numerous softwares which are recorded on computers : systems softwares for example.

Therefore, one can obtain directly IID sequences by transforming the files of computers. In this case, it is not necessary to use machines in order to have true random numbers.

On the contrary some electronics files can be studied logically. Then obtained numbers are surer than thoses obtained by machines which can have also malfunctions.

Of course, one can apply also our methods to noises furnished by machines: of course, our results are much surer than those of current methods.

1.1.3 Definition of randomness

To produce a really random sequence, it is thus necessary initially to have a definition of the randomness. It is a subject which was studied much. But, it is extremely complex. Philosophical questions are even involved. A summary of this study is in the book of Knuth [1] pages 149-183. One reminds some definitions in section 2.1. In fact, one will see that no current definition is really satisfactory.

Thus in some definitions one uses numerical approximations which one notes in the following way.

Notations 1.1.1 *One notes the approximation by $\approx$: for $x, y \in \mathbb{R}$, one sets $x \approx y$ if numerically x is nearly equal to y.*

Indeed, one can think to define randomness by the following way.

Definition 1.1.2 *: Let L be the Lebesgue measure. A sequence $x_n \in [0,1]$ is said random if, for all Borel set Bo, for all n+1, if the past $x_1, x_2, , x_n$ is given, one cannot predict the place of x_{n+1} with a probability very different from that of the uniform distribution : $P_e\{x_{n+1} \in Bo | x_1, ..., x_n\} \approx L(Bo)$, where $P_e\{x_{n+1} \in Bo | x_1, ..., x_n\}$ is the empirical conditional probability of Bo when the past is given.*

This type of definition it is that which one wishes. Unfortunately, it has a defect : one does not have specified enough the approximation. On the one hand, the definition of $\approx$ is very undetermined mathematically. On the other hand, one would like a definition closer to the statistics definitions. But it is almost impossible to obtain such a definition : cf section 2.1.1.

But these questions of mathematical definition will not obstruct us because we will circumvent this problem by using sequences which are really samples of sequences of random variables X_n.

Unfortunately, an infinity of models X_n corresponds to the sequence x_n. Then, there is the problem of the choice of the model X_n. We will avoid this problem by proving that x_n behave as an IID sequence for all the logical possible models.

1.2 Idea of the solution

Our method rests on a simple idea: to transform random noises by adapted transformations.

Like random noises, one can use those provided by the machines. It is what Vazirani, Neumann, Elias and others (cf [46] [4], [8]) wanted to do, but with too restrictive assumptions.

One can also use some electronic files. It is what Marsaglia did with the Rap music (cf [1], [20]). But he has transformed these data in a too elementary way (cf chapter 3).

Then, one has sequences of random noises y_n. One can regard these y_n as a realization of a sequence of (not IID) random variables $y_n = Y_n(\omega)$. The associated random variables will be always thus noted throughout this report.

Notations 1.2.1 *When one has a sequence of real numbers which one can regard as one realization of a sequence of random variables, one will always note with small letters the data and with CAPITAL LETTERS the random variables which one will suppose defined on a probability space* (Ω, A, P).

For example, for all n, $a_n = A_n(\omega)$ *for all i=1,...,N., where* $\omega \in \Omega$.

When the y_n's mean random noises, to consider that $y_n = Y_n(\omega)$ is a traditional and normal assumption. It is also true for the y_n extracted from certain electronic files. We thus have data $y_n = Y_n(\omega)$ where Y_n is a sequence of random variables not completely deterministic (and same often Qd-dependent).

1.3 Fundamental properties

Into this section, we introduce the properties which are at the heart of our study. We will use the following notations.

Notations 1.3.1 *: The notation Ob(.) is that of the classical "O(.)" with the additional condition* $|Ob(1)| \leq 1$.

The sequences $j_1, j_2, ..., j_p$, $p \in \mathbb{N}^*$, *mean alway finite injective sequences* $j_s \in \mathbb{Z}$, *such that* $j_1 = 0$. *On the other hand, the sequences* $j'_1, j'_2, ..., j'_p$ *satisfy moreover* $0 = j'_1 < j'_2 < ..... < j'_p$.

The notation $P\big\{X_n \in Bo\big|x_2, ..., x_p\big\}$ *means always the conditional probability that the random variable* X_n *belongs to the Borel Set Bo given* $X_{n+j_2} = x_2$,.....,$X_{n+j_p} = x_p$.

Let $m \in \mathbb{N}^*$. *We set* $F(m) = \{0/m, 1/m, ..., (m-1)/m\}$ *and* $F^*(m) = \{0, 1, ..., m-1\}$. *We note by* μ_m *and* μ^*_m *the uniform measures on F(m) and* $F^*(m)$ *, respectively :* $\mu_m(k/m) = 1/m$.

Fundamental transformation

Because one is in the finite case, the densities always exist and one can suppose that the condition of Lipschitz is always checked.

Proposition 1.3.1 *: Let $Z_n \in F(m)$, n=1,...,N, be a sequence of random variables.*

For all injective sequence j_s, we note by $f_{n,j(.)}\big\{z|z_2,.....,z_p\big\}$ the conditional density of Z_n with respect to μ_m given $Z_{n+j_s} = z_s$, s=2,3,...,p.

Then, there exists $K_0 > 0$, such that, for all $n \in \mathbb{N}$, for all sequence j_s, for all $(z, z') \in F(m)^2$,

$$\Big|f_{n,j(.)}\big\{z|z_2,.....,z_p\big\} - f_{n,j(.)}\big\{z'|z_2,.....,z_p\big\}\Big| \leq K_0|z - z'| \ . \tag{1.1}$$

One can then state the fundamental theorem.

Theorem 1 *: We keep the previous notations. Then for all $\epsilon > 0$, there exists an application $T^* : F(m) \to F(m)$ such that, for all interval I, for all $n \in \mathbb{N}$, for all sequence j_s,*

$$P\big\{X_n \in I\big|x_2,.....,x_p\big\} = L(I) + Ob(1)\epsilon \ \ , \tag{1.2}$$

where ϵ is function of K_0 and I and where $X_n = T^*(Z_n)$, $x_n = T^*(z_n)$.

This theorem is proved in chapter 4.1. It will be understood that ϵ depend on K_0. In the chapter 6.1, this theorem is studied in a more precise way when one uses the Fibonacci congruence to define T^*.

Transformation of Fibonacci

What is important, it is that ϵ can be enough small according to the choice of T^*. However, one carries out that the functions associated with the Fibonacci congruence give the smallest ϵ.

Definition 1.3.2 *Let fi_n be the Fibonacci sequence : $fi_1 = fi_2 = 1$, $fi_{n+2} = fi_{n+1} + fi_n$. Let T be a congruence $T(x) \equiv ax$ modulo m such that there exists $n_0 > 3$ satisfying $a = fi_{n_0}$ and $m = fi_{n_0+1}$. Then T is said a Fibonacci's congruence with parameters a and m (or more simply m),*

In theorem 1, one can choose $T^*(x/m) = \overline{T}(x)/m$ where $\overline{T}$ is defined by the following way.

Notations 1.3.3 *Let $h \in \mathbb{Z}$ and $m \in \mathbb{N}^*$. We define $\overline{h}^m$ by the following way*
1) $\overline{h}^m \equiv h$ modulo m.
2) $0 \leq \overline{h} < m$.

If the choice of m is obvious, we simplify $\overline{h}^m$ into $\overline{h}$.

In the same way, if the choice of m is obvious, and if T is a congruence : $T(x) \equiv ax + c$ modulo m, we set $\overline{T}(x) = \overline{T(x)}^m$.

One uses also the following notations.

Notations 1.3.4 *Let $h \in F(m)$. We define $\overline{h}^1$ by $\overline{h}^1 = \overline{mh}^m/m$. Often we simplify $\overline{h}^1$ into $\overline{h}$.*

The reduction of Fibonacci congruences to their first bits will be very useful for our study.

Definition 1.3.5 *Let $q, d \in \mathbb{N}^*$. Let T be the congruence of Fibonacci modulo m.*

We define the function of Fibonacci $T_q^d : F(m) \to F(d^q)$ by $T_q^d = Pr_q^d \circ \widehat{T}$ where
1) $\widehat{T}(x) = \overline{T}(x)/m$
2) $Pr_q^d(z) = \overline{\overline{0, d_1 d_2 .... d_q}}$ where $z = \overline{\overline{0, d_1 d_2 ...}}$ is the writing of z in base d.

If d=2, we simplify T_q^d in T_q and Pr_q^d in Pr_q.

Standardization by the functions of Fibonnacci

These functions T_q make uniform marginal distributions of sequences of random variables $Z_n \in F(m)$. In order to better understand that, we need the following notations.

Notations 1.3.6 *Let X_G be a random variable which has the distribution N(0,1) : $X_G \sim N(0,1)$. For all $b > 0$, we define $\Gamma(b)$ as being the probability that X_G does not belong to $[-b, b]$: $\Gamma(b) = P\{|X_G| \geq b\}$.*

Suppose that the probabilities of F(m) are chosen randomly with the uniform distribution. We set $X_n = T_q(Z_n)$. Then, we shall prove in chapter 8 that, for all Borel set Bo,

$$P\{X_n \in Bo\} = L(Bo)\Big[1 + \frac{2bOb(1)}{\sqrt{12L(Bo)m}}\Big] ,$$

with a probability larger than $1 - 2\Gamma(b)$.

Example 1.3.1 *Suppose $m \geq 2^{100}$, $q = 50$. Then, with a probability larger than $1 - \frac{2}{10^{340}}$,*

$$P\{X_n \in Bo\} = L(Bo)\Big[1 + \frac{2 * 40Ob(1)}{\sqrt{12 * 2^{50}}}\Big] = L(Bo)\Big[1 + \frac{Ob(1)}{2^{20}}\Big] .$$

As a matter of fact, we have a more important result : it means that, for almost ALL the possible random variables $Z_n \in F(m)$ such as Z_n is a model for a sequence of real number z_n ($z_n = Z_n(\omega)$), (even if Z_n is a bad model)

$$P\{X_n \in Bo\} = L(Bo)\Big[1 + \frac{Ob(1)}{2^{20}}\Big] .$$

This result is false only with a probability smaller than $2/10^{340}$. One can same decrease this probability by adding a pseudo-random sequence g_n : $X_n = T_q\big(\overline{Z_n + g_n}\big)$. All these results are proved in chapter 8.

1.4 Method of construction

We now describe a method of construction of an IID sequence using the properties which we have just described. We build it starting from certain electronic files. But this method can also be used with the random noises provided by the machines.

1.4.1 The method

One uses a sequence of data a(j) resulting from several independent files. These a(j) are thus independent by group. One can them consider like the realization of a stochastic process A_j : $a(j) = A(j)(\omega)$. One then builds an IID sequence of bits $b^0(n')$ in the following way (this construction is detailed in section 11.1).

a) One groups the data together in order to have a sequence of data $e^1(j) \in F^*(m^1) = \{0, 1, ..., m^1 - 1\}$ where m^1 belongs to the Fibonacci sequence.

b) One makes uniform the marginal distributions by using Fibonacci functions T_1^m.

b-1) One sets $e^2(j) = \overline{e^1(j) + rand_0(j)}$ where $rand_0(j) \in F^*(m^1)$ is a pseudo-random generator of period m^1.

b-2) One sets $e^3(j) = mT_1^m(e^2(j)/m^1) \in F^*(m) = \{0, 1, ..., m - 1\}$ where m belongs also to the Fibonacci sequence.

As a mater of fact, T_1^m make also independent the $e^3(j)$.

c) One uses the CLT (Central Limit Theorem)

c-1) One rewrites the sequence $e^3(j)$ in the form of table $\{f(i, n)\}$, i=1,...,S, with S=10 or 15.

One removes possibly some $e^3(j)$ in order to ensure independence between the lines: the lines will result from different files.

c-2) The columns are summoned : $g(n) = \sum_{i=1}^{S} f(i, n)$ for n = 1,...,N.

c-3) One uses these results modulo m: $h(n) = \overline{g(n)}^m$.

d) One applies the function of Fibonacci T_{q_0}.

d-1) One sets $x(n) = T_{q_0}(h(n)/m)$.

d-2) One rewrites x(n) in the form of bits b_s^n and one joins the b_s^n obtained to have the sequence of IID bits $b^0(n')$.

This construction is detailed in the chapter 11.

1.4.2 Clarification of the method

The previous construction now is clarified.

Step b : By making uniform the marginal distributions with the functions T_1^m, the result is true for ALL the logical models $A(j)$ associated with the sequence $a(j) : a(j) = A(j)(\omega)$ [1].

Step c : The use of the CLT enables the data to check the equation 1.1 with a K_0 not too big where

$$\Big| f_{n,j(.)}\big\{z|z_2, ..., z_p\big\} - f_{n,j(.)}\big\{z'|z_2, ..., z_p\big\}\Big| \leq K_0|z - z'| \ .$$

To check that, one will need to know the rate of convergence of H(n) and G(n) to their limit distributions (by CLT). This study is carried out in the chapter 5.

In fact the step c-3) corresponds to a second limit theorem much more efficient than the CLT that one will call XOR Limit Theorem (XORLT). In this case, the sums modulo m are not other than the XOR (Or Exclusive) generalized.

Step d : By using the rate of convergence one knows how to apply the fundamental theorem 1 with the function of Fibonacci. A good choice of the parameters will imply that the equation 1.2 is true not only for intervals I, but also for all Borel sets Bo :

$$P\big\{X(n) \in Bo\big|x_2, ..., x_p\big\} = L(Bo) + Ob(1)\epsilon \ ,$$

where ϵ is small enough.

If the parameters are well selected, the sequence $b^0(n')$ could be regarded as IID.

1.4.3 Choice of parameters

It is thus necessary to choose the parameters q_0, m and m^1. This choice is carried out according to the size $q_0 N$ of the sample $b^0(n')$ that we want to have.

Indeed, one will impose

$$2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{m}}{\sqrt{q_0 N}} \ ,$$

[1] To consider the models A(j) is equivalent to consider the models $E^3(j)$: cf properties of functions T_q in section 1.3

where $a_2 \approx 4^{-q_0}$ and where $\alpha \leq 0.02$. The reasons of these choices are detailed in section 2.4.3

1.5 Quality of the sequences $b^0(n')$

We will understand that the sequence $b^0(n')$ can not be differentiated from an IID sample.

1.5.1 Obtained relations

First, for all bit b,

$$P\{B^0(n') = b | b_2, ..., b_p\} = 1/2 + \frac{Ob(1)\alpha}{\sqrt{q_0 N}} \ ,$$

where $q_0 N$ means the size of the sample of the bits $b^0(n')$.

Now let us note by P_e the empirical probability that $B^0(n' + j_s) = b_s$ for s=1,2,..,p : $P_e = P_e\big\{\{B^0(n' + j_1) = b_1\} \cap ..... \cap \{B^0(n' + j_p) = b_p\}\big\}$, where $b = (b_1, ..., b_p) \in \{0, 1\}^p$. Then,

$$P\{\sqrt{Nq_0}|P_e - 1/2^p)| > \sigma_p b\} \leq \Gamma\big(b[1 - \eta]\big),$$

where σ_p^2 is the variance of P_e in the IID case and where η is rather close to 0.

Note by P_e^C the empirical conditional probability that $B^0(n') = b_1$ given $B^0(n' + j_s) = b_s$ for s=2,3,...,p. Then,

$$P\{\sqrt{Nq_0}|P_e^C - 1/2| > \sigma_p^C b\} \leq \Gamma\big(b[1 - \eta']\big),$$

where $(\sigma_p^C)^2$ is the variance of P_e^C in the IID case and where η' is rather close to 0.

It is important to notice that these results are also checked for all subsequence $b^0(\phi(n'))$ where ϕ is an injective function.

1.5.2 It is impossible to differentiate $b^0(n')$ from an IID sequence

Indeed, it is known that if there is sequence really IID, P_e is close to $1/2^p$ with a certain probability: it is completely possible that P_e is rather different from $1/2^p$, but the probability that happens is weak.

Now, under the assumptions of the theorem 1, it is also possible that P_e is rather different from $1/2^p$, but that does not have much more chances to happen than in case really IID.

It will be thus difficult to differentiate the $b^0(n')$ from a really IID sample.

1.5.3 Checking of the definitions of randomness

The sequences $b^0(n')$ thus checks a probabilistic relation similar to that of the definitions of a random sequence : cf section 1.1.3 and definition 1.1.2. This definition is known to be that which one expects from an IID sequence.

The sequences $b^0(n')$ check also empirical relations similar to those of the traditional definitions of the random sequences : cf sections 2.1.

These results are true for all logical model A(j) associated to the sequence of the data a(j). *They are thus true for ANY models* $B^0(n')$ deduced from the A(j)'s by the constructions defined in section 1.4.1.

By this way, we circumvented the problem of definition of random sequences of real numbers and even that of the choice of model.

1.5.4 A file of an IID sequence

We have concretely built an IID sequence of real $b^1(n')$ by using the method described here : cf section 11.2. One can obtain this sequence $b^1(n')$ by asking it to rene.blacher@imag.fr (Laboratory LJK, University Joseph Fourier of Grenoble, France). More precisions on this subject will found in [18].

1.5.5 Tests

Theoretically, some tests can not be checked by the sequences $b^0(n')$. But that will happen only with a probability equivalent to that of an IID sequence.

On the sequence $b^1(n')$, we carried out the traditional tests of Diehard : cf section 2.1.6. All were checked : cf section 11.2.14.

1.6 Conclusion

For all these reasons, it is not possible to differentiate the sequences $b^0(n')$ from an IID sequence : one can consider that $b^0(n')$ is an IID sequence.

There is thus well a total and simple answer to the problem of random numbers. In particular, this solution will have all its interest
1) For delicate calculations.
2) In cryptography: an IID sequence being by nature unbreakable.
3) In simulation, analysis, etc : one avoids having to test the provided sequence.

The advantages compared to the current methods are clear:

1) It was **proven** that the numbers obtained are random.

2) There is not to test these numbers, especially in simulation where it had to be done for each new practical application.

3) The method is applicable directly on the computers: it is as easy as to use a function "random". Moreover, there does not need to add a machine or an additional chip to the computer.

4) If one uses the random noises (Machines, chips, software programs), one removes all the dependence, which generally the current methods do not do. Moreover that can remove certain dysfunctions of the machines.

A more detailed comparison with the current methods is carried out in section 2.3.1.

1.7 Other possible transformations

There are other possible transformations of the data to obtain IID sequences. We detail this point in the Chapter 12. In section 12.1, we show another method explicitly.

Chapter 2

Quality of obtained sequences

In this chapter, we detail what is the interest of our method.

2.1 Criteria of randomness

2.1.1 Mathematical definitions

To determine the quality of a generator, one needs a definition of the randomness of a sequence of real numbers x_n. Many studies were made to give reasonable definitions: there is a good summary of these studies in chapter 3-5 of Knuth : cf [1]. In this section 2.1.1, we summarize the study of Knuth.

The common wish when one tries to obtain random sequences, it is to obtain a sequences of real numbers x_n which can be regarded as a sample of an IID sequence of random variables X_n : $x_n = X_n(\omega)$. Then, one can propose the following definition.

Definition 2.1.1 *: Let x_n, n=1,2,....,N, be a sequence of real numbers in [0,1]. Then, x_n is random if there exists an IID sequence of random variables $X_n \in [0,1]$ defined on a probability space (Ω, A, P) such that $x_n = X_n(\omega)$ where $\omega \in \Omega$.*

But there is a problem with this definition : for example, if it is admitted, x_n can be increasing. Of course, it is possible only with a negligible probability. But it is possible. Then, another definition thus should be used. Thus, Franklin proposed the following definition.

Definition 2.1.2 *: Let x_n, n=1,2,....,N, be a sequence of real numbers in [0,1]. Then, x_n is random if it has each property that is shared by all samples of an IID sequence of random variables from uniform distribution.*

This definition is not precise and one could even deduce from it that no really random sequence exists (cf [1], Knuth page 149).

One must thus define differently what is a random sequence (or IID sequence). Also, the following definitions were introduced.

Definition 2.1.3 *: For all finite sequence of intervals $I_s \subset [0,1]$, we denote by P_e the empirical probability : $P_e = (1/N_4) \sum_{n=1}^{N_4} 1_{I_1}(x_n) 1_{I_2}(x_{n+1}).....1_{I_p}(x_{n+p})$ where $N_4 = N - p$.*

The sequence $\{x_n\}$ is said p-distributed if $|P_e - L(I)| \leq N_4^{-1/2}$ for all $I = I_1 \otimes I_2 \otimes ... \otimes I_p$.

Definition 2.1.4 *The sequence x_n is random if it is p-distributed for all $p \leq Log_2(N_4)$.*

Unfortunately, this definition does not take into account the randomness of subsequences $x_{t_1}, x_{t_2}, ...x_{t_m}$. However, it is known that one cannot extend this definition to all the transformations $s \to t_s$ which define these subsequences : for example, this definition cannot be satisfied by the sequences x_{t_s} increasing. It is necessary thus that the application $s \to t_s$ is too not complicated. Also Knuth proposes the following definition.

Definition 2.1.5 *: The sequence x_n is random with respect to a set of algorithms A, if for all sequence $x_{t_1}, x_{t_2}, ...x_{t_m}$, determined by A, it is p-distributed for all $p \leq Log_2(N)$.*

Remark 2.1.1 *One imposes $p \leq Log_2(N)$ because that does not have any meaning to consider P_e if $p > Log_2(N)$, e.g., if p=5, and if one has a sample of size 10, that has not meaning to study its dependence in $32 = 2^5$ cubes of width 1/2.*

These definitions summarize those given by Knuth, [1] page 108. In fact he has especially studied the infinite case. But because in practice, there are always samples of finite size, we are limited to this case in this report.

This type of definition was the subject of many studies. In 1966, Knuth had thought that definition 3 defines the randomness perfectly: cf [1] page 163. It seems that he changed opinion since. In any case, none of these definitions is fully satisfactory. Knuth speaks philosophical debate on this subject. Thus, he points out that, according to certain principles, all the finite sequences can be regarded like determinist (cf pages 167-168) [1]).

Remark 2.1.2 *In any case, the above definition of Knuth is uncertain: can one choose algorithms which mix the past and the future? The algorithms should not they not be limited to mix some x_{n-t} where t is not too large and $t > 0$? It seems that not. But is this sure? Cf definitions 2.1.8 and 2.1.9. These remarks show that the problem is complicated.*

2.1.2 Statistical definitions

Now, the definitions above are not satisfactory statistically. Indeed, it is known that if x_n is really an IID sample, $\frac{N^{1/2}(P_e-L(I))}{\sigma}$ has approximatively the normal distribution where σ is the variance associated to P_e. For example, suppose p=1, $L(I) = 1/2$, $\sigma^2 = 1/4$. Then, $P\{N^{1/2}|P_e - 1/2| \geq 1\} \approx 0.045$. That means that it is possible that the event $|P_e - L(I)| > N^{-1/2}$ occurs. In fact, if one considers $|P_e - L(I)|$ for the set of intervals I, it are almost certain that it will occur. This property is thus different from the definition 2.1.5 de Knuth.

Therefore, one specifies statistically the definitions 2.1.4 and 2.1.5 in the following way.

Definition 2.1.6 *: Let us suppose that x_n belongs to the set of the numbers with q decimals bases d. One says that x_n is random if, for all the sequences x_{t_s} defined by a set of algorithms A, for all suitable p, for all $I = I_1 \otimes I_2 \otimes .. \otimes I_p$, (where the I_s are intervals of $\{0, 1/d^q, 2/d^q, ..., d^q/d^q\}$), it checks all the tests associated to $N^{1/2}|P_e - L(I)|$ with the same frequency that a really IID sequence would do it.*

By "same frequency", we understand that a really IID sequence will not check all the associated tests, but will check them only with a certain probability. For example, if all the tests to 1 percent was checked that would be abnormal.

This definition is completely natural. Indeed, what would make a priori an observer not knowing anything about x_n if he wants to know if this sequence is IID? He will carry out, in an certain order, a succession of tests which will define randomness completely.

Now, the definition is limited to the intervals. It does not include Borel sets and it is an important gap.

For example let us suppose that $x_n = \overline{\overline{0, d_n^1 d_n^2 d_n^3 d_n^4}}$ is the writing of x_n bases 10. Let us suppose that the definition 2.1.6 is checked only for the intervals. Then, because $\{d_n^1 = 1\} = \{x_n \in [0.1, 0.2[\}$, the event $\{d_n^1 = 1\}$ will probably check acceptable conditions. But $\{d_n^4 = 1\} = P\{x_n \in \cup I_s\}$, where $I_1 = [0.0001, 0.0002[$, $I_2 = [0.00011, 0.0012[$, $I_3 = [0.0021, 0.0022[$,..... It is thus easy to obtain examples where $d_n^4 = 1$. Of course, in this case, x_n could not be regarded like random. However, the definition 2.1.6 is checked

Therefore, one specifies statistically definitions 2.1.4 and 2.1.5 in the following way.

Definition 2.1.7 *: Let $P'_e = (1/N_4) \sum_{n=1}^{N_4} 1_{Bo_1}(x_n) 1_{Bo_2}(x_{n+1}).....1_{Bo_p}(x_{n+p})$ where the Bo_i's are Borel sets.*

It is said that x_n is random if, for all the sequences x_{t_s} defined by a set of algorithms A, for all $Bo = Bo_1 \otimes ..... \otimes Bo_p$, it checks all the tests associated to $N^{1/2}|P'_e - L(Bo)|$ with the same frequency that a really IID sequence would do it.

Now, it is known that one will always find Borel sets which does not check tests of randomness even for a really IID sequence.

This fact is not annoying: this case is envisaged by the use of the terms "with the same frequency ".

On the other hand, it is not obvious that one has forgets not any dependence in the previous definitions. Also, to avoid gaps, one introduces a new definition (generalization of the definition 1.1.2).

Definition 2.1.8 *: It is said that x_n is random if, for any n, for all Borel set Bo, for any injective sequence j_s, if one knows $x_{n+j_2}, x_{n+j_3}, ..., x_{n+j_p}$, one cannot predict the site of x_n with a probability very different from that of uniform distribution : $P_e\{x_n \in Bo | x_{n+j_2}, ..., x_{n+j_p}\} \approx L(Bo)$, where $P_e\{x_n \in Bo | x_{n+j_2}, ..., x_{n+j_p}\}$ is a good estimate of the conditional probability $P\{x_n \in Bo | x_{n+j_2}, ..., x_{n+j_p}\}$.*

Such a definition will pose problems, for example for the sequences x_{n+j_s} which are increasing. Then, it is noticed that one can simplify it: if all the conditional probabilities check $P\{x_n \in Bo | x_{n+j_2}, ..., x_{n+j_p}\} \neq L(Bo)$ for all n, it is equivalent to the fact that all the conditional probabilities knowing the past check $P\{x_{n+1} \in Bo | x_1, ..., x_n\} \neq L(Bo)$ for all n : it is enough to regard the values of probabilities of the $Bo_1 \otimes ..... \otimes Bo_p \subset [0,1]^p$. One can thus be satisfied to use the conditional probabilities knowing only the past.

Definition 2.1.9 *: It is said that x_n is random if, for all Borel set Bo, for all n+1, if the past $x_1, x_2, ..., x_n$ is given, one cannot predict the site of x_{n+1} with a probability very different from that of uniform distribution : $P_e\{x_{n+1} \in Bo | x_1, ..., x_n\} \approx L(Bo)$, where $P_e\{x_{n+1} \in Bo | x_1, ..., x_n\}$ is a good estimate of the conditional probability $P\{x_{n+1} \in Bo | x_1, ..., x_n\}$*

This definition seems a priori a good definition of the randomness. Indeed, it says that, knowing the past, one cannot predict the future with a probability too different from that of the uniform distribution. Intuitively, it is understood well that it is well the independence of the X_n which one defines thus.

Besides, it is this condition which one wishes for the random sequences in much books. However, in these books, one does not adopt this definition. Indeed, the definition 2.1.9 is imprecise : one does not have specified the approximation.

In fact, it is also the case in the definition 2.1.6 where one does not have specified the frequency. However, that will pose problems as for the definition of Franklin. It would thus be necessary to specify our definitions and to make a theoretical study. But it is not the goal of this report more especially because all the studies on this subject were delicate

Especially, that will not be necessary because we have to avoid this problem by using sequences which are really samples of random variables and by studying the properties that one has a priori on such sequences.

Of course, we have proposed the previous theoretical definitions especially to show that the results that we obtain in section 2.4.3 are well what is supposed and what one always has wished for random sequences.

2.1.3 Use of random Variable

Use of really random Variable

Then, we use really random variables. It is this technique what one uses with machines and the method of Von Neumann, Vazirani, and others ones : cf [4], [46], [8].

As a matter of fact, they assume that x_n is the realization of a sequence (not IID) X_n defined on a probability space (Ω, A, P) : $x_n = X_n(\omega)$ where $\omega \in \Omega$.

Then they transforms $\{X_n\}$ in a sequence $\{X_n^T\} = \mathcal{F}(\{X_n\})$ such that X_n^T is IID. But they obtain often this result under some assumptions whose one is not sure that they are checked.

Now, the matter of samples reappears : it is necessary to choose a correct model for X_n. It is known that some ones are bad. Now there exists an infinity of such models. How is it possible to choose the good one?

By using our method, we avoid this difficulty by proving that the sequence X_n^T which we obtain are IID for *all* the logical models X_n.

In particular these sequences X_n^T satisfy properties compatible to definitions 2.1.8 and 2.1.7 for ALL logical models. It is a very strong result which resolves the problem of definition.

Choice of a model : case of machines

We know that one cannot choose definition 2.1.1 as definition of randomness, that is to say the following definition is bad.

Definition 2.1.10 *(Definition 2.1.1) : We say that x_n is random if there exists an IID sequence of random variables $X_n \in [0, 1]$ such that $x_n = X_n(\omega)$.*

For example, x_n can be increasing (with a negligible probability).

In fact, this definition is not a problem solely when it is known a priori that the sequence X_n is IID: it is the type of assumption which one makes when x_n is provided by physical phenomena in the machines generating random numbers electronically.

Unfortunately, it is an assumption which is not completely sure. In fact, a priori this assumption does not hold for these machines: the intruments of measurements distort the physical phenomenon and induce bias and dependences (cf section 2.2.2, page 33). The reasoning which follows from there thus become partially dubious.

In fact, this definition does not pose any problem solely when one knows completely a priori the IID sequence X_n, like the case of a mechanical roulette or a mechanical lotto.

In this case, one starts from a machine and one extracts a sample from it.

But this technique is not thus appropriate inevitably when one starts from a real sequence y_n: cf counterexample of the increasing sequences.

Choice of a model : general case

To a sequence of real, it corresponds an infinity of models. Even if x_n can be regarded as a sample of an IID sequence X_n, it can be also logically regarded as the realization of an infinity of other models X_n^a (thus not IID).

The question thus should be asked: if one associates a model to a sequence x_n which criteria make it possible to be sure that this model is correct? It is a study which was never made until.

Then, which is the theory on the models? Generally, the following facts are admitted:

1) There never exists single model: a model is always related so that one wants to make of it. The same object will not be translated in a model in the same way according to in what one is interested.

2) Even when the goal is fixed, there are always several possible models, which all can be as valid the ones as the others.

Then how to be sure that a model is the good? That seems impossible.

Choice of a model in a group of models

Group of models One can admit that a set of models contains correct models.

For example, when all the x_n are different, one can admit that X_n has a differentiable density with a Lipschitz coefficient K_0 not too large. That can be increased thanks to the value $Min(x_{t(n+1)} - x_{t(n)})$ where t is the permutation checking $x_{t(n+1)} > x_{t(n)}$.

In fact this hypothesis is usually admitted by the statisticians especially those which use functional estimate.

Transformation in a group of models In order to avoid the problem of the definition of the random sequence, one can transform them : $\{X_n^T\} = \mathcal{F}(\{X_n\})$ as Von Neumann, Vazirani, and others ones do : cf [4], [46], [8]. But their transformations do not have all the good properties necessary.

It is necessary that these transformations impose sufficiently useful properties on the models. For example, one can use the CLT which will impose

for the majority of the models this assumption of continuity of the densities. But it is simpler to use the Fibonacci functions T_q which have extremely strong properties.

Fibonacci transformation If a sequence y_n has reasonable models in the group of models with continuous density, then, with these models it will check

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] ,$$

where I is an interval and N(I) be the number of $k/m \in I$.

More precisely, all the models with continuous density and associated Lipschitsz coefficient K_0 will check this equality. Therefore, one is well sure to have finally

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)[1 + \epsilon] ,$$

where ϵ is enough small.

Then, one is sure that $T_q(Y_n)$ could be regarded reasonably as an IID sequence.

By this way, one turns well the problem of the definition of an IID sequence.

The problem of other models

Then a question is asked : if a model is correct and does not belong to the models with K_0 rather small, is what it will produce the same properties? If it produces another one, it will be a conctradiction. There will be two possible logical conclusions.

But does this problem really exist there? a priori not! It seems absurd that a sequence can be at the same time regarded as IID and not IID. However, it is not completely obvious: there one finds the problem of the choice of the definitions.

Then, does this problem really exist there?

As a matter of fact, the question is: will these models be they correct? In a certain way, certainly not! But it is not so simple. It is in fact very difficult to answer this question without making a study which proves it.

The total answer : true for all the logical models

Now, by using the Fibonacci functions, one avoids the problem. In section 13.3, one proves mathematically that, *for almost all the models*, one has well

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] \ .$$

It is a very satisfactory result. Indeed, it is known well that if one takes all the possible models without no a priori, there will be an infinity of bad models. Here, we find of it only a number negligible: it is already extraordinary.

Moreover, there is a still better result. One indeed finds that for some data, for example those resulting from texts, ALL the logical models associated with y_n will check (cf chapter 13)

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] = L(I)\big[1 + \epsilon\big] \ .$$

One could better wish with difficulty like result.

Unfortunately, there remain the ϵ. But this problem can be resolved by choosing ϵ with respect the size of the sample.

2.1.4 An answer to the problem of definitions

Indeed, we will understand that the bits b_n that we regard as proven IID check the following properties :

for all $p \leq Log(N)/Log(2)$, for all injective sequence j_s, for all logical models B_n,

$$P\left\{\sqrt{N}\big|P_e - \frac{1}{2^p}\big| \geq \sigma_B x\right\} = K_1\Big(\big[1 - \eta\big]x\Big) \ ,$$

$$P\left\{\sqrt{N}\left|\frac{P_e}{p_e} - (1/2)\right| > \sigma_{cp}\ x\right\} = K_2\Big(\big[1 - \eta'\big]x\Big) \ ,$$

where η and η' are small enough and σ_B^2 and σ_{cp}^2 are the variances associated to P_e and P_e/p_e when $P_e = (1/n_1)\sum_{n=1}^{n_1} 1_b(B_{t_n})1_{b_2}(B_{t_{n+j_2}}).....1_{b_p}(B_{t_{n+jp}})$ and $p_e = (1/n_1)\sum_{n=1}^{n_1} 1_{b_2}(B_{t_{n+j_2}}).....1_{b_p}(B_{t_{n+jp}})$ where t is a permutation and $n_1 \leq N$.

Moreover, if B_n is IID, the same equalities holds with $\eta = \eta' = 0$. That means well that, if η and η' are small enough, one cannot differentiate B_n from an IID sequence. This result is proved in this report.

These results correspond well to definitions 2.1.7 and 2.1.8 which we think to be good mathematical definitions of randomness. As a matter of fact, theses properties are equivalent to these definitions. Then, by this way, one proves that B_n satisfies these definitions for all logical models B_n.

Thus, we solve the problem of mathematical definition and also the problem of the choice of the model.

One can be convinced of the certainty of this result when $\epsilon = \frac{1}{2^{50.000.000}}$ for a sample of size 25402545 bits : cf section 12.1.9. In this case, $\eta = O(\epsilon)$ and $\eta' = O(\epsilon)$. Indeed, on the current computers which produce results with a few tens of decimals, one cannot differentiate the numbers which have the same order that $\frac{1}{2^{50.000.000}}$.

2.1.5 Number of associated conditions

Concretely how can one check that a sequence x_n satisfies these criteria? If there is no special informations about x_n, the natural method consists in executing the tests associated with these definitions. However, the number of tests which must be carried out is colossal.

For example, let us suppose that one wants to test all independences between two numbers. First, it would be necessary to test that all the X_n and X_{n+p} are independent. There is thus approximately N independences to test if there is a sequence x_n of size N.

After, it will be necessary to test the independence of 3 numbers: X_n, X_{n+p}, X_{n+q}. For that, it is necessary to make approximately C_N^2 tests. For example, $C_N^2 \approx 5000000000$ if N= 100000. To test the independence of 4 numbers: $X_n, X_{n+p}, X_{n+q}, X_{n+t}$, it is necessary to make approximately C_N^3 tests, i.e. approximately 166665000000000 tests ..... etc; and so on until p numbers where $p \leq 16 \leq Log(N)/log(2)$.

Moreover, the set of all possible sequences defined by the algorithms A is much vaster than subsequences of the type: $X_n, X_{n+p}, X_{n+q}, X_{n+t}$ where one only considers some possible subsequences.

In addition, it will also be necessary to choose the Borel sets $Bo_1 \otimes Bo_2 \otimes ... \otimes Bo_p$, on which one will carry out the tests. Of course, there is a enormous number of possible Borel sets.

The number of tests which would have to be carried out is thus colossal. One cannot do them all. In practice, one thus needs less restrictive quality standards.

2.1.6 Other criteria of randomness

In order to practically know the quality of a generator, the majority of the authors use the natural method of the statistical tests. Considering the number of the tests to be carried out, there is an insurmountable gap between the mathematical definitions and the practical tests. In practice less restrictive criteria are thus used.

The most common criterion is the following.

Criteria 2.1.3 *: One forces x_n to pass a series of significant tests.*

Currently, one often uses the Diehard battery of tests (cf [2] p 155. In cryptography one adds tests like that of Maurer (cf [3]).

As these tests are only a negligible part of the possible tests, one fills this lack in the following way.

Criteria 2.1.4 *: One forces x_n to check certain essential mathematical properties.*

For example, one imposes on congruences that their period is long enough.

Now, even by imposing this new criterion, the generators will check always only a negligible part of necessary properties. Then, Knuth suggests that a sequence of numbers will be identified random if it cannot be differentiated from a IID sample by lack of time. Of course, such a criterion will be especially useful in cryptography.

More generally, all depends on the use which one wants to do of the generator. Thus in simulation, most of the time, it will be necessary to make tests to know if the generator is usable for the envisaged application : cf [2] page 151.

2.2 Current random generators

In this section, we study the current random generators.

2.2.1 Various current techniques

To obtain random numbers, various methods exist. Let us quote in particular,

1) Pseudo Random generator of type $x_n = f(x_{n-1}, ..., x_{n-q})$: for example,

1-A) Pseudo Random generator for simulation .
1-A-1) Congruences $x_n \equiv ax_{n-1} + c$ mod m
1-A-2) Add with carry (AWC) $x_n \equiv x_{n-j} + x_{n-k} + c_n$ mod (m)
1-A-3) Mersenne Twister.

1-B) Pseudo Random generator for cryptography.
1-B-1) BBS (Blum, Blum, Shub) generator based on $x_{n+1} = x_n^2$ mod m.
1-B-2) RSA generator
1-B-3) Feedback shift register

2) Irrational numbers : for example π and e. On use the writing base 10 of theses numbers.

3) Use of numerical data : for example music Rap for the Marsaglia Cd-Rom.

4) Hardware-based random bit generators : they exploit the randomness which occurs in some physical phenomena. They use machine or chips. Examples of such physical phenomena include :

4-1 elapsed time between emission of particles during radioactive decay

4-2 thermal noise from a semiconductor diode or resistor

4-3 the frequency instability of a free running oscilator

4-4 Quantum phenomena

5) Processes upon which software random bit generators may be based include

5-1 the system clock;

5-2 elapsed time between keystrokes or mouse movement;

5-3 content of input/output buffers;

6) Numerical tables, by example the CD-ROM of Marsaglia: in general, these tables are made with sequences of numbers obtained from the previous methods.

2.2.2 Study

Generators using algorithms

For these generators (pseudo random generators or irrational numbers), it is admited that it will never provide really random sequence : "Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin" : John Von Neumann (1951) (cf Knuth [1] page 1).

These generators will certainly not check the definitions of randomness given by Knuth. One cannot thus mathematically regard them as random sequences.

1-a) Generators used in cryptography.
These generators have certain mathematical properties sometimes very nice (e.g. the BBS). In particular, they are built to be unbreakable before a very long time. It is precisely the work of cryptanalysts to prove that it is false.

1-b) Generators used in simulation, analysis, etc.
One forces to them to check certain tests. The current generators appeared satisfactory in *certain* calculations, and certain calculations only. But contrary to the generator BBS, that is proven by empirical way and not by a mathematical way.

1-c) Generators using irrational numbers.

For "e", one has proved that it does even not satisfy the test of uniformity.

Concerning π, it satisfies the majority of tests. But one could nothing proven mathematically. Moreover, this sequence is rich with remarkable patterns to a numerologist (Knuth p 41, 151). As a matter of fact, the question is to know if the randomness provided by π is better than that of the other pseudo-random

generators.

Use of random noise

A true random bit generator requires a naturally occuring sources of randomness. Designing a hardware device or software program to exploit this randomness and produce a bit sequence that is free of biases and correlation is a difficult task. Moreover, random bit generators based on natural sources of randomness are subject to influence by external factors, and also to malfunctions. It is imperative that such devices be tested periodically (cf [3]). Thus, Marsaglia has found that some machines give numbers which have very poor quality : they fail for some Diehard tests. Moreover, it is impossible to reproduce calculations exactly a second time when cheking out a program.

Moreover

For the Hardware (Machines, chips) : these machines use some physical noise (diode noise, quantum noise,...). Machines have tended to suffer from malfunctions that are extremely difficult to detect.

For the Software-based generators : The behavior of such processes can vary considerably depending on various factors, such as the computer platform.

The major defect of all these systems is that there can be correlations and bias in the generated sequences. The underlying physical process can be random. But there are many measuring devices between the digital part of the computer and the physical device. These intruments can thus introduce bias and correlations (cf [5] ch 17.14 Bias and Correlation).

One removes these bias and these linear correlations by various mathematical transformations like that of Neuman or Vazirani ([4] [46]). One can also use hash functions (cf 17.14, [5]).

However, the linear correlation are only one of the possible correlation. There exists correlations of higher order (quadratic cubic, etc : Cf [10] or section A.3). If there are bits, the correlation of higher order are the multilinear correlation between 3,4,5, etc bits. If one does not remove the correlation of higher order, one will not have independence. However, the made transformations aim especially at removing only the correlation linear between two successive numbers (e.g [4]). Indeed, it is very difficult to remove all the correlations, especially the ones of high order.

Therefore, a priori the sequences of numbers built by the current methods to remove the correlation are not IID. It is a serious defect of the hardware device or software.

Tables of random numbers

One can obtain such tables by mechanical processes like the lotto or the roulette. They are the alone tables having results which are guaranteed IID.

But most of the time, these tables are obtained by the previous methods. They thus have the defects of them

A particular case is the CD-Rom of Marsaglia. Indeed, the random bits of this CD-ROM were made by combining music rap with sources of electronic white noise and the output from the best of the latest crop of deterministic random number generators, based on Marsaglia's "mulitply-with-carry" method. "They seem to pass all tests I have put to them – and I have some very stringent tests," Marsaglia says.

But, the randomness of the obtained sequence was not proved mathematically. In this report, we wanted to know logically if this sequence were random: cf chapter 3. This study shows that to have more certainty, it is necessary that these sequences are built by a certain way

That led us to take up the idea of Marsaglia: to regard certain electronic files as random noises, but to apply transformations a little more complex to them. That thus makes it possible to obtain true random numbers with a computer alone and to do without machines and chips. But, one can also use the numbers produced by machines

In any case, these tables have a major defect: they are limited by their size.

Conclusion

On none of the current generators there is certainty which the obtained sequences are random: that which approaches more this result is the Cd-Rom of Marsaglia.

However, much of users think that the provided generators are completely reliable and use them without precaution. All this already led to some scientific errors (cf [1] page 32). Thus, a more reliable solution should be obtained.

2.3 Advantages of our method

2.3.1 Comparison with the current methods

First, it is proved that the obtained numbers are really random. That had been obtained in no other method. Moreover,

A) Comparison with the pseudo-random generators

First, the usual opinion, it was that no generator built on computer is random: it is understood that it is an error. In fact as it thinks for example Von Neumann, the truth is probably that no generator built by algorithm is random.

Our method thus brings obvious concrete advantages. In particular, in cryptography, there is no risks that the system can be broken. In simulation, there is not to test the numbers obtained.

B) Comparison with generators based on natural sources of randomness
B-1) When one directly uses the program on a computer.

1) There does not need to add an additional machine to the computer.

2) There are no possible malfunctions as on the machines. Therefore, there is not to regularly test them like those.

3) The sequence obtained starting from the electronic files can be reproduced (it is useful for the checking of calculations).

B-2) When one uses the program on a source of random noises

1) That removes all the dependences, and maybe even certain effects of the malfunctions.

2) One can have very long sequences quickly (contrary to the methods using software).

C) Comparison with the CD-Rom of Marsaglia.

1) The results are proved.

2) The CD-ROM has a limited size.

3) A priori, it is possible that the sequence of the CD-Rom have defects.

2.3.2 Advantage for the definition of randomness

Our results bring even solutions to the problems of definition of randomness, including the philosophical problems: our results are true for *all* logical model associated to data.

2.4 Proof of randomness

The proofs of the randomness of sequences $b^0(n')$ are throughout this report. For better understanding these proofs, we summarize them here.

2.4.1 Randomness of used data

Our results are based on the fact that the data $a(j)$ which we use can be regarded as a realization of a sequence of random variables $A(j) : a(j) = A(j)(\omega)$. It is thus appropriate to show that one can admit this assumption with certainty and then to study the models A(j).

Let us announce that one studied this problem in a more complete way in section 2.4.1 and chapter 10 .

The data mean really random phenomena

In this paragraph, we show that the assumption $a(j) = A(j)(\omega)$ is a sure assumption.

First, it is a scientist assumption universally admited when the a_j represents a physical phenomenon.

Moreover, even for a completely deterministic sequence, one can always write $a(j) = A(j)(\omega)$: in this case the model A(j) is deterministic. Moreover, one can admit for this sequence A(j) various other realistic and nondeterministic models.

But what interests us is to have a logical nondeterministic model. It is the case for the real phenomena associated with hardware device or software programs.

It is also the case when a_j is resulting from certain electronic files. First, these files often represent real phenomena. Then, one can even prove logically that certain files result from nondeterministic models. It is the case for the numbers obtained starting from a dictionary: the defined words (or groups of words) represent independent facts. The numbers obtained from these definitions are thus independent.

If one wants to go into detail this question, one finds complex philosophical problems. But in any case, some phenomena are independent from each other. Therefore, one has to note that the philosophies which would refuse the assumption $a_j = A_j(\omega)$ would not be in conformity with what exists in reality. It appears normal to take not account of such philosophies. For more detailed study, one can refer to the section 2.4.1 and to chapter 10.

For all these reasons, we thus adopt the assumption $a(j) = A(j)(\omega)$.

Asymptotic independance

In fact most of the time, one can admit that there is asymptotic independence.

With regard to the machines, that is due to the choice of the used real phenomena which one wants independent as much as possible, for example quantum phenomena.

With regard to the electronic files, that is due to their nature. The case of the dictionaries was understood above. One can also take the encyclopaedias. In this case, there is independence by groups, and thus Qd-dependence (cf [21] page 369).

Moreover, one often uses several electronic files completely independent from each other, for example, a dictionary and a computer program. The sequences of numbers which they provide are thus independent.

More generally, for logical reasons one can be sure that the data of certain electronic files have a certain asymptotic independence: e.g. the data obtained from text (cf section 2.4.1 et chapter 10).

Thus for certain electronic files, asymptotic independence is a logical assumption. For the machines, so that it does not have asymptotic independence there, one would need a serious malfunction.

General case

Our results do not need asymptotic independence to be checked. It is enough for a vast majority of possible models that the A(j)'s are not determinist. But in this case, it is difficult for the moment to specify which are the models which are not appropriate.

On the other hand, if there is asymptotic independence, one can prove that, under this assumption, the sequence $b^0(n')$ is IID. It is thus surer to be under this assumption. Thus for certain electronic files, one can prove by logical reasoning that it is well the case.

Let us repeat that, for the machines, there would be a serious malfunction if there is not asymptotic independence. Now, if the numbers provided by the machines were deterministic, it would be much more serious. Because of that, our method applied to machines makes possible to remove many dysfunctions. It is enough in general that they continue to produce nondeterministic numbers.

2.4.2 Mathematical proofs

In this section, we want to show in a simple way why we are sure that the sequence $b^0(n')$ is IID.

Because most of the time one can admit that there is asymptotic independence, we place oneself under this assumption. For a more general case, it is necessary to refer to the various chapters of this report.

When the assumption of asymptotic independence is admitted, there are two transformations which it is practical to apply under this hypothesis : the Central Limit Theorem (CLT) and the congruence of Fibonaci. We have them uses both in the transformations defined into section 1.4.

We will now understand the effects of these transformations. We thus study the steps b, c, d, defined above in 1.4.1 .

b) To make uniform the marginal distributions.

b-1) We set $e^2(j) = \overline{e^1(j) + rand_0(j)} \in F^*(m^1)$ for j=1,2,.....,J where $J/m^1 \approx 0$.

A priori, because of this transformation, one can admit that $P\{E^2(j) = x\}$ cannot be negligible, but equal about 1/m. One obtains these conclusions by logical reasoning which we will detail in chapter 10.

Remark 2.4.1 *Let us notice that it is supposed only that the model $E^2(j)$ has marginal distributions which are not concentrated nearly a small number of points : it is an assumption much weaker. One is thus sure logically that one chooses assumptions which are correct. Also let us notice that it is with this alone transformation that Marsaglia had concluded that the sequence of its CD-ROM is IID. We thus choose a conclusion much weaker.*

Now, let us suppose for example that the data $e^1(j)$ result from texts. If one did not add $rand_0(j)$, one would know logically, a priori, that for some "x", $P\{E^1(j) = x\} = 0$. But to know it starting from the sequences $e^1(j)$, it would

be necessary that one realized that the $e^1(j)$'s results from an English text : it is doubtful that one can obtain this conclusion if one has only the sequence $e^1(j)$. This shows that, maybe one can do without adding $rand_0(j)$ (cf chapter 8.1, sections 8.3 and 8.2), i.e. that, because one adds $rand_0(j)$, the result is very sure.

Now one applies mathematical results.

b-2) One applies the Fibonacci functions T_1^m : $e^3(j) = mT_1^m(e^2(j)/m^1)$.

We proved in the chapter 8.1 that $P\{E^3(j) = x\} \approx 1/m^1$ for all logical models, except maybe for a negligible number of them. Moreover, we proved also that this applications makes the $E^3(j)$ independent.

c) Use of limit theorems.

c-1) One rewrites the $e^3(j)$'s in the form of table with independent lines f(i,n).

c-2) Lines are summoned : $g(n) = \sum_{i=1}^{S} f(i,n)$.

By using the very traditional Central Limit Theorem, one knows, that for any injective sequence j_s such that $j_1 = 0$, for any interval I,

$$P\{G(n)/m \in I \mid G(n+j_s) = g_s, s = 2, ..., p\} \approx P\{X_{G_\sigma} \in I\} \ ,$$

where $X_{G_\sigma} \sim N(0, \sigma^2)$. It is also known that convergence is extremely fast

c-3) One takes this sequence modulo m : $h(n) = \overline{g(n)}^m$.

As a matter of fact, that amounts considering a second limit theorem (the XOR Limit Theorem, XORLT): $h(n) = \overline{\sum_i f(i,n)}$. This sequence satisfies

$$P\big\{H(j)/m \in I \ \big| \ H(j+j_s) = h_s, s = 2, ..., p\big\} \approx P\{H(j)/m \in I\} \approx L(I) \ .$$

This result is written also $P\big\{H(j)/m \in I \ \big| \ H(j+j_s)/m = h_s, s = 2, ..., p\big\} = L(I) + Ob(1)\epsilon'$.

This limit theorem corresponds to famous OR exclusive (XOR) modulo m. In fact, it is much stronger than the Central Limit Theorem. We study it in section 5.2.

d) One uses again the Fibonacci function.

d-1) We set $x(n) = T_{q_0}(h(n)/m)$.

One applies the functions of Fibonacci T_{q_0} with a suitably chosen parameter q : the smaller q is, the less there are Borel sets. Therefore, by choosing the parameters well, and because of the above result in c-3),

$$P\big\{X(n) \in Bo \ \big| \ X(n+j_s) = x_s, s = 2, ..., p\big\} = L(Bo) + \frac{Ob(1)\alpha}{\sqrt{Nq_0}} \ ,$$

for all Borel set Bo. We shall prove this result in chapter 11.

d-2) One obtains the sequence of bits $b^0(n')$, $n' = 1, ..., q_0N$, formed by the writing bases 2 of each x(n). Because of the previous property,

$$P\big\{B^0(n') = b \mid B^0(n' + j_s) = b_s, s = 2, ..., p\big\} = 1/2 + \frac{Ob(1)\alpha}{\sqrt{Nq_0}} \ .$$

Finally all is proven mathematically or logically.

2.4.3 Choice of the parameters

It is thus necessary to choose the parameters q_0, m and m^1. This choice is carried out according to the size q_0N of sample $b^0(n')$ that we want to have and according to the quality that we wish, i.e. according to ϵ.

In practice, one chooses ϵ according to q_0N, the sample size $b^0(n')$. Indeed, one will impose and

$$2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{m}}{\sqrt{q_0N}} \ ,$$

where $a_2 \approx 4^{-q_0}$ and where $\alpha \leq 0.02$.

Choice of the parameters according to the sample size

One thus chooses the parameters according to the sample size. In this paragraph, we will clarify this point.

Let us suppose that we have a really IID sequence with uniform distribution on [0,1/2] and]1/2,0] and with a probability such as $P\{[0, 1/2]\} = 0, 501$. Then, this sequence has not the uniform distribution on [0,1]. However, if we have a sample with size 10, we will absoluetely not understand it. To understand this difference, one will need samples with size larger than 1000.

One will thus solve the problem of the choice of ϵ in the same way: according to q_0N, the wished size of the sample, one will choose ϵ and thus T^*. Let us translate that mathematically.

Let us note by P_e the empirical probability of an interval I associated with a sequence $x_n^* = X_n^*(\omega)$, n=1,2,....,N. Then, if X_n^* is a sequence of IID random variables with uniform distribution, if N is big,

$$P\{N^{1/2}|P_e - L(I)| > \sigma b\} \approx \Gamma(b) \ ,$$

where $\sigma^2 = L(I)[1 - L(I)]$.

Now, if X_n^* checks only equation 1.2 , i.e. $P\{X_n^* \in I | x_2^*, ..., x_p^*\} = L(I) + Ob(1)\epsilon$, one can prove that

$$P\big\{N^{1/2}|P_e - L(I)| > \sigma b\big\} \leq \Gamma\big\{b[1 - \eta(\epsilon)]\big\},$$

where $\eta(\epsilon) \geq 0$ and $\eta(\epsilon) \to 0$ as $\epsilon \to 0$.

For example, let us suppose that we built T^* so that $\eta(\epsilon) = 0.1$. In this case, for b=1,5

$$P\{N^{1/2}|P_e - L(I)| > \sigma.1.5\} \leq 0,134 \ \ under\ IID\ hypothsesis,$$

$$P\{N^{1/2}|P_e - L(I)| > \sigma.1.5\} \leq 0,148 \ \ under\ hypothesis\ of\ equation\ 1.2.$$

However, it is known that if there is a really IID sequence, P_e is close to L(I) with a certain probability: it is completely possible that P_e is enough different from L(I), but the probability that occurs is weak.

Now, under the assumptions of equation 1.2, it is also possible that P_e is enough different from L(I), but that is not likely much more to occur than in really IID case.

With such a result, it will be thus difficult to differentiate the x_n^* from a really IID sample.

Of course, if it is necessary, one can impose $\eta(\epsilon)$ smaller : for example, $\eta(\epsilon) = 0.01$. In this case,

$$P\{N^{1/2}|P_e - L(I)| > \sigma.1.5\} \leq 0,135 \ \ under\ hypothesis\ of\ equation\ 1.2.$$

This type of result holds again for $I_1 \otimes .... \otimes I_p$ where the I_i's are intervals. Moreover, one obtains a similar result for the empirical conditional probability $P_e^C = P_e\{x_n^* \in I | x_2^*, ..., x_p^*\}$:

$$P\big\{N^{1/2}\big|P_e^C - L(I)\big| > \sigma_p^C b\big\} \leq \Gamma\big(b[1 - \eta'(\epsilon)]\big),$$

where $\eta'(\epsilon) \to 0$ as $\epsilon \to 0$. This results are proved in chapter 9.

Finally if it is wanted that one cannot differentiate the sequence x_n^* from an IID sample on *intervals*, $\epsilon = O(N^{-1/2})$ should be chosen.

Case of Borel sets

Then, in order to obtain the previous results, one uses the theorem 1 applied to T^* : we deduce $P\{X_n^* \in I | x_1^*, ..., x_p^*\} = L(I) + ob(1)\epsilon$ for the intervals I.

Now, we need an equivalent result for the Borel sets Bo. For that, one uses $X_n = Pr_q^2(X_n^*)$: i.e one restricts X_n^* to his q first bits (cf def 1.3.5). Then, for all Borel set Bo,

$$P\{X_n \in Bo | x_2, ..., x_p\} = L(Bo) + Ob(1)2^q\epsilon \ .$$

It is thus enough to choose q not too large and ϵ enough small so that $2^q\epsilon$ is also enough small. In fact, it is necessary to choose $2^q\epsilon = O((qN)^{-1/2})$.

It is what one will make to define the q_0 of the application T_{q_0} at step d. Finally, with the method of construction defined in section 1.4.1, one obtains for all borel set Bo

$$P\{X(n) \in Bo | x_2, ..., x_p\} = L(Bo) + \frac{Ob(1)\alpha}{\sqrt{q_0 N}} \ , \tag{2.1}$$

where $\alpha \leq 0.02$ et ou $q_0 N$ is the size of sequence $b^0(n')$.

Then, one deduced the relations about the empirical probabilities of Borel set similar to those with the intervals.

Relations about $B^0(n')$

The previous results being true for all Borel sets, one deduced equivalents results about the bits $b^0(n')$ provided by the writing of x(n) bases 2.

One deduces from equation 2.1 that, for all bits b,

$$P\{B^0(n') = b | b_2, ..., b_p\} = 1/2 + \frac{Ob(1)\alpha}{\sqrt{q_0 N}} \ ,$$

where $q_0 N$ means the size of the sequence $b^0(n')$.

2.4.4 Empirical proofs of the randomness of $b^0(n')$

Result d-2 ensures that the associated empirical probabilities could not be differentiated from those of an IID sequence.

Indeed, with the notations of section 2.4.3, one understands that empirical probabilities P_e and P_e^C will check the equations

$$P\{N^{1/2}|P_e - 1/2^p)| > \sigma_p x\} \leq \Gamma\big(x[1-\eta]\big),$$

$$P\{N^{1/2}|P_e^C - 1/2)| > \sigma_p^C x\} \leq \Gamma\big(x[1-\eta]\big),$$

where η is enough close to 0.

Moreover, a numerical study shows that one cannot differentiate the empirical probabilities associated with the $b^0(n')$ from those of an IID sequence.

2.5 Precise wording of the result

2.5.1 Wording

It is thus proven that, for the model $B^0(n')$ built from ANY models of the data a(j) (except maybe for a negligible minority, according to the case), that the sequence $B^0(n')$ cannot be differentiated from a sequence of IID random variables.

In particular, it satisfies the properties

$$P\{B^0(n) = b \mid B^0(n + j_s) = b_s\} = 1/2 + \frac{Ob(1)\alpha}{\sqrt{N q_0}} \ ,$$

$$P\{N^{1/2}|P_e^C - 1/2)| > \sigma_p^C x\} \leq \Gamma\big(x[1-\eta]\big) \ ,$$

which correspond theoretically and empirically to the definition 2.1.9 of the randomness.

It satisfies also

$$P\{N^{1/2}|P_e - 1/2^p)| > \sigma_p x\} \leq \Gamma\big(x[1-\eta]\big) \ ,$$

which corresponds empirically to the definition 2.1.7 of the randomness.

That means that the sequence $b^0(n')$ cannot be differentiated from a sample of IID random variables.

2.5.2 Alone risks of errors

In fact, the alone risk that a sequence $b^0(n')$ built by our method is not random, it is first that an human error slipped in some place. But this risk can be eliminated by checking carefully each step of calculation and of the study of data: contrary to the machines, if there is an error, one can detect it keenly.

The second risk is that computer itself have a failure. This risk cannot never be eliminated completely. Anyway, in this case, any calculation is likely to be false.

2.6 Uses of these results

2.6.1 Direct programming on computer

A great number of the electronic files recorded on hard disk provide sequences of data a(j) which one can use to generate random numbers.

There is thus a simple program which provides numbers guaranteed random without machine or additional chip. It is as simple to use as the function "random".

It is thus a method quite superior to the current generators.

2.6.2 Application to hardware devices

One can choose like data those provided by machines or chips. One then applies our transformations defined into section 1.4. That offers several advantages.

On the one hand any dependence is removed, (and not only linear correlations).

On the other hand, some of their defects could disappear. Indeed, our method can be applied as soon as there is a certain asymptotic independence. In

fact it is not even necessary. It is often enough that the data are not completely deterministic.

It is certainly the case for the data provided by the machines even if they have malfunctions. If not, the machine has a very serious problem.

It is thus a new method which one proposes to transform the noises provided by these machines. The advantage, it is that it needs extremely weak assumptions to be applied. It is thus much surer. However one produces a little less quickly random numbers.

2.6.3 Application to software methods

One can choose as data those provided by the software methods if they have sequences of enough large size.

But normally, it will not be useful. If the files of the computers are used, our results are much surer than those provided by the software methods. Moreover it is simpler to use texts than the system clock for example.

2.6.4 Use of files of IID sequences

By using the method described in this report, one can develops files of numbers which are proved IID.

They could thus be placed for public use in the form of files to download, of files recorded on hard disk, of DVD or of CD-Rom as it is the case for the CD-Rom of Marsaglia (cf Internet site [20]) or for the file which we built according to our method.

2.6.5 Transformations of $b^0(n')$

From sequences $b^0(n')$ which are proved random, one can obtain a multitude of others by using any sequences y_n provided by generators which are pseudo-random or even different. Indeed, $\overline{b^0(n') + y_{n'}}$ modulo 2 is also IID (cf theorem 5).

2.6.6 Software for data external to the computer

One can build softwares allowing to transform the majority of data external to the computer in random numbers. They will apply with safety only to some types of data, for example, texts.

2.6.7 Complete construction

It is the matter to completely use the method of programming defined in this report with new data and choice of new parameters.

This method can be used when one wants, for various reasons, to obtain new sequences x_n completely reliable.

For that, it is initially necessary to collect the data, to transform them in numbers, to check that the necessary assumptions are satisfied and that there are no errors: for example, the transformation of the text files Apple Work 2006 currently transforms, not only the texts, but also the framework. That revealed a too great number of 0, which distorts calculations.

Then, it is necessary to choose the parameters according to the wished results, after to program the transformations and to carry out this program.

A certain work thus should be done. It is the normal price to pay to have the certainty that one wishes about the randomness of the obtained numbers.

2.6.8 Combination of several methods

If one wants to avoid any risk of human error, of machine's error, of computer's error or others, one can build several sequences $b^0(n')$ as described above in section 2.6.7.

One will thus build them with different data. In this case, one can also use machines, even different machines. One can even employ the files of random numbers which exist over the world.

Then one will summon modulo the 2 various sequences of obtained bits.

That will reduce the probability of any potential error, human or different.

Indeed, if one summons modulo 2 : $b_n = \overline{\sum_{s=1}^{I} b_n^s}$, it is enough that only one sequence b_n^s is a random sequence, so that b_n is it.

That means that the probability that the sequence b_n is not IID is the product of the probabilities that each sequence b_n^s is not IID. One finds very quickly very negligible probabilities of failure (and including even the human errors and the dysfunctions of the computers).

One can employ this technique if it is though that it is worthwhile, for example if one want to build a rocket being worth some billion dollars.

Remark 2.6.1 *As a matter of fact, we do in this way by using several times methods which make a sequence IID (Fibonacci transform, XORLT, transformations similar to permutations : cf chapter 11 and 12).*

2.7 Conclusion

One thus obtains a sequence $b^0(n')$ which can be considered like an IID sample.

However, there is a probability that it is not it. But this one has the same order to occur as a really IID sample fails for a test of randomness.

One thus obtained well a sequence which one will be able regarded as IID with a probability infinitely close to 1.

One can obtain this sequence $b^1(n')$ by asking it to rene.blacher@imag.fr (Laboratory LJK, University Joseph Fourier of Grenoble, France) [1]. Of course, by prudence, we have carries out on this sequence all the known tests of randomness.

There is thus well thus a total solution to the problem of random numbers. It is simple to use. This solution will have all its interest
1) For sensitive calculations.
2) In cryptography : an IID sequence being inherently unbreakable.
3) In simulation, analysis, etc, by avoiding having to test the provided sequence.

[1] More precisions on this subject will found in [18].

Chapter 3

Cd-Rom of Marsaglia

In this chapter, one will study the method which Marsaglia employed to create its CD-ROM. As a matter of fact, Marsaglia never described its method. It is only an assumption which one finds in various documents about this subject.

The Marsaglia Random Number CD-ROM contains some 5 billion random bits, divided into sixty 10-megabyte files.

The random bits were made by combining three sources of electronic white noise with the output from the best of the latest crop of deterministic random number generators, based on Marsaglia's "mulitply-with-carry" method. "They seem to pass all tests I have put to them – and I have some very stringent tests," Marsaglia says.

A truly random stream of bits remains random when it's combined with any other stream of bits, no matter how patterned. Marsaglia mixed digital tracks from rap and classical music selections and even a few digital pictures into some of the 10-megabyte files on the CD-ROM. Both the untouched and mixed files seem to pass his randomness tests, he reports.

Then, Marsaglia has studied his CD-Rom by using tests. In this chapter one will study this method by logical reasoning.

One should thus study the summation modulo m of three generators $my'_n = \overline{g_n + my_n + mz_n} \in F^*(m)$ where $y_n \in F(m)$ and $z_n \in F(m)$ are random sequence (non IID) and where $g_n \in F^*(m)$ is a pseudo random generator. But we will be satisfied to study the case $my'_n = \overline{g_n + my_n} \in F^*(m)$.

One will note that it is likely to obtain IID sequences by this method if the parameters well are chosen. Unfortunately there is no complete certainty.

In this chapter, one supposes that the y_n's derive from a text, e.g dictionary, report, etc. One employs the data which we use in the section 11.2. We studied their behavior in the chapter 10 .

Remark 3.0.1 *Marsaglia has not used texts but Rap music. It is no important. In this chapter we want only to study logically the method of Marsaglia. Then, we use texts because we studied them in a detailed way.*

3.1 Theoretical study

We will understand now that the behavior of the Y'_n depends on the way in which one builds the y_n.

3.1.1 Counterexample

First, let us notice that in some cases, to add g_n and my_n is not enough to obtain an IID sequence.

Example 3.1.1 *Suppose that m=10000 and that* $my_n \in \{1000, 2000, ..., 9000\}$.

Study It is not difficult to understand that there will be a dependence between the y'_n: $P\{Y'_n \in I \mid y'_2, ., y'_p\} \neq L(I)$ for certain intervals I. Indeed, let $\mathcal{R}(g_n) = \overline{\overline{d_2 d_3 d_4}}$ where $\overline{\overline{d_1 d_2 d_3 d_4}}$ is the writing base 10 of g_n. Then, $\mathcal{R}(g_n)$ is a pseudo-random generator.

In particular, it is clear that, for each n, the y'_n will be concentrated nearly a small number of points. That will mean that one will have points k such as $P_{Y'_n}\{y'_n = k\}$ is much larger than 1/m (where $P_{Y'_n}$ is the probability associated with the sequence Y'_n). ■

3.1.2 Case of 2-dependence

In section 11.2, we understand that the data can be regarded as 2 dependent. Then, we suppose now that the sequence y_n is 2-dependent.

First, study the behavior of $(y'_{n+j'_1}, y'_{n+j'_2}, ..., y'_{n+j'_p})$ when $0 = j'_1 < j'_2 < .... < j'_p$.

Clearly, if $j'_{s_0+1} - j'_{s_0} \geq 2$,

$$(y'_{n+j'_1}, y'_{n+j'_2}, ..., y'_{n+j'_p}) = \big((y'_{n+j'_1}, y'_{n+j'_2}, ..., y'_{n+j'_{s_0}})(y'_{n+j'_{s_0+1}}, ..., y'_{n+j'_p})\big),$$

where $(y_{n+j'_1}, y_{n+j'_2}, ..., y_{n+j'_{s_0}})$ and $(y_{n+j'_{s_0+1}}, ..., y_{n+j'_p})$ are independent (because the 2-dependence).

Therefore, to study the dependence of y'_n, it is sufficient to study the case $j'_s = s - 1$, i.e. the $(y'_n, y'_{n+1}, ..., y'_{n+p-1})$.

Now, one knows it is more logical to limit oneself to study the cases where $p \leq log(N)/log(2)$: cf remark 2.1.1. For example suppose $p \leq 22$. In this case, if the g_n's produce sequences where $(g_n, g_{n+1}, ..., g_{n+22})$ are independent, the $(y'_n, y'_{n+1}, ..., y'_{n+22})$ are independent. Therefore, one can consider that $(y'_{n+j'_1}, y'_{n+j'_2}, ..., y'_{n+j'_p})$ are also independent.

Then, it is enough to choose pseudo-random generators such that $(g_n, g_{n+1}, .$ $...., g_{n+22})$ are independent. In this case, to suppose that the y'_n are independent will be a reasonable assumption.

As it is considered that the generators generally have uniform marginal distributions, to suppose that the y'_n are IID will be also a reasonable assumption.

Therefore the method of Marsaglia can be sufficient to obtain IID sequences if the parameters have to be suitably chosen. It is the same for the the pseudo-generator g_n and for the type of data : if one uses texts, one can study it by logical reasoning (cf chapter 10). But is it possible for Rap Music?

However it remains to be checked that the marginal distributions are quite uniform: the tests of uniformity of the g_n means that some tests are checked, for example for intervals. But is this case for all Borel sets? It is similar for independences: they are independences for some hypercubes of the g_n: what is it for the others?

3.1.3 Transformation of datas

Now, one can use the data directly. One can also transform them. It is what we do for the transformation of the sequence c(j) defined in section 11.1.2 during the construction of the sequences of random bits $b^1(n')$. We now remind the definition of these d(j).

Example 3.1.2 *Let $c(j) \in \{0, 1, ..., 31\}$ be a sequence of data. Let $r_0 \in \mathbb{N}^*$. We set $d(j) = \sum_{r=1}^{r_0} c(r_0(j-1)+r)32^{r-1}$.*

Then, we understand now that the behavior of y'_n depend on the choice of transformation and on parameters

Size of r_0 and conditional dependence

One will now understand that one can choose r_0 large in order to decrease the dependences.

For that, always let us choose data resulting from texts. If one finds a ". ", it there a strong probability so that it is followed by a "space character".

Therefore, it is possible that it has there some strong dependences between c(j) and c(j+e) (where c(j) are the letters modulo 32) especially for e=1. But this dependence decreases very quickly if e increases.

That will mean that the possible concentrations of d(j+1) given $d(j) = \sum_{r=1}^{r_0} c(r_0(j-1)+r)32^{r-1}$ will be less strong if r_0 increases. In practice, it will be found that the associated probabilities $P\{D(j+1) = d | D(j) = d_1\}$ are always much smaller (and by far) than $32^{r_0/8}$ as soon as r_0 is enough large : for example $r_0 \geq 20$ as the study of the conditional probabilities or the numerical calculations cited in this report proves it.

Let us suppose, for example, $r_0 = 1$. Thus D(j)=C(j). Let us suppose that $d(j-1)$ means a "." . Then, there is much chance that $d(j) = d_{Es}$ where d_{Es} is

the numerical value of the "space character". The conditional probability will be thus concentrated close to the point d_{Es}.

Let us suppose now that $r_0 = 60$ and that $d(j-1)$ means a piece of text ending in a "." . Then, d(j) belongs to the set of the part of texts starting with a "space character". There is 32^{59} such sets. Thus the conditional probability is about $1/32^{59}$ for these points and equal to 0 for the others. As the uniformity would be of $1/32^{60}$ for all the points, it is understood that it is different - but not too - from conditions of an IID sequence.

Indeed, there is not much relationship between the last decimals of D(j) and the first of D(j-1). That means that the associated distribution functions are not too far away from independence. A fortiori, that means also that conditional distributions of $\overline{D(j-1)+g'(j-1)}$ given $\overline{D(j)+g'(j)} = d'_j$ - where $g' \in F^*(32^{r_0})$ is a pseudo-random generator - are not concentrated nearly a small number of points but are well distributed on $F^*(32^{r_0})$.

The previous result is checked numerically : in section 3.2.4, one studies the dependence and the conditional density on samples of y_n and y'_n .

3.1.4 Independence induced by the data

Independence of times of emergence of the "."

We use again the example of ".". We note by p_o their numerical value. Let $z_t \in \{1, 2, ..., m\}$ be the value of successive "n" such as $y_n = p_o$, i.e. $y_{z_t} = p_o$. This sequence z_t is random : one can write $z_t = Z_t(\omega_5)$ where Z_t is a sequence of variable increasing in a random way, defined on a probability space $(\Omega_5, \mathcal{A}_5, Proba_5)$. Then, in order to obtain $my'_n = \overline{g_n + my_n}$, we add g_{z_t} to $y_{z_t} = p_o$.

In practice, we understand that $Z_{t+1} - Z_t$ is close to an IID sequence (not necessarily with uniform distribution). It is enough to make some numerical simulation to realise that.

Test about the z_t In order to check the independence of the z_t, one made the chi-squared test for the $(z_{t+p+1} - z_{t+p}, z_{t+1} - z_t)$ on various text file for different p.

One used sample where the number of points p_o is between 584 and 2415. One makes the chi-squared test of independence with estimate of parameters over partitions (6,6), i.e. with the chi-squared statistics χ_2 with 25 degrees of freedom.

Then, one has obtainded for theses samples of $(z_{t+2} - z_{t+1}, z_{t+1} - z_t)$ the following chi-squared statistics :

χ_2	25.354	17.901	39.102	31.012	24.980	42.557

One reminds that the 0.01 significance level at 1 percent (with 25 degrees of freedom) is 44.31. Then one can deduce that the distribution is close to independence.

Conclusion This result means that the $z_{t+1} - z_t$ has a behavior close to an independent sequence.

Therefore, this result means that $my'_{z_t} = \overline{g_{z_t} + my_{z_t}} = \overline{g_{z_t} + mp_o}$, has a behavior close to an IID sequence because g_{z_t} can be regarded as chosen randomly.

Then, it is not possible to predict mY'_n given $Y_n = p_o$. Therefore, it is not possible to predict mY'_n given $Y'_n = y'_n$ thanks to $y_n = p_o$.

The same reasoning holds for other value of y_n. Therefore mY'_n has a behavior close to an IID sequence.

3.1.5 Study of $y_n = \lfloor d(n)(m/32^{r_0}) \rfloor$

In section, 11.1.2, one uses the sequences y'_n when $y_n = \lfloor d(n)(m/32^{r_0}) \rfloor$ where $m \in \mathbb{N}^*$. Then, we want to know if this transformation brings sufficient improvements. As a matter of fact it brings little improvement. However, in this cas also, the parameters have to be suitably chosen.

First, recall that, in the example of subsection 3.1.1, the problem derives from what the support of y_n is too small (cf also section 5.4). With data resulting from texts, one can find a problem of this type because the number of possible texts y_n is tiny compared to the number of all the sequences with values in$\{0, 1, ..., m-1\}$. It is the case if the d(j) belong to a larger set, i.e. if $m/32^{r_0}$ is large.

On the other hand if m is smaller than 32^{r_0}, one does not have maybe problem of this type. Admittedly all the possible texts belong to a subset fixed. But this set is unknown. It is even more unknowable if there are samples of size not too large compared to m. It is thus difficult to conclude from it logically that the probability is concentrated close to a small number of points. It will be thus better to choose $m \leq 32^{r_0}$ and samples d(j) of size not too large compared to m.

This example shows that it is necessary to choose well the parameters m and r_0.

3.2 Numerical results

Thie previous conclusion is confirmed by numerical study.

3.2.1 Test of the y'_n's

One carried out tests of independence between various sequence y'_n and y'_{n+p} .

It is known that the marginal distributions are uniform. One thus uses chi-squared test of independence without estimate of parameters : $\chi^2_I = \chi^2_{X,Y} - \chi^2_X - \chi^2_Y$: cf proposition A.1.1.

It is used with partitions (20,20) . Therefore χ^2_I has $D = 19^2$ degrees of freedom. Because D is big, the statistics $\chi^2_{Nor} = \sqrt{2\chi^2_I} - \sqrt{2D-1}$ has aproximatively the N(0,1) distribution when there is independence : cf proposition A.1.2

One uses this statistics for various samples of (y'_n, y'_{n+p}).

Then, for various "p", for 20 various samples, the following table of the maximum of $|\chi_{Nor}|$ has been obtained.

p	2	3	4	7	10	50
$Max\|\chi_{Nor}\|$	2.158	1.989	2.198	1.8054	1.6879	1.1457

All the carried out tests conclude to independence.

Also let us recall that Marsaglia affirms that the numbers of its CD-ROM passes all the test known. This result thus confirms its conclusions. But the construction of its sequence y'_n is a little more complex and thus even more random.

3.2.2 Test between y'_n and g_{n+p}

One makes tests of independence between various sequences y'_n and g_{n+p}. It is known that the marginal distributions are uniform. Then, one uses $\chi^2_I = \chi^2_{X,Y} - \chi^2_X - \chi^2_Y$ with partitions (20,20) . The statistics has $D = 19^2$ degrees of freedom.

Then, for various "p", for 20 various samples, the following table of the maximum of $|\chi^2_{Nor}|$ has been obtained:

p	2	3	4	7	10	50
$Max\|\chi^2_{Nor}\|$	2.170	2.054	2.024	1.7754	1.4493	2.2017

The tests that we have carried out conclude to independence.

3.2.3 Test between y'_n and y_{n+p}

We made the test of independence between various sequences y'_n and y_{n+p}. In this case, it is necessary to estimate the marginal distributions.

This test has been used with partitions (20,20) with the classical statistics which we note χ^2_{Es}. Then, for various "p", for 20 various samples, the following

table of the maximum of $|\chi^2_{Es}|$ has been obtained:

p	2	3	4	7	10	50
$Max\|\chi^2_{Es}\|$	2.130	2.0455	2.201	1.745	1.789.	1.634

3.2.4 Other tests

Now, one uses the polynomial correlation coefficients of higher order $\rho_{i,j}$ between y'_{n+1} and $(y_n, y_{n-1}, ...)$, $\rho'_{i,j}$ between y'_{n+1} and $(y'_n, y'_{n-1}, ...)$, and also $\rho"_{i,j}$ between y_{n+1} and $(y_n, y_{n-1}, ...)$: cf [10]; cf also section A.3. We know that $\sum_j \rho^2_{1,j} \leq 1$ and that $y'_{n+1} = f(y_n, y_{n-1}, ...)$ if and only if $\sum_j \rho^2_{1,j} = 1$: cf [10]. One estimates $\rho_{i,j}$ by using empirical correlation coefficients of higher order .

Then, one has estimated $\sum_j \rho^2_{1,j}$, for 20 samples of y_n and y'_n with size varying from 10.000 to 1.000.000. One varies r_0. One obtains the table representing the maximum of the $\sum_j \rho^2_{1,j}$:

$r_0 =$	5	6	7	8	9	10	11	12
$\sum_j (\rho'_{1,j})^2 \leq$	0.041	0.03	0.011	0.021	0.012	0.006	0.025	0.004
$\sum_j (\rho_{1,j})^2 \leq$	0.05	0.043	0.03	0.07	0.06	0.04	0.03	0.008
$\sum_j (\rho"_{1,j})^2 \leq$	0.2	0.16	0.13	0.1	0.08	0.06	0.055	0.01

In the same way one also estimated the conditional densities by using the density of dependence (cf th 5-3, page 8 [10]) and the empirical orthogonal functions page 10 [10].

On the figure 3.1 one has the curve of the conditional density f'^y of y'_{n+1} given y_n. On the figure 3.2 one has the curve of the conditional density f^y of y_{n+1} given y_n. It is in question the density where the maximum of $sup|f^y(y_n)|$ is reached.

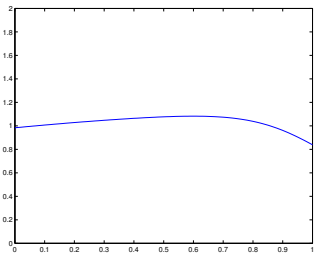

Figure 3.1: Conditional density of y'_{n+1} given y_n

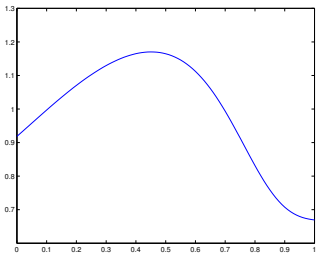

Figure 3.2: Conditional density of y_{n+1} given y_n

3.3 Conclusion

The previous study shows that one can improve the result by choosing better the parameters.

If they are well chosen, there is many reasons to think that y_n is IID. But we have not a certainty : that is difficult to specify mathematically. Maybe a thorough study would allow to arrive at certainties.

But it is simpler to use other transformations whose properties are appropriate well for the construction of an IID sequence and can be studied more easily. It is the aim of this report.

Chapter 4

Basic properties

4.1 Basic theorem

In this section the basic theorem 1 is proved. To this end, one uses functions T^* defined by the following way.

Notations 4.1.1 *Let $d \in \mathbb{N}^*$, $d \geq 2$, $p, q \in \mathbb{N}^*$, $p, q \geq 2$. Let $G(x) \equiv d^p x \ mod \ m$ defined on $F^*(m) = \{0, 1, ..., m-1\}$ where $m = d^{p+q} - 1$. We denote by $T^* : F(m) \to F(m)$ the function such that $m.T^*(k/m) \equiv G(k) \ mod \ m$ and $0 \leq T^*(k/m) < 1$ for k=0,1,...,m-1 .*

4.1.1 Proof of theorem 1

New statement of the theorem

With the previous notations , the theorem 1 can be specified by the following way.

Proposition 4.1.1 *Assume p=q. Let $Z \in F(m)$ be a random variable. Let f be the density of Z with respect to μ_m ($\mu_m(k/m) = 1/m$) : $f(x) = m.P\{Z = x\}$. Let $K_0 > 0$ such that, for all $z, z' \in F(m)$, $|f(z) - f(z')| \leq K_0|z - z'|$.*

Let I be an interval of F(m). Let N(I) be the number of points of F(m) which belong to I.

Then, if $N(I) \geq d^p$,

$$P\big\{T^*(Z) \in I\big\} = N(I)/m + e(K_0, d, p),$$

where

$$e(K, d, p) = 2(3K_0 + 4)/d^p + O(K_0/m),$$

with

$$O(K_0/m) = \frac{2(K_0+1)}{m}\ \Big[2 + \frac{(d^p+1)}{d^p(d^p-1)}\Big] + \frac{3(K_0+1)(1+1/d^p)}{d^p(d^p-1)}\ .$$

Proof of proposition 4.1.1

First, we simplify notations used in this proof.

Notations 4.1.2 *In order to simplify the notations, we set* $T = T^*$. *Let* $G'(x) \equiv d^q x \ mod\ m$ *defined on* $F^*(m)$. *We denote by T' the function* $T' : F(m) \to F(m)$ *such that* $mT'(k/m) \equiv G'(k)$ *modulo m and* $0 \leq T'(k/m) < 1$ *for k=0,1,...,m-1.*

Then the following lemmas hold.

Lemma 4.1.1 *The following equalities hold :* $G^{-1} = G'$, $(T^*)^{-1} = T'$.

Lemma 4.1.2 *Let I = [a,b[where a = k'/m and b= k*/m. We set* $T'(a) = \alpha$. *Let* $\widetilde{T}$ *be the function such that* $\widetilde{T}(k/m) - T'(k/m) = 0$ *or 1, and* $\alpha \leq \widetilde{T}(k/m) < \alpha + 1$.

Then, $T^{-1}(I) = T'([a, b[)$

Now one needs the following notations

Notations 4.1.3 *One defines* $\widetilde{f}[\alpha, \alpha + 1[\to \mathbb{R}$ *by* $\widetilde{f}(t) = f(t)$ *if* $t \in [\alpha, 1[$ *and* $\widetilde{f}(t) = f(t-1)$ *if* $t \in [1, \alpha + 1[$.

Notations 4.1.4 *We set* $D = d^q/(d^{p+q} - 1)$. *Let* $[\alpha, \alpha+1[= U_1 \cup U_2 \cup ..... \cup U_{d^p}$, *where* $U_s = [\alpha + (s-1)D, \alpha + sD[$ *for* $s = 1, 2, ..., d^p - 1$, *and* $U_{d^p} = [\alpha + 1 - (d^q - 1)/(d^{p+q} - 1)\ ,\ \alpha + 1[$.

Notations 4.1.5 *We denote by* $N(I) = Hd^p + r$ *the Euclidean division of N(I) by* d^p.

Remark that N(I) is the number of k/m such that $a \leq k/m < b$. Therefore N(I) is the number of $k/m \in F(m)$ which belongs to $\widetilde{T}([a, b[)$.

Then the following result holds.

Lemma 4.1.3 *For* $t = 1, ..., d^p - 1$

$$F(m) \cap U_t \cap \widetilde{T}(I) = \left\{ \alpha + (t-1)D + h/m \ \middle|\ h = 0, 1, ..., H - 1 + e_t \right\} ,$$

where $e_t = 1$ *if* $t \leq r$ *and* $e_t = 0$ *if not.*

Proof : At first $F(m) \cap \widetilde{T}([a, b[) = \{\widetilde{T}(a + k/m) \mid k = 0, 1, ..., N(I) - 1\}$.

Let $\widetilde{G}(x)$ be the function such that $\widetilde{G}(x) \equiv G'(x)$ and $m\alpha \leq \widetilde{G}(x) < m + m\alpha$. Then, $\widetilde{T}(a + k/m) = \widetilde{G}(ma + k)/m$ for k=0,1,...,N(I)-1.

Let $k = hd^p + s$ the Euclidean division of k by d^p. Then, $G'(ma + k) \equiv m\alpha + G'(k) \equiv m\alpha + d^q(hd^p + s) \equiv m\alpha + d^{p+q}h + sd^q \equiv m\alpha + (m+1)h + sd^q \equiv m\alpha + h + sd^q$.

Now $h + sd^q < m$.
Indeed, $h \leq d^q - 1$ and $s < d^p$. Then, $h + sd^q \leq (d^q - 1) + (d^p - 1)d^q = d^q - 1 + d^{p+q} - d^q = d^{p+q} - 1 = m$.
Moreover, if $h = d^q - 1$, $m > k = hd^p + s = (d^q - 1)d^p + s = d^{p+q} + s - d^p$. Then, if $s = d^p - 1$, $m > d^{p+q} - 1 = m$: that is impossible. Therefore, $s \leq d^p - 2$. Then, $h + sd^q \leq d^q - 1 + (d^p - 2)d^q < m$.

Then, $G'(ma + k)/m \equiv \widetilde{T}(a + k/m) = \alpha + h/m + sd^q/m = \alpha + h/m + sd^q/(d^{p+q} - 1) = \alpha + h/m + sD$.

Now, if $k \in [0, N(I) - 1] \cap \mathbb{N}$, $h \in [0, H - 1] \cap \mathbb{N}$ or $h \in [0, H] \cap \mathbb{N}$.
If h= 0, $s \in [0, d^p - 1] \cap \mathbb{N}$.
if h= 1, $s \in [0, d^p - 1] \cap \mathbb{N}$.
.....
if h= H-1, $s \in [0, d^p - 1] \cap \mathbb{N}$.
if h= H, $s \in [0, r - 1] \cap \mathbb{N}$.

We deduce the lemma. ■

We deduce the following result.

Lemma 4.1.4 *There are $H+e_t$ points of $\widetilde{T}(I)$ which belongs to U_t and H points of $\widetilde{T}(I)$ which belongs to U_{d^p}.*

Remark 4.1.5 *The points of F(m) which belongs to U_t are the H or H+1 first points of U_t .*

Then, the following lemmas holds

Lemma 4.1.6 *We set $\Delta_k = H + e_k$. Then, $\sum_{k=1}^{d^p} \Delta_k = d^p H + r = N(I)$.*

Lemma 4.1.7 *The following equality holds : $sup(f) = K_0 + 1$.*

Proof Indeed, f is defined on F(m). Then, there exists s_0 such that $f(s_0) \leq 1$. Then, by our assumptions $|f(s) - f(s_0)| \leq K_0|s - s_0| \leq K_0$ for all $s \in F(m)$. ■

Lemma 4.1.8 *Let $\widetilde{Z}$ be the sequence $\widetilde{Z}(\omega) = Z(\omega)$ if $Z(\omega) \geq \alpha$ and $\widetilde{Z}(\omega) = Z(\omega) + 1$ if $Z(\omega) < \alpha$ for all $\omega \in \Omega$. Then,*

$$\big\{T(Z) \in [a, b[\big\} = \cup_k \big\{\widetilde{Z} \in U_k \cap \widetilde{T}([a, b[)\big\}.$$

Proof The following equalities hold,

$$\big\{T(Z) \in [a, b[\big\} = \big\{Z \in T^{-1}([a, b[)\big\} = \big\{Z \in T'([a, b[)\big\}$$

$$= \Big\{Z \in \big\{[0, \alpha[\cap T'([a, b[)\big\}\Big\} \cup \Big\{Z \in \big\{[\alpha, 1[\cap T'([a, b[)\big\}\Big\}$$

$$= \Big\{ Z+1 \in [1, \alpha+1[\cap \big\{ T'([a,b[)+1) \big\} \Big\} \cup \Big\{ Z \in \big\{ [\alpha, 1[\cap T'([a,b[) \big\} \Big\}$$

$$= \Big\{ Z+1 \in \big\{ [1, \alpha+1[\cap \widetilde{T}([a,b[) \big\} \Big\} \cup \Big\{ Z \in \big\{ [\alpha, 1[\cap \widetilde{T}([a,b[) \big\} \Big\}$$

$$= \Big\{ \widetilde{Z} \in \big\{ [1, \alpha+1[\cap \widetilde{T}([a,b[)) \big\} \Big\} \cup \Big\{ \widetilde{Z} \in \big\{ [\alpha, 1[\cap \widetilde{T}([a,b[) \big\} \Big\}$$

$$= \big\{ \widetilde{Z} \in \widetilde{T}([a,b[) \big\}$$

$$= \cup_k \big\{ \widetilde{Z} \in U_k \cap \widetilde{T}([a,b[) \big\} \ . \ \blacksquare$$

Lemma 4.1.9 *The Euclidean Division of m by d^p is $m = Hd^p + r$ with $r = d^p - 1$ and $H = d^q - 1$. Moreover, $H/m = Ob(1)/d^p$.*

Proof Indeed, $m = d^{p+q} - 1 = Hd^p + r = d^p(d^q - 1) + r = d^{p+q} - d^p + r$: $r = d^p - 1$.

Moreover, $(H/m) \leq (d^q - 1)/(d^{p+q} - 1) = (1/d^p)(d^{p+q} - d^p)/(d^{p+q} - 1) \leq (1/d^p)$. $\blacksquare$

Lemma 4.1.10 *Let k_1 such that $1 \in U_{k_1}$. Then*

$$P\big\{T(Z) \in [a,b[\big\} = (H/m) \sum_{k \neq k_1} f\big\{\alpha + (k-1)D\big\} + (K_0+1) Ob(1) \Big(\frac{1}{d^q} + \frac{2}{d^p} + \frac{2}{m} \Big) .$$

Proof Indeed,

$$P\big\{T(Z) \in [a,b[\big\} = \sum_k P\big\{ \widetilde{Z} \in U_k \cap \widetilde{T}([a,b[) \big\}$$

$$= \sum_k \sum_{\tau=0}^{\Delta_k - 1} P\big\{ \{ \widetilde{Z} = \alpha + (k-1)D + \tau/m \big\}$$

$$= \sum_k \int_{0 \leq t \leq \ (\Delta_k - 1)/m} f\big\{ \{ t + \alpha + (k-1)D \big\} . \mu_m(dt)$$

$$= \sum_{k \neq k_1} \int_{0 \leq t \leq (\Delta_k - 1)/m} f\big\{ \{ t + \alpha + (k-1)D \big\} . \mu_m(dt)$$

$$+ \sum_{k = k_1} \int_{0 \leq t \leq (\Delta_k - 1)/m} f\big\{ \{ t + \alpha + (k-1)D \big\} . \mu_m(dt)$$

$$= \sum_{k\neq k_1} \int_{0\leq t\leq(\Delta_k-1)/m} f\big\{\{\alpha+(k-1)D\big\}.\mu_m(dt)$$

$$+\sum_{k\neq k_1} \int_{0\leq t\leq(\Delta_k-1)/m} \Big[f\big\{\{t+\alpha+(k-1)D\big\}-f\big\{\{\alpha+(k-1)D\big\}\Big].\mu_m(dt)$$

$$+\int_{0\leq t\leq(\Delta_{k_1}-1)/m} sup(f)Ob(1)(t).\mu_m(dt)$$

$$= \sum_{k\neq k_1} (\Delta_k/m) f\big\{\{\alpha+(k-1)D\big\}$$

$$+\sum_{k\neq k_1} \int_{0\leq t\leq(\Delta_k-1)/m} [K_0(\Delta_k-1)/m)]Ob(1)(t).\mu_m(dt)$$

$$+Ob(1)(\Delta_{k_1}/m)sup(f)$$

$$= \sum_{k\neq k_1} (H/m+e_k/m) f\big\{\{\alpha+(k-1)D\big\}$$

$$+\sum_{k\neq k_1} [K_0(H/m)Ob(1)(\Delta_k/m)]+Ob(1)[(H+1)/m]sup(f)$$

$$= (H/m+Ob(1)/m)\sum_{k\neq k_1} f\big\{\{\alpha+(k-1)D\big\}$$

$$+K_0(H/m)\sum_{k\neq k_1} Ob(1)(\Delta_k/m)+Ob(1)[(H+1)/m]sup(f)$$

$$= (H/m)\sum_{k\neq k_1} f\big\{\{\alpha+(k-1)D\big\}$$

$$+Ob(1)(d^p/m)sup(f)$$

$$+Ob(1)K_0(H/m)(N(I)/m)+Ob(1)[(H+1)/m]sup(f)$$

$$= (H/m)\sum_{k\neq k_1} f\big\{\{\alpha+(k-1)D\big\}$$

$$+Ob(1)(K_0+1)\big[d^p/m+(H/m)(N(I)/m)+(H+1)/m\big].$$

By the previous lemma 4.1.9, $H/m = Ob(1)/d^p$. Then, $(H/m)(N(I)/m) = Ob(1)(1/d^p)$. At last, $d^p/m = d^p/(d^{p+q}-1) = (1/d^q)[d^{p+q}/(d^{p+q}-1)] = (1/d^q)[1+1/(d^{p+q}-1)] = (1/d^q)(1+1/m)$.

Therefore,

$$(H/m)\sum_{k\neq k_1} f\big\{\{\alpha+(k-1)D\big\}$$
$$+Ob(1)(K_0+1)\big[d^p/m + (H/m)(N(I)/m) + (H+1)/m\big].$$

$$= (H/m)\sum_{k\neq k_1} f\big\{\alpha+(k-1)D\big\}$$
$$+(K_0+1)Ob(1)\big(1/d^q + 2/d^p + 2/m\big)\ . \blacksquare$$

Lemma 4.1.11 *The following equality holds.*

$$(H/m)\sum_{k\neq k_1} f\big\{\{\alpha+(k-1)D\big\}$$

$$= N(I)/m + \frac{2Ob(1)(1-d^p)}{d^q-1}$$
$$+(K_0+1)Ob(1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]\Big[1+\frac{1}{d^q-1}+\frac{1}{d^p(d^q-1)}\Big]\ .$$

Proof : One chooses [a,b[= [0,1[in the previous lemma. If $N(I) = m = Hd^p + r$, $r = d^p - 1$ and $H = d^q - 1$.

Then,

$$H/m = (d^q-1)/(d^{p+q}-1) = (1/d^p)(d^{p+q}-d^p)/(d^{p+q}-1)$$
$$= (1/d^p)(d^{p+q}-1)/(d^{p+q}-1) + (1/d^p)(1-d^p)/(d^{p+q}-1)$$
$$= (1/d^p) + Ob(1)/(d^{p+q}-1) = (1/d^p) + Ob(1)/m.$$

Then, by the previous proof,

$$1 = P\big\{T(Z)\in[0,1[\big\}$$

$$= (H/m)\sum_{k\neq k_1} f\big\{\{\alpha+(k-1)D\big\} + (K_0+1)Ob(1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]$$

$$= \frac{d^q - 1}{d^{p+q} - 1} \sum_{k \neq k_1} f\big\{\{\alpha + (k-1)D\big\} + (K_0 + 1)Ob(1)\Big[\frac{1}{d^q} + \frac{2}{d^p} + \frac{2}{m}\Big] \ .$$

Then,

$$\frac{d^{p+q} - 1}{d^q - 1}$$

$$= \sum_{k \neq k_1} f\big\{\{\alpha + (k-1)D\big\} + (K_0 + 1)Ob(1)\Big[\frac{1}{d^q} + \frac{2}{d^p} + \frac{2}{m}\Big]\frac{d^{p+q} - 1}{d^q - 1} \ .$$

We know that $H/m \leq 1/d^p$. Then,

$$\frac{H}{m} \ \frac{d^{p+q} - 1}{d^q - 1} = \frac{Ob(1)}{d^p} \ \frac{d^{p+q} - 1}{d^q - 1}$$

$$= \frac{Ob(1)(d^{p+q} - d^p)}{d^{p+q} - d^p} + \frac{Ob(1)(d^p - 1)}{d^{p+q} - d^p}$$

$$= Ob(1) + \frac{Ob(1)(d^p - 1)}{d^p(d^q - 1)} = Ob(1) + \frac{Ob(1)d^p}{d^p(d^q - 1)} + \frac{Ob(1)}{d^p(d^q - 1)}$$

$$= Ob(1) + \frac{Ob(1)}{d^q - 1} + \frac{Ob(1)}{d^p(d^q - 1)} \ .$$

Moreover

$$\frac{H}{m} \ \frac{d^{p+q} - 1}{d^q - 1}$$

$$= \frac{N(I)}{md^p} \ \frac{d^{p+q} - 1}{d^q - 1} - \frac{r}{md^p} \ \frac{(d^{p+q} - 1)}{d^q - 1}$$

$$= \frac{N(I)}{m} \ \frac{d^q - 1/d^p}{d^q - 1} - \frac{r}{d^p} \ \frac{1}{d^q - 1}$$

$$= \frac{N(I)}{m} + \frac{N(I)}{m} \ \frac{1 - 1/d^p}{d^q - 1} + \frac{Ob(1)(d^p - 1)}{d^{p+q} - d^p}$$

$$= \frac{N(I)}{m} + \frac{2Ob(1)(d^p - 1)}{d^{p+q} - d^p} \ .$$

Then,

$$\frac{H}{m} \sum_{k \neq k_1} f\big\{\alpha + (k-1)D\big\}$$

$$= \frac{H}{m}\ \frac{d^{p+q}-1}{d^q-1} + \frac{(K_0+1)Ob(1)H}{m}\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]\frac{d^{p+q}-1}{d^q-1}$$

$$= \frac{N(I)}{m}+\frac{2Ob(1)(1-1/d^p)}{d^q-1}+(K_0+1)Ob(1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]\Big[1+\frac{1}{d^q-1}+\frac{1}{d^p(d^q-1)}\Big]. \blacksquare$$

Lemma 4.1.12 *The following equality holds :*

$$P\big\{T(Z)\in[a,b[\big\} = \frac{N(I)}{m} + Ob(1).e(K_0,d,p,q)\ ,$$

where

$$e(K_0,d,p,q) = \frac{2(1-1/d^p)}{d^q-1}+(K_0+1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]\Big[2+\frac{1}{d^q-1}+\frac{1}{d^p(d^q-1)}\Big]\ .$$

Proof The following equalities hold :

$$P\big\{T(Z_n)\in[a,b[\big\}$$
$$= \frac{H}{m}\sum_{k\neq k_1} f\big\{\alpha+(k-1)D\big\} + (K_0+1)Ob(1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]$$

$$= \frac{N(I)}{m} + \frac{2Ob(1)(1-1/d^p)}{d^q-1}$$
$$+(K_0+1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]\Big[1+\frac{1}{d^q-1}+\frac{1}{d^p(d^q-1)}\Big]$$
$$+(K_0+1)Ob(1)\Big[\frac{1}{d^q}+\frac{2}{d^p}+\frac{2}{m}\Big]. \blacksquare$$

.

Lemma 4.1.13 *If p=q,*

$$e(K_0,d,p,p) = \frac{2(3K_0+4)}{d^p} + O(K_0/m).$$

where

$$O(K_0/m) = \frac{2(K_0+1)}{m}\ \Big[2+\frac{(d^p+1)}{d^p(d^p-1)}\Big] + \frac{3(K_0+1)(1+1/d^p)}{d^p(d^p-1)}\ .$$

Proof The following equality holds

$$e(K_0,d,p,p) = \frac{2(1-1/d^p)}{d^p-1} + (K_0+1)\Big[\frac{1}{d^p}+\frac{2}{d^p}+\frac{2}{m}\Big]\Big[2+\frac{1}{d^p-1}+\frac{1}{d^p(d^p-1)}\Big]$$

$$= \frac{2}{d^p} + (K_0+1)\Big[\frac{3}{d^p}+\frac{2}{m}\Big]\Big[2+\frac{1}{d^p-1}+\frac{1}{d^p(d^p-1)}\Big]$$

$$= \frac{2}{d^p} + \frac{6(K_0+1)}{d^p} + O(K/m) = \frac{2(3K_0+4)}{d^p} + O(K_0/m)\ ,$$

where

$$O(K_0/m) = \frac{2(K_0+1)}{m}\Big[2+\frac{1}{d^p-1}+\frac{1}{d^p(d^p-1)}\Big]+\frac{3(K_0+1)}{d^p}\Big[\frac{1}{d^p-1}+\frac{1}{d^p(d^p-1)}\Big]$$

$$= \frac{2(K_0+1)}{m}\ \Big[2+\frac{(d^p+1)}{d^p(d^p-1)}\Big] + \frac{3(K_0+1)(1+1/d^p)}{d^p(d^p-1)}\ .\ \blacksquare$$

.

Proof 4.1.14 *In order to prove proposition 4.1.1, it is enough to apply the previous lemmas.*

4.1.2 Calculation of T^*

Now, in practice, the form of T^* is very particular. Indeed T^* permutes the first decimals with the last ones. For example, if d = 10, p=3, q= 2, $x = 45873$, $T^*(x) = 87345$.

Proposition 4.1.2 *We keep the notations of proposition 4.1.1. Let $m = d^{p+q}-1$, $a = d^q$. Let $x \in F^*(d^{p+q}-1)$. Let $x = x'd^p + x"$ the Euclidean division of x by d^p.*

Then, $mT'(x) = r = x"d^q + x'$.

Proof By definition $x \leq d^{p+q}-2$ and $x" \leq d^p-1$. Moreover, $x' \leq d^q-1$. Moreover, if $x' = d^q-1$, $x" \leq d^p-2$.

Let $X = d^q x$. Then, $X = d^q x \leq d^{p+2q} - 2d^q$.

Then, $X = x'd^{p+q} + x"d^q$ with $x"d^q \leq d^{p+q} - d^q$ and if $x' = d^q-1$, $d^q x" \leq d^{p+q} - 2d^q$.

Therefore, $X = x'd^{p+q} + x"d^q = x'(d^{p+q}-1) + x"d^q + x'$. Moreover, $x"d^q + x' < d^{p+q} - 1 = m$

Then, $X = x'm + x"d^q + x' = x'm + r$ where $r < m$.

Therefore, $T^*(x) = r = x"d^q + x'$. ■.

Write x in base d : $x = \overline{\overline{x_{p+q}, x_{p+q-1}, ..., x_{p+1}, x_p, ..., x_1}}$.
Then, $T^*(x) = \overline{\overline{x_p, ..., x_1, x_{p+q}, x_{p+q-1}, ..., x_{p+1}}}$.

Indeed $x = \overline{\overline{x_{p+q}, x_{p+q-1}, ..., x_{p+1}}}d^p + \overline{\overline{x_p, x_{p-1}, ..., x_1}}$, where $\overline{\overline{x_p, x_{p-1}, ..., x_1}} \leq d^p - 1$.

For example, if d = 10, p=4, q= 4, x = 21058453, X= 21058453000, $T^*(x) =$ 84532105.

4.2 Some properties

Let $m \in \mathbb{N}^*$. Let $X_n \in F(m)$ be a sequence of random variables. In this section we study some properties of conditional probabilites when $P\big\{X_n \in I|x_2, ..., x_p)\big\} = L(I) + Ob(1)\epsilon$.

First, the following result is obvious.

Proposition 4.2.1 *Assume $P\{X_n \in I\} = L(I) + Ob(1)\epsilon$ for all $n \in \mathbb{N}^*$ and all interval $I \subset [0, 1]$.*

Let Bo be a Borel set : $Bo = \cup_{s=1}^k I_s$ where the I_s are disjoint interval $I_s \subset [0, 1]$. Then, $P\{X_n \in Bo\} = L(Bo) + Ob(1)k\epsilon$.

Then, one obtains the probability of Borel sets of $F(m)^p$.

Proposition 4.2.2 *Let Bo be a Borel set of $F(m)^p$, $Bo = Bo_1 \otimes Bo_2 \otimes ..... \otimes Bo_p$. Assume that, for all $s \in \{1, 2, ..., p\}$, for all sequence x_s, s=1,...,p, and for all $n \in N^*$, $P\big\{X_n \in Bo_s|x_2, ..., x_p)\big\} = L(Bo_s) + Ob(1)\epsilon$.*

Then, for all injective sequence $j_s \in \mathbb{Z}$ such that $j_1 = 0$,

$$P\Big\{\{X_{n+j_1} \in Bo_1\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\} = \Big[L(Bo_1) + Ob(1)\epsilon\Big].....\Big[L(Bo_p) + Ob(1)\epsilon\Big] \ .$$

In order to prove this proposition the following lemma is needed

Lemma 4.2.1 *Let $Y_s \in F(m)$, s=1,2,...,N, be a sequence of random variables defined over a probability space (Ω, A, P). Let $f \in L^1$ be a measurable function defined over $Y^-(\Omega)$ where $Y^- = (Y_1, Y_2, ..., Y_{n-1}, Y_{n+1}, ..., Y_N)$ and $n \in \{1, 2, ..., N\}$. Let Bo_1 be a Borel set of F(m).*

Assume $P\big\{Y_n \in Bo_1|y_1, ..., y_{n-1}, y_{n+1}, ..., y_N\big\} = L(Bo_1) + Ob(1)\epsilon$ for all $(y_1, ..., y_{n-1}, y_{n+1}, ..., y_N)$.

Then,

$$E\big\{1_{Bo_1}(Y_n)f(Y^-)\big\} = L(Bo_1)E\big\{f(Y^-)\big\} + Ob(1)\epsilon E\big\{|f(Y^-)|\big\}\ .$$

Proof Let Q be the distribution of $(Y_1, Y_2, ..., Y_N)$ and let Q^- be the distribution of $(Y_1, Y_2, ..., Y_{n-1}, Y_{n+1}, ..., Y_N)$. Let $Q(.|y_1, ..., y_{n-1}, y_{n+1}, ..., y_N)$ be the distribution of Y_n given $Y_s = y_s$, for s=1,2,...,n-1,n+1,...,N.

Let $y^- = (y_1, ..., y_{n-1}, y_{n+1}, ..., y_N)$. Then,

$$E\big\{1_{Bo_1}(Y_n)f(Y^-)\big\} = \int 1_{Bo_1}(y_n)f(y^-)Q(dy)$$

$$= \int \Big(\int 1_{Bo_1}(y_n)Q(dy_n|y_1, ..., y_{n-1}, y_{n+1}, ..., y_N)\Big)f(y^-)Q^-(dy^-)$$

$$= \int P\big\{Y_n \in Bo_1|y_1, ..., y_{n-1}, y_{n+1}, ..., y_N)\big\}f(y^-)Q^-(dy^-)$$

$$= L(Bo_1)\int f(y^-)Q^-(dy^-) + \int Ob(1)\epsilon\big(y^-)\big)f(y^-)Q^-(dy^-),$$

where $|\epsilon\big(y^-)\big)| \leq \epsilon$.

Then,

$$E\big\{1_{Bo_1}(Y_n)f(Y^-)\big\} = L(Bo_1)\int f(y^-)Q^-(dy^-) + Ob(1)\epsilon\int |f(y^-)|Q^-(dy^-)\ .\ \blacksquare$$

Proof 4.2.2 *We prove the proposition 4.2.2*

We use the lemma 4.2.1 with N=p, $X_{n+j_s} = Y_s$. Moreover, we choose $f(Y^-) = 1_{Bo_2}(Y_{n+j_2})$.....$1_{Bo_p}(Y_{n+j_p})$. Then,

$$P\Big\{\{X_{n+j_1} \in Bo_1\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}$$

$$= \big(L(Bo_1) + Ob(1)\epsilon\big)P\Big\{\{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}\ .$$

Then, we prove the proposition by recurence. ■

Now, one obtains a similar resul about conditional probability.

Proposition 4.2.3 *Let Bo be a Borel set of $F(m)^p$, $Bo = Bo_1 \otimes ..... \otimes Bo_p$. Assume that $P\{X_n \in Bo_1 | x_2, ..., x_p\} = L(Bo_1) + Ob(1)\epsilon$.*

Then,

$$P\Big\{X_n \in Bo_1 \Big| \{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\} = L(Bo_1) + Ob(1)\epsilon.$$

Proof By using the proof 4.2.2 ,

$$P\Big\{X_{n+j_1} \in Bo_1 \Big| \{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}$$

$$= \frac{P\Big\{\{X_{n+j_1} \in Bo_1\} \cap \{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}}{P\Big\{\{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}}$$

$$= \frac{\Big(L(Bo_1) + Ob(1)\epsilon\Big) P\Big\{\{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}}{P\Big\{\{X_{n+j_2} \in Bo_2\} \cap ..... \cap \{X_{n+j_p} \in Bo_p\}\Big\}}$$

$$= L(Bo_1) + Ob(1)\epsilon. \blacksquare$$

The proof of the following theorem is a consequence of proposition 4.2.2.

Proposition 4.2.4 *The sequence X_n, n=1,2,.....,N, is IID if and only if, for all $p \in \{1, 2, ..., N-1\}$, for all $n \in \mathbb{N}^*$, for all Borel set Bo, for all sequence x_s, s=1,...,p*

$$P\Big\{X_n \in Bo \Big| X_{n+j_2} = x_2, ..., X_{n+j_p} = x_p\Big\} = L(Bo) \ .$$

4.3 Some properties of random bits

In this section, we study the property of sequence of bits defined by the following way.

Notations 4.3.1 *Let $q \in \mathbb{N}^*$. Let $R_n \in F(m)$ be a sequence of random variables such that $P(R_n \in Bo | x_2, ..., x_p) = L(Bo) + Ob(1)k\epsilon$ for all Borel set $Bo = \cup_{t=1}^{k} I_t$ where the I_t's are intervals such that $I_t \neq \emptyset$.*

Write X_n base d: $R_n = \overline{0, D_1^n D_2^n .....}$. We set $X_n = \overline{0, D_1^n D_2^n ....D_q^n}$ and $D_{qn-r+1} = D_r^n$ for n=1,...,N and r=1,...,q.

The $D_{n'}$'s are obtained by taking succesively $D_q^n, D_{q-1}^n ...., D_1^n$ .....

The proof of the following lemma is almost obvious.

Lemma 4.3.1 *Let* $1 \leq r \leq r' \leq q$. *Let* $j_1' = 0 < j_2' < .... < j_p'$ *where* $r + j_p' = r' \leq q$. *Then,*

$$\{D_{r+j_1'}^n = d_1\} \cap \{D_{r+j_2'}^n = d_2\} \cap ... \cap \{D_{r+j_p'}^n = d_p\}\Big\}$$

$$= \bigcup_{d_s \;\; s \leq r', s \neq r+j_s'} \Big\{ X_n \in [\overline{0, d_1 d_2 .... d_{r'}}, \overline{0, d_1 d_2 .... d_{r'}} + d^{-r'}[\Big\} \; .$$

For example, d=10, r'=3 et r= 2. Then,

$$\{D_2^n = 5\} \cap \{D_3^n = 3\}$$

$$= \{X_n \in [0.053, 0.054[\} \cup \{X_n \in [0.153, 0.154[\} \cup ..... \cup \{X_n \in [0.953, 0.954[\}.$$

For example, d=10, r'=4 et r= 2. Then,

$$\{D_2^n = 5\} \cap \{D_4^n = 3\}$$

$$= \{X_n \in [0.0503, 0.0504[\} \cup \{X_n \in [0.1503, 0.1504[\} \cup ..... \cup \{X_n \in [0.9503, 0.9504[\}$$
$$\cup \{X_n \in [0.0513, 0.0514[\} \cup \{X_n \in [0.1513, 0.1514[\} \cup ..... \cup \{X_n \in [0.9513, 0.9514[\}$$
.....
$$\cup \{X_n \in [0.0593, 0.0594[\} \cup \{X_n \in [0.1593, 0.1594[\} \cup ..... \cup \{X_n \in [0.9593, 0.9594[\}.$$

We deduce the following lemma

Lemma 4.3.2 *We keep the notation of the previous lemma. Then,*

$$P\Big\{\{D_{r+j_1'}^n = d_1\} \cap \{D_{r+j_2'}^n = d_2\} \cap ... \cap \{D_{r+j_p'}^n = d_p\}\Big\}$$

$$= \frac{1}{d^p} + Ob(1) d^{r'-p} \epsilon = \frac{1}{d^p} + Ob(1) d^{q-1} \epsilon \; .$$

Then, one can prove the following property.

Proposition 4.3.1 *Suppose* $1/2 \leq 1 - d^q \epsilon$. *Then,*

$$P\big\{D_{n+j_1} = d_1 \big| D_{n+j_2} = d_2, ..., D_{n+j_p} = d_p\big\} = \frac{1}{d}\Big[1 + \frac{2Ob(1) d^q \epsilon}{1 - d^q \epsilon}\Big] \; .$$

Proof There exists n^* such that $\{D_{n+j_1} = d_1\} = \{X_{n^*} \in Bo_1\}$.

For all $s = 1, 2, ..., p$, let $n + j_s = qn_s + r_s$ be the Eucidean division of $n + j_s$ by q. One can assume that $n + j_s = qn^* + r_s$ if and only if $s = 1, 2, ..., e$.

Then, there exists two Borel sets Bo_1 and Bo_1^* such that

$$\{D_{n+j_1} = d_1\} \cap ..... \cap \{D_{n+j_e} = d_e\} = \{X_{n^*} \in Bo_1\}$$

$$\{D_{n+j_2} = d_2\} \cap ..... \cap \{D_{n+j_p} = d_p\} = \{X_{n^*} \in Bo_1^*\} \ .$$

Remark that it is possible that $Bo_1^* = \emptyset$

More generally, there exists a sequence $i_1 = 0$, $i_2, ..., i_{p'}$, $p' \leq p$, and a sequence of Borel sets Bo_s s=1,2,...,p', such that

$$\{D_{n+j_1} = d_1\} \cap \{D_{n+j_2} = d_2\} \cap ..... \cap \{D_{n+j_p} = d_p\}$$

$$= \{X_{n^*+i_1} \in Bo_1\} \cap \{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\} \ .$$

Moreover,

$$\{D_{n+j_2} = d_2\} \cap ..... \cap \{D_{n+j_p} = d_p\}$$

$$= \{X_{n^*+i_1} \in Bo_1^*\} \cap \{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{\{X_{n^*+i_{p'}} \in Bo_{p'}\}$$

if $e > 1$.

On the other hand,

$$\{D_{n+j_2} = d_2\} \cap ..... \cap \{D_{n+j_p} = d_p\} = \{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\},$$

if e=1.

Therefore, if $e = 1$, by lemmas 4.2.1 and 4.3.2,

$$P\big\{D_{n+j_1} = d_1 \big| D_{n+j_2} = d_2, ..., D_{n+j_p} = d_p\big\}$$

$$= \frac{P\Big\{\{D_{n+j_1} = d_1\} \cap \{D_{n+j_2} = d_2\} \cap ..... \cap \{D_{n+j_p} = d_p\}\Big\}}{P\Big\{\{D_{n+j_2} = d_2\} \cap ..... \cap \{D_{n+j_p} = d_p\}\Big\}}$$

$$= \frac{P\Big\{\{X_{n^*+i_1} \in Bo_1\} \cap \{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}{P\Big\{\{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}$$

$$= \frac{\big[L(Bo_1) + Ob(1)d^{r'-e}\epsilon\big] P\Big\{\{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}{P\Big\{\{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}$$

$$= L(Bo_1) + Ob(1)d^{r'-e}\epsilon = \frac{1}{d^e} + Ob(1)d^{r'-e}\epsilon = \frac{1}{d} + Ob(1)d^{r'-1}\epsilon \ .$$

If $e > 1$, one can write

$$P\big\{D_{n+j_1} = d_1 \big| D_{n+j_2} = d_2, ..., D_{n+j_p} = d_p\big\}$$

$$= \frac{P\Big\{\{X_{n^*+i_1} \in Bo_1\} \cap \{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}{P\Big\{\{X_{n^*+i_1} \in Bo_1^*\} \cap \{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}$$

$$= \frac{\big[L(Bo_1) + Ob(1)d^{r'-e}\epsilon\big] P\Big\{\{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}{\big[L(Bo_1^*) + Ob(1)d^{r"-(e-1)}\epsilon\big] P\Big\{\{X_{n^*+i_2} \in Bo_2\} \cap ..... \cap \{X_{n^*+i_{p'}} \in Bo_{p'}\}\Big\}}$$

$$= \frac{L(Bo_1) + Ob(1)d^{r'-e}\epsilon}{L(Bo_1^*) + Ob(1)d^{r"-(e-1)}\epsilon}.$$

$$= \frac{(1/d^e) + Ob(1)d^{r'-e}\epsilon}{(1/d^{e-1}) + Ob(1)d^{r'-(e-1)}\epsilon}$$

$$= \frac{1}{d}\frac{1 + Ob(1)d^{r'}\epsilon}{1 + Ob(1)d^{r'}\epsilon} = \frac{1}{d}\Big[1 + \frac{1 + Ob(1)d^{r'}\epsilon}{1 + Ob(1)d^{r'}\epsilon} - 1\Big]$$

$$= \frac{1}{d}\Big[1 + \frac{Ob(1)d^{r'}\epsilon + Ob(1)d^{r'}\epsilon}{1 + Ob(1)d^{r'}\epsilon}\Big] = \frac{1}{d}\Big[1 + \frac{2Ob(1)d^{r'}\epsilon}{1 - d^{r'}\epsilon}\Big]$$

$$= \frac{1}{d} + \frac{2Ob(1)d^{r'-1}\epsilon}{1 - d^{r'}\epsilon} \ .$$

Moreover, $d^{r'-1}\epsilon \leq \frac{2d^{r'-1}\epsilon}{1-d^{r'}\epsilon}$. Then, in all cases,

$$P\big\{D_{n+j_1}=d_1\big|D_{n+j_2}=d_2,...,D_{n+j_p}=d_p\big\}=\frac{1}{d}\Big[1+\frac{2Ob(1)d^{r'}\epsilon}{1-d^{r'-1}\epsilon}\Big].$$

$$=\frac{1}{d}\Big[1+\frac{2Ob(1)d^{q}\epsilon}{1-d^{q}\epsilon}\Big]\ .\ \blacksquare$$

We deduce the following result.

Proposition 4.3.2 *Suppose d=2. Then*

$$P\big\{D_{n+j_1}=d_1\big|D_{n+j_2}=d_2,...,D_{n+_p}=d_p\big\}=\frac{1}{d}+\frac{Ob(1)d^{q}\epsilon}{1-d^{q}\epsilon}\ .$$

Chapter 5

Limit Theorems

In this chapter, we study the classical Central Limit Theorem (CLT) and a new limit theorem : the XOR Limit Theorem (XORLT). It corresponds with OR Exclusive (XOR) but used modulo m (and not only modulo 2).

5.1 Central Limit Theorem

The Central Limit Theorem is a classical result. It produces the limit distribution of $(X_1 + ...X_n)/\sigma$ when X_n is a sequence of random variables such that $\mathbb{E}\{X_n\} = 0$ and σ^2 is the variance.

First, it has been proved for independent sequences of random variables X_n. These results have been generalized under various hypotheses of asymptotical independence. The most known results have been proved by Ibragimov under the strong mixing condition or under martingale assumptions : cf [21] and [41].

Moreover, some authors have also proved the convergence of moment : cf Cox-Kim [36], Ibragimov-Lifshits [37], Soulier [38], Rozovsky [39]. Bernstein, [27] , Yokohama ([34],[35]), Brown [28] , Eissein-Janson [29], Hernndorf [30], Birkel [31], Krugov [32], Mairoboda [33]. Recall that this convergence implies the convergence in distribution.

Now, the strong-mixing condition is too strong for most of datas. Indeed, it is a very strong condition. Then, some authors have introduce weaker hypotheses : Versik Ornstein ([22], [23]), Cogburn [25] Rosenblatt [26], Pinskers [7] , Doukhan-Louichi [40]

For example, Withers [24] has introduced the ℓ-mixing condition.

Definition 5.1.1 *: Let $\sigma(N)^2$ be the variance of $\sum_{j=m}^{n} X_j$ where N=n-m+1. Let $u \in \mathbb{R}$, $0 \leq k \leq n-m$, $N \geq 1$.*

Let $\ell_N(k,u) = max_{m \leq j \leq n-k} Sup\big|cov(e^{iuP}, e^{-iuF})\big|$,
where cov is the covariance, where $P = \frac{1}{\sigma(N)} \sum_{\ell=m}^{j} \delta_\ell X_\ell$ and $F = \frac{1}{\sigma(N)} \sum_{\ell=j+k}^{n} \delta_\ell X_\ell$ and the sup is over $\delta_\ell = 0$ or 1.

We set $\ell(k,u) = sup_{N:k \leq N-m} \ell_N(k,u)$.

Then, X_n is $\ell-$mixing if, for all $u\in\mathbb{R}$, $\ell(k,u)\to 0$ as $k\to\infty$

Then, X_n is strongly $\ell-$mixing if, for all u, there exists $K(u)<\infty$ such that $\ell(k,u)\leq \ell(k)K(u)$ where $\ell(k)\to 0$ as $k\to\infty$

Doukhan-Louhichi [40] have introduced the weak-dependence.

Definition 5.1.2 *: Let $\mathcal{L}=\cup_{p=1}^{\infty}\mathcal{L}^p$ where $\mathcal{L}^p=\{f\ :\ \mathbb{R}^p\to\mathbb{R}\}$. Let $\Psi:\mathcal{L}\otimes\mathcal{L}\otimes(\mathbb{N}^*)^2\to\mathbb{R}_+$ and $(\theta_r)_{r\in\mathbb{N}}\searrow 0$.*

The sequence $\{X_n\}_{n\in\mathbb{Z}}$ is $(\theta,\mathcal{L},\Psi)$ weakly dependent if
$\forall r\in\mathbb{N}$, $\forall u,v\in\mathbb{N}^*$, $\forall(h,k)\in\mathcal{L}^u\otimes\mathcal{L}^v$,
$\forall\ i_1<i_2<.....<i_u<i_u+r\leq j_1<.....<j_v$,

$$\big|Cov\big(h(X_{i_1},....,X_{i_u}),k(X_{j_1},....,X_{j_v})\ \big)\big|\leq\theta_r\Psi(h,k,u,v).$$

But they did not studied the moments's convergence. Moreover theses assumptions are still strong. Indeed, we need very weak dependence assumptions in order to build easily IID sequences if we want to use the method of this report.

Fortunately, another look is possible : one can use higher order correlation coefficients (cf Lancaster [9], Blacher [10]; cf also section A.3). Then, in [11] we have turned the convergence of moments into an equivalent condition on these coefficients. Then, one has minimal condition for convergence of moments. For example we have proved the following theorem.

Theorem 2 *Assume that the X_n have the same distribution with variance σ^2 and that there exists $bo>0$ such that $|X_n|\leq bo$. Assume that*

$$\frac{\sum_{s=1}^n\sum_{r\neq s}\big[\mathbb{E}\{(X_s)^2(X_r)^2\}-\mathbb{E}\{(X_s)^2\}\mathbb{E}\{(X_r)^2\}\big]}{n^2}\to 0\ .$$

Let $\mu_p=\mathbb{E}\{(X_G)^p\}$ where $X_G\sim N(0,1)$. Then, for all $p\in\mathbb{N}^*$,

$$\mathbb{E}\Big\{\Big(\frac{X_1+X_2+.....+X_n}{\sqrt{(N_2+\sigma^2)n}}\Big)^p\Big\}\to\mu_p\ \ as\ n\to\infty$$

if and only if, for all $p\in\mathbb{N}^*$,

$$p!\frac{\sum_{t_1=1}^n\sum_{t_2=t_1+1}^n....\sum_{t_p=t_{p-1}+1}^n\mathbb{E}\{X_{t_1}X_{t_2}.....X_{t_p}\}}{n^{p/2}}\to(N_2)^p\mu_p\ .$$

From this type of results we have deduced CLT with minimal assumptions whose the conditions are close to strong mixing assumptions. In particular, one has defined condition H_{mI} which is close to minimal assumptions : cf [14] .

For the convergence in distribution, equivalent conditions H_I have been obtained : cf [15]. They are close to minimal assumptions. Then, we use these conditions in order to be sure that the CLT holds for some of our datas.

We recall them now.

At, first, one decomposes $X_1 + X_2 + ... + X_n$ in $X_1 + X_2 + ... + X_u$, $X_{u+1} + X_{u+2} + ... + X_{u+t}$ and $X_{u+t+1} + X_{u+t+2} + ... + X_{u+t+u}$ where u=u(n) and t=t(n). In this purpose, one uses the following notations

Notations 5.1.3 *We denote by $\kappa(n) \in \mathbb{N}$, an increasing sequence such that $\kappa(1) = 0$, $\kappa(n) \leq n$ and $\kappa(n)/n \to 0$. We define the sequences u(n) and t(n) by : u(1)=1, $u(n) = max\big\{m \in \mathbb{N}^* \big| 2m + \kappa(m) \leq n\big\}$, and t(1)=0, t(n) = n-2u(n) if $n \geq 2$.*

Notations 5.1.4 *Let $\sigma(u)^2 = \mathbb{E}\{(X_1 + X_2 + ... + X_u)^2\}$. One sets*

$$S_u = \frac{X_1 + X_2 + ... + X_u}{\sigma(u)} \ , \ \xi_u = \frac{X_{u+1} + X_{u+2} + ... + X_{u+t}}{\sigma(u)}$$

$$and \quad S'_u = \frac{X_{u+t+1} + X_{u+t+2} + ... + X_{u+t+u}}{\sigma(u)} \ .$$

Then, one can define almost minimal assumptions for the convergence of moments.

Notations 5.1.5 *: Let $k \in \mathbb{N}^*$. We define conditions $H_{mS}(k)$ and $H_{mI}(k)$ by the following way :*
$H_{mS}(k)$: $\forall p \in \mathbb{N}$, $p < k + 1$, $E\big\{(S_u)^p\big\} - E\big\{(S'_u)^p\big\} \to 0$ as $n \to \infty$.
$H_{mI}(2k)$: $\forall (p, q) \in (\mathbb{N}^*)^2$, $p + q < k + 1$,

$$E\big\{(S_u)^p (S'_u)^q\big\} - E\big\{(S_u)^p\big\} E\big\{(S'_u)^q\big\} \to 0$$

as $n \to \infty$.

Equivalent conditions can be defined for the convergence in distribution.

Notations 5.1.6 *: We define condition H_S and H_I by the following way.*

$$H_S : \forall k \in \mathbb{N}, \forall j \in \mathbb{N}, P\{A_{k,j}\} - P\{B_{k,j}\} \to 0 \quad as \quad n \to \infty \ ,$$

$$H_I : \forall k \in \mathbb{N}, \forall (j, j') \in \mathbb{N}^2, P\{A_{k,j} \cap B_{k,j'}\} - P\{A_{k,j}\} P\{B_{k,j'}\} \to 0$$

as $n \to \infty$, where $A_{k,j} = \big\{S_u \in I_{k,j}\big\}$ and $B_{k,j} = \big\{S'_u \in I_{k,j}\big\}$ with $I_{k,j} = \big[j.4^{-k}, (j + 1)4^{-k}\big[\big\}$.

Then the following CLT holds : cf [14] [15].

Theorem 3 *We assume that $H_{mS}(\infty)$ and $H_{mI}(\infty)$ hold. We assume also that, for all $p \in \mathbb{N}^*$, $\mathbb{E}\{(\xi_u)^p\} \to 0$ as $n \to \infty$.*

Then, $S_n \xrightarrow{D} N(0,1)$.

Theorem 4 *We assume that H_S , H_I , $H_{mS}(4)$ and $H_{mI}(4)$ hold. We assume also that $\mathbb{E}\{(S_u)^2\} - \mathbb{E}\{(S'_u)^2\} \to 0$ and $\mathbb{E}\{\xi_u^2\} \to 0$ as $n \to \infty$.*

Then, $S_n \xrightarrow{D} N(0,1)$.

It is this CLT that we use with our datas (cf chapter 10).

5.2 XOR Limit Theorem

Our secund limit theorem is based on the property of XOR. But it holds also modulo m. Then, by misusing of language, we call this this result "XOR Limit Theorem" (XORLT).

5.2.1 Presentation

In order to build the sequences of random bits $b^1(n')$ in section 11.2, we suppose that our datas are asymptotically independent. But, we are not sure that it is the case for all datas, for example datas obtained by some informatic files. In this case, the CLT cannot be used. Now, the XORLT holds under many weaker hypotheses.

The XORLT does not use the sums

$$\frac{X_1 + X_2 + ... + X_n}{\sigma(n)}$$

but the sums $\overline{X_1 + X_2 + ... + X_n}$ modulo 1 (in this section, $\overline{X_1 + X_2 + ... + X_n} \equiv X_1 + X_2 + ... + X_n$ modulo 1).

Definition 5.2.1 *Let $(X_n^1, X_n^2, ...X_n^p) \in \mathbb{R}^{\mathbf{P}}$ be a sequence of random vectors. For s=1,...,p, let $\sigma_s(n)^2 = \mathbb{E}\{(X_1^s + ... + X_n^s)^2\}$. Then, we set*

$$S_n^s = \frac{X_1^s + ... + X_n^s}{\sigma_s(n)} \ .$$

The XOR limit theorem holds for $(X_n^1, X_n^2, ...X_n^p)$ if there exists p sequences $\alpha_s(n) \to \infty$ as $n \to \infty$, such that $(\overline{\alpha_1(n)S_n^1}, ..., \overline{\alpha_p(n)S_n^p})$ has asymptotically the uniform distribution on $[0,1[^p$.

For the data used to obtain the sequence $b^1(n')$ in section 11.2, we always obtained that $\overline{X_1 + X_2 + ... + X_n}$ has asymptotically the uniform distribution on $[0, 1[$ i.e. $\alpha_(n) = 1/\sigma(n)$. This result is true for many datas.

We can find a first reason to it by considering the following example. Suppose that $\mathbb{E}\{X_1^2\} = 1 : S_n = \frac{X_1+...+X_n}{\sqrt{n}}$. Suppose that X_n is IID and has a continuous probability density function. Choose $n = k^2$ where $k \in \mathbb{N}^*$. Then, $X_1 + ... + X_n = \sqrt{n}S_n = kS_n$. Then, $\overline{X_1 + ... + X_n} = \overline{\sqrt{n}S_n} = \overline{kS_n}$ modulo 1.

Then, let us observe the behavior of the $\overline{kS_n}$'s as soon as k is enough big. The y's such that $x = \overline{ky}$ are distributed in a way close to uniformity in [0,1] : for example assume k=10. Then, if x=0.1, y=0.01 ,0.11, 0.21,....., 0.91. Now, by the CLT, S_n has a probability density function f such that $f(x) \approx (1/\sqrt{2\pi})e^{-x^2/2}$. Then, $P\{\overline{kS_n} \in [x, x + e[\} \to e : \overline{kS_n}$ is asymptotically uniformly distributed.

Now we recall the following theorem.

Theorem 5 *Let X and Y be two independent random vectors, $X,\ Y \in F^*(m)^p$. Assume that X has the uniform distribution. Then, $\overline{X + Y} \in F^*(m)^p$ has also the uniform distribution.*

Proof The following equalities hold

$$P\{X + Y = k\} = \sum_{t \in F^*(m)^p} P\big\{\{\overline{X + Y} = k\} \cap \{Y = t\}\big\}$$

$$= \sum_{t \in F(m)^p} P\big\{\{\overline{X + t} = k\} \cap \{Y = t\}\big\}$$

$$= \sum_{t \in F^*(m)^p} P\big\{\{X = \overline{k - t}\} \cap \{Y = t\}\big\}$$

$$= \sum_{t \in F^*(m)^p} P\{X = \overline{k - t}\}P\{Y = t\}$$

$$= \frac{1}{m^p} \sum_{t \in F^*(m)^p} P\{Y = t\} = \frac{1}{m^p} \ . \blacksquare$$

One deduces the following corollaries .

Corollary 5.2.1 *Let $X_n \in [0, 1[$ be a sequence of random variables. Assume that X_1 has the uniform distribution and that X_1 is independent of $(X_2, ..., X_n)$. Then $\overline{X_1 + X_2 + ... + X_n}$ has the uniform distribution on [0,1[.*

Corollary 5.2.2 *Let $X_n \in [0, 1[$ be an IID sequence of uniform random variables. Then $\overline{X_1 + X_2 + ... + X_n}$ has the uniform distribution on [0,1[.*

In this case $\sigma(n)^2 = \sigma^2 n$. Moreover, $\overline{\sigma\sqrt{n}S_n}$ has asymptotically the uniform distribution. Then, the CLT and the XORLT hold for the sequence X_n.

This result is general : if the CLT is satisfied, then the XORLT is satisfied.

The XORLT is more general than the CLT

Indeed, the following result holds.

Proposition 5.2.3 *: Let X_n be a sequence of random variables defined over a probability space. Assume that $\mathbb{E}\{X_n\} = 0$ and that $S_n = \frac{X_1+....+X_n}{\sigma(n)} \xrightarrow{D} S$ with $\mathbb{E}\{S^2\} = 1$. Assume that S has a probability density function f with respect to the Lebesgue measure such that $|f(x) - f(x')| \leq K_0|x - x'|$.*

Then, there exists a sequence $\alpha(n) \to \infty$ as $n \to \infty$ such that, for all $0 \leq t \leq 1$, $P\{\overline{\alpha(n)S_n} \in [0, t[\} \to t$ as $n \to \infty$.

Proof : Let $0 \leq t \leq 1$. For all $K \in \mathbb{N}^*$, let

$$E_K^n = \Big\{ \cup_{k=-K}^{K-1} \big\{ S_n \in \big[\frac{k}{K^{3/4}}, \frac{k+t}{K^{3/4}} \big[\big\} \Big\}$$

$$L_K = \Big\{ \cup_{k=-K}^{K-1} \big\{ S \in \big[\frac{k}{K^{3/4}}, \frac{k+t}{K^{3/4}} \big[\big\} \Big\} .$$

Then, $P\{E_K^n\} = P\{L_K\} + \epsilon_n^K$, where $\epsilon_n^K \to 0$ as $n \to \infty$.

Now, let $k \in \{-K, -K+1, ..., K-1\}$. Then, there exists $x_0 \in \big[\frac{k}{K^{3/4}}, \frac{k+t}{K^{3/4}} \big[$ [1] such that

$$\int_{[k/K^{3/4},(k+t)/K^{3/4}[} f(x)dx = \int_{[k/K^{3/4},(k+t)/K^{3/4}[} \big[f(x_0) + \frac{Ob(1)K_0 t}{K^{3/4}} \big] dx$$

$$= \frac{t.f(x_0)}{K^{3/4}} + \frac{Ob(1)K_0 t^2}{K^{3/2}} .$$

Then,

$$\frac{f(x_0)}{K^{3/4}} = \int_{[k/K^{3/4},(k+1)/K^{3/4}[} f(x)dx + \frac{Ob(1)K_0}{K^{3/2}} .$$

Then,

$$\int_{[k/K^{3/4},(k+t)/K^{3/4}[} f(x)dx = t \int_{[k/K^{3/4},(k+1)/K^{3/4}[} f(x)dx + \frac{2Ob(1)K_0}{K^{3/2}} .$$

[1] One can choose $x_0 = k/K^{3/4}$.

Therefore,

$$P\big\{S \in [k/K^{3/4}, (k+t)/K^{3/4}[\big\} = t.P\big\{S \in [k/K^{3/4}, (k+1)/K^{3/4}[\big\} + \frac{2Ob(1)K_0}{K^{3/2}} .$$

Then,

$$P\{E_K^n\} = P\{L_K\} + \epsilon_n^K = \sum_{k=-K}^{K-1} P\big\{S \in [k/K^{3/4}, (k+t)/K^{3/4}[\big\} + \epsilon_n^K$$

$$= \sum_{k=-K}^{K-1} t.P\big\{S \in [k/K^{3/4}, (k+1)/K^{3/4}[\big\} + \sum_{k=-K}^{K-1} \frac{2Ob(1)K_0}{K^{3/2}} + \epsilon_n^K$$

$$= t.P\big\{S \in [-K/K^{3/4}, K/K^{3/4}[\big\} + (2K)\frac{2Ob(1)K_0}{K^{3/2}} + \epsilon_n^K$$

$$= t.P\big\{S \in [-K^{1/4}, K^{1/4}[\big\} + \frac{4Ob(1)K_0K}{K^{3/2}} + \epsilon_n^K$$

$$= t - tP\big\{S \notin [-K^{1/4}, K^{1/4}[\big\} + \frac{4.Ob(1)K_0}{K^{1/2}} + \epsilon_n^K$$

$$= t + \frac{t.Ob(1)}{K^{1/2}} + \frac{4.Ob(1)K_0}{K^{1/2}} + \epsilon_n^K \ ,$$

by the Bienayme-Tchebycheff Inequality.

Let $F_n = \Big\{\cup_{k\in\mathbb{Z}}\big\{S_n \in \big[\frac{k}{K^{3/4}}, \frac{k+t}{K^{3/4}}\big[\big\}\Big\}$. Then, by the Bienayme-Tchebycheff Inequality, $P\{F_n\} = P\{E_K^n\} + \frac{Ob(1)}{K^{1/2}}$.

Then, $P\{F_n\} = t + \frac{t.Ob(1)}{K^{1/2}} + \frac{4.Ob(1)K_0}{K^{1/2}} + \epsilon_n^K + \frac{Ob(1)}{K^{1/2}}$.

One chooses a sequence $K_n \to \infty$ such that $\epsilon_n^{K_n} \to 0$ as $n \to \infty$. We deduce that $P\{F_n\} = t + \epsilon'_n$ where $\epsilon'_n \to 0$ as $n \to \infty$.

Now, $F_n = \Big\{ \cup_{k\in\mathbb{Z}} \big\{K_n^{3/4} S_n \in \big[k, k+t\big[\big\}\Big\} = \big\{\overline{K_n^{3/4} S_n} \in \big[0, t\big[\big\}$.

Then, $P\big\{\overline{K_n^{3/4} S_n} \in \big[0, t\big[\big\} \to t$ as $n \to \infty$. ■

Remark that analog results hold also for random vectors $(S_1^n, S_2^n, ..., S_p^n) \in \mathbb{R}^p$.

In general, $\sigma(n)\alpha(n) = 1$: i.e. $\overline{(X_1 + ...X_n)}$ has asymptotically the uniform distribution on $[0, 1[$. For example, the proposition 5.2.4 show a such result .

In order to state this proposition, the measure μ'_n is needed.

Notations 5.2.2 *We denote by* μ'_n *the measure defined on* $\{\frac{k}{m\sigma(n)}|k \in \mathbb{Z}\}$ *by* $\mu'_n(\frac{k}{m\sigma(n)}) = \frac{1}{m\sigma(n)}$ *for all* $k \in \mathbb{Z}$.

Then, μ'_n is the measure equivalent at the Lebesgue measure in the discrete case. For example the following property holds.

Lemma 5.2.1 *The following limit holds :* $\mu'_n([0,1]) \to 1$ *as* $n \to \infty$.

Proof We have $\mu'_n([0,1]) = \sum_{0\leq k\leq m\sigma(n)} \mu'_n\{\frac{k}{m\sigma(n)}\} = \sum_{0\leq k\leq \lfloor m\sigma(n)\rfloor} \frac{1}{m\sigma(n)}$ $= [\lfloor m\sigma(n)\rfloor + 1]\frac{1}{m\sigma(n)} = \frac{\lfloor m\sigma(n)\rfloor}{m\sigma(n)} + \frac{1}{m\sigma(n)} = 1 - \frac{|Ob(1)|}{m\sigma(n)} + \frac{1}{m\sigma(n)} = 1 + \frac{Ob(1)}{m\sigma(n)}$. ■

Then the following XORLT holds.

Proposition 5.2.4 *Let* $(S_1^n, S_2^n, ..., S_p^n) \in \mathbb{R}^{\mathbf{p}}$ *be a random vector such that* $E\{(S_n^s)^2\} = 1$ *for s=1,2,...,p.*

Let μ^A *be a measure on* $\mathbb{R}$ *: one assumes that* $\mu^A = \mu^1 \otimes .... \otimes \mu^p$ *where* $\mu^s = \mu$ *the Lebesgue measure for s=1,...,p or where* $\mu^s = \mu'_n$ *for s=1,..,p. Assume that* $(S_1^n, S_2^n, ..., S_p^n)$ *has a probability density function* f_n *with respect to* μ^A *such that* $|f_n(x_1, ..., x_p) - f_n(x'_1, .., x'_p)| \leq K_0 max(|x_s - x'_s|)$.

Let $\alpha(n)$ *be a sequence such that* $\alpha(n) \to \infty$ *as* $n \to \infty$.

Then $\overline{\alpha(n)(S_1^n, S_2^n, ..., S_p^n)}$ *has asymptotically the uniform distribution over* $[0,1[^p$.

Proof 5.2.2 *We begin to prove the proposition 5.2.4*

Let $t_s \in [0,1[$ for s=1,...,p. If $\mu^s = \mu'_n$ for s=1,..,p, one assumes that $t_s = \frac{h_s}{m}$ where $h_s \in \{0, 1, ..., m\}$ for s=1,2,...p. Then,

$$P\{\overline{\alpha(n)(S_n^1, ..., S_n^p)} \in [0, t_1[\otimes ..... \otimes [0, t_p[\}$$

$$= P\Big\{\alpha(n)(S_n^1, ..., S_n^p) \in \bigcup_{k_1\in\mathbb{Z}} .... \bigcup_{k_p\in\mathbb{Z}} [k_1, k_1 + t_1[\otimes ..... \otimes [k_p, k_p + t_p[\Big\}$$

$$= \sum_{k_1\in\mathbb{Z}} .... \sum_{k_p\in\mathbb{Z}} P\Big\{\alpha(n)(S_n^1, ..., S_n^p) \in [k_1, k_1 + t_1[\otimes ..... \otimes [k_p, k_p + t_p[\Big\}$$

$$= \sum_{k_1\in\mathbb{Z}} .... \sum_{k_p\in\mathbb{Z}} P\Big\{(S_n^1, ..., S_n^p) \in [\frac{k_1}{\alpha(n)}, \frac{k_1 + t_1}{\alpha(n)}[\otimes ..... \otimes [\frac{k_p}{\alpha(n)}, \frac{k_p + t_p}{\alpha(n)}[\Big\}$$

$$= \sum_{k_1} .... \sum_{k_p} \int_{[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)}[} .... \int_{[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)}[} f_n(x_1, ...x_p)\mu^A(dx_1, ..., dx_p) \ . \ ■$$

In order to continue the proof of proposition 5.2.4 the following lemma is needed.

Lemma 5.2.3 *The following equality holds*

$$\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[}f_n(x_1,...x_p)\mu^A(dx_,1..,.dx_p)$$

$$= t_1...t_p\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+1}{\alpha(n)}\right[}f_n(x_1,...x_p)\mu^A(dx_,1..,.dx_p)+\frac{2Ob(1)K_0}{\alpha(n)^{p+1}}\ .$$

Proof : First, suppose that $\mu=\mu'_n$. Then, one assumes that $t_s=\frac{h_s}{m}$. Let $s\in\{1,..,p\}$. Then,

$$\Big[\frac{k_s}{\sigma(n)},\frac{k_s+t_s}{\sigma(n)}\Big[\cap\Big\{0,\frac{1}{m\sigma(n)},\frac{2}{m\sigma(n)},....\Big\}$$

$$=\Big\{\frac{k_sm}{m\sigma(n)},\frac{k_sm+1}{m\sigma(n)},....,\frac{k_sm+h_s-1}{m\sigma(n)}\Big\}\ .$$

Indeed

$$\frac{k_s+t_s}{\sigma(n)}=\frac{k_sm+t_sm}{m\sigma(n)}=\frac{k_sm+h_s}{m\sigma(n)}\ .$$

Then,

$$card\Big(\Big[\frac{k_s}{\sigma(n)},\frac{k_s+t_s}{\sigma(n)}\Big[\cap\Big\{0,\frac{1}{m\sigma(n)},\frac{2}{m\sigma(n)},....\Big\}\Big)=h_s\ .$$

Then, $\int_{[k_s/\sigma(n),(k_s+t_s)/\sigma(n)[}\mu'_n(dx)=\frac{h_s}{m\sigma(n)}=\frac{t_s}{\sigma(n)}$.

Now, if μ is the Lebesgue measure, $\int_{[k_s/\sigma(n),(k_s+t_s)/\sigma(n)[}\mu(dx)=\frac{t_s}{\sigma(n)}$ obviously.

Then, in all cases

$$\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[}\mu^A(dx_1,..,.dx_p)=\frac{t_1.....t_p}{\sigma(n)^p}\ .$$

Let $(x_1,...,x_p)$ and $(x'_1,...,x'_p)\in\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[\otimes....\otimes\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[$. Then, $|x_s-x'_s|\leq\frac{1}{\alpha(n)}$ for s=1,2,...,p. Then, $|f_n(x_1,...,x_p)-f_n(x'_1,...,x'_p)|\leq\frac{K_0}{\alpha(n)}$.

Let $x_0^s=k_s/\sigma(n)$ for s=1,...,p. Then,

$$\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[}f_n(x_1,...x_p))\mu^A(dx_1,..,.dx_p)$$

$$=\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[}f_n(x_1^0,...x_p^0)\mu^A(dx_1..,.dx_p)$$
$$+\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[}\frac{Ob(1)K_0}{\alpha(n)}\mu^A(dx_1,..,.dx_p)$$

$$=f_n(x_1^0,...x_p^0)\frac{t_1...t_p}{\alpha(n)^p}+\frac{Ob(1)K_0}{\alpha(n)}\ \frac{t_1...t_p}{\alpha(n)^p}\ .$$

Then,

$$\frac{f_n(x_1^0,...x_p^0)}{\alpha(n)^p}$$
$$=\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+1}{\alpha(n)}\right[}f_n(x_1,...x_p)\mu^A(dx_1..,.dx_p)+\frac{Ob(1)K_0}{\alpha(n)^{p+1}}\ .$$

Then,

$$\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+t_1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+t_p}{\alpha(n)}\right[}f_n(x_1,...x_p)\mu^A(dx_1..,.dx_p)$$

$$=f_n(x_1^0,...x_p^0)\frac{t_1...t_p}{\alpha(n)^p}+\frac{Ob(1)K_0}{\alpha(n)}\ \frac{t_1...t_p}{\alpha(n)^p}\ .$$

$$=t_1...t_p\int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+1}{\alpha(n)}\right[}....\int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+1}{\alpha(n)}\right[}f_n(x_1,...x_p)\mu^A(dx_1..,.dx_p)+\frac{2Ob(1)K_0}{\alpha(n)^{p+1}}\ .\ \blacksquare$$

Proof 5.2.4 *Now, we continue the proof of the proposition 5.2.4*

Soit $a\in\mathbb{R}_+$. By Bienaymé-Tchebycheff Inequality, $P\{|S_n^s-e_n^s|\geq a\}\leq\frac{1}{a^2}$, where $e_n^s=\mathbb{E}\{S_n^s\}$ for s=1,...,p. Moreover, by Schwarz Inequality, $|e_n^s|^2\leq\mathbb{E}\{(S_n^s)^2\}=1$.

Let $A_n\to\infty$ as $n\to\infty$. There exists a sequence $\epsilon_{A_n}\to 0$ as $n\to\infty$ such that $P\{|S_n^s-e_n^s|\geq A_n\}\leq\epsilon_{A_n}$.

Then,

$$\sum_{k_1} .... \sum_{k_p} \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)}\right[} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$

$$= \sum_{(k_1,...,k_p)\in E_{A_n}} \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)}\right[} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$

$$+ \sum_{(k_1,...,k_p)\notin E_{A_n}} \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)}\right[} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p) \ ,$$

where

$$E_A = \Big\{(n_1, ..., n_p) \in \mathbb{Z}^p \ \Big| \ |\frac{n_s}{\alpha(n)} - e_n^s| \le A, \ and \ |\frac{n_s + t_s}{\alpha(n)} - e_n^s| \le A, \ \forall s \in \{1, .., p\}\Big\}.$$

Then, $(k_1, ..., k_p) \in E_A$ if and only if $-A + e_n^s - \frac{t_s}{\alpha(n)} \le \frac{k_s}{\alpha(n)} \le A + e_n^s - \frac{t_s}{\alpha(n)}$ and $-A + e_n^s \le \frac{k_s}{\alpha(n)} \le A + e_n^s$ for s= 1,...,p.

Then, $(k_1, ..., k_p) \in E_A$ if and only if $-A + e_n^s \le \frac{k_s}{\alpha(n)} \le A + e_n^s - \frac{t_s}{\alpha(n)}$ for s= 1,...,p.

Then, if $(k_1, ..., k_p) \notin E_A$, there exists $s \in \{1, 2, ..., p\}$ such that $\frac{k_s}{\alpha(n)} < -A + e_n^s$ or $A + e_n^s - \frac{t_s}{\alpha(n)} < \frac{k_s}{\alpha(n)}$.

Then, if $(k_1, ..., k_p) \notin E_A$, there exists $s \in \{1, 2, ..., p\}$ such that $k_s \in B_s(A)$ where $B_s(A) = \Big\{k \in \mathbb{Z} \Big| \frac{k}{\alpha(n)} < -A + e_n^s \ or \ A + e_n^s - \frac{t_s}{\alpha(n)} < \frac{k}{\alpha(n)}\Big\}$.

Then,

$$\sum_{(k_1,...,k_p)\notin E_{A_n}} \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)}\right[} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$

$$\le \sum_{s=1}^{p} \sum_{k_s\in B_s(A_n), k_t\in\mathbb{Z}, t\neq s} \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)}\right[} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$

$$\le \sum_{s=1}^{p} \sum_{k_s\in B_s(A_n)} \int_{\left[\frac{k_s}{\alpha(n)}, \frac{k_s+t_s}{\alpha(n)}\right[} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$

$$\leq \sum_{s=1}^{p} \int_{x_s \leq -A_n + e_n^s + \frac{t_s}{\alpha(n)}} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$
$$+ \sum_{s=1}^{p} \int_{A_n + e_n^s - \frac{t_s}{\alpha(n)} \leq x_s} f_n(x_1, ...x_p)\mu^A(dx_1.., .dx_p)$$

$$= \sum_{s=1}^{p} \Big[P\big\{ S_n^s \leq -A_n + e_n^s + \frac{t_s}{\alpha(n)} \big\} + P\big\{ A_n + e_n^s - \frac{t_s}{\alpha(n)} \leq S_n^s \big\} \Big]$$

$$\leq p\epsilon'_n \ ,$$

where $\epsilon'_n \to 0$ as $n \to \infty$ because $\big| \frac{t_s}{\alpha(n)} \big| \leq 1$.

Now, $(k_1, ..., k_p) \in E_{A_n}$ if and only if, for s= 1,...,p,

$$-A_n\alpha(n) + e_n^s\alpha(n) \leq k_s \leq A_n\alpha(n) + e_n^s\alpha(n) - t_s \ .$$

Let $s \in \{1, ..., p\}$. Let

$$N_s = card\big\{ k \in \mathbb{Z} \big| - A_n\alpha(n) + e_n^s\alpha(n) \leq k \leq A_n\alpha(n) + e_n^s\alpha(n) - t_s \big\} \ .$$

Then, for s= 1,...,p, $N_s \leq 2A_n\alpha(n) + 1$.

Then,

$$P\big\{ \overline{\alpha(n)(S_n^1, ..., S_n^p)} \in [0, t_1[\otimes ..... \otimes [0, t_p[\big\}$$

$$= \sum_{(k_1,...,k_p) \in E_{A_n}} \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+t_1}{\alpha(n)} \right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+t_p}{\alpha(n)} \right[} f_n(x_1, ...x_p)\mu^A(dx_1, .., .dx_p) + Ob(1)p\epsilon'_n$$

$$= \sum_{(k_1,...,k_p) \in E_{A_n}} t_1.....t_p \int_{\left[\frac{k_1}{\alpha(n)}, \frac{k_1+1}{\alpha(n)} \right[} .... \int_{\left[\frac{k_p}{\alpha(n)}, \frac{k_p+1}{\alpha(n)} \right[} f_n(x_1, ...x_p)\mu^A(dx_1, .., .dx_p)$$
$$+ \sum_{(k_1,...,k_p) \in E_{A_n}} \frac{2K_0 Ob(1)}{\alpha(n)^{p+1}} + Ob(1)p\epsilon'_n$$

$$= \sum_{(k_1,...,k_p)\in E_{A_n}} t_1.....t_p \int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+1}{\alpha(n)}\right[} f_n(x_1,...x_p)\mu^A(dx_1..,.dx_p)$$

$$+(2A_n\alpha(n)+1)^p \frac{2K_0 Ob(1)}{\alpha(n)^{p+1}} + Ob(1)p\epsilon'_n$$

$$= \sum_{(k_1,...,k_p)} t_1.....t_p \int_{\left[\frac{k_1}{\alpha(n)},\frac{k_1+1}{\alpha(n)}\right[} .... \int_{\left[\frac{k_p}{\alpha(n)},\frac{k_p+1}{\alpha(n)}\right[} f_n(x_1,...x_p)\mu^A(dx_1,..,.dx_p) + Ob(1)p\epsilon'_n$$

$$+2^{p+1}K_0 Ob(1)\big(1+\frac{1}{2A_n\alpha(n)}\big)^p \frac{(A_n)^p}{\alpha(n)} + \ Ob(1)p\epsilon'_n$$

$$= t_1.....t_p + 2^{p+1}K_0 Ob(1)\big(1+\frac{1}{2A_n\alpha(n)}\big)^p \frac{(A_n)^p}{\alpha(n)} + 2Ob(1)p\epsilon'_n \ .$$

Then, $\frac{(A_n)^p}{\alpha(n)} \to 0$ as $n \to \infty$ is supposed. Then,

$$P\big\{\overline{\alpha(n)(S_n^1,....,S_n^p)} \in [0,t_1[\otimes..... \otimes [0,t_p[\big\} \to t_1...t_p \quad as\ n \to \infty \ . \ \blacksquare$$

In particular, assume that p=1. Suppose that $X_s \in F(m)$. Suppose that $|f_n(x) - f_n(x')| \leq K_0|x-x'|$. Then, $P\big\{\overline{X_1 + ... + X_n} \in [0,t[\big\} \to t$ as $n \to \infty$ where $t \in F(m)$.

Now, if $X_s \in F(m)$, there is always a probability density function f_n with respect to μ'_n such that $\exists K_n \ : \ |f_n(x) - f_n(x')| \leq K_n|x-x'|$.

Then, the condition " $\exists K_0 \ : \ |f_n(x) - f_n(x')| \leq K_0|x-x'| \ \forall \ n \in \mathbb{N}^*$ " is not a necessary condition of the CLT. Then, in some cases, the hypotheses of proposition 5.2.4 are stronger than those of the CLT.

Generally, condition " $|f_n(x) - f_n(x')| \leq K_0|x-x'| \ \forall \ n \in \mathbb{N}^*$ " holds: one has examples for some distributions in [43], e.g. page 54. As a matter of fact, under the assumptions of datas studied in this report, we have never found a single case where it is not verified.

If we want that it is not verified, it is necessary to build specially X_n for it. For example one can choose a sequence X'_n such that densities f_n satisfy

$max(f_n) \to \infty$: then, one builds a sequence X_n such that $S_n = X'_n : X_1 = X'_1$, $X_2 = X'_2 - X'_1$, $X_3 = X'_3 - X'_2$....

Now proposition 5.2.4 suggests that if the CLT holds, then, $\overline{X_1 + ... + X_n}$ has asymptotically the uniform distribution.

It shows also that the XORLT is more general than the CLT : for example, the XORLT holds if $X_s = X_1$ for all s : $\overline{X_1 + ... + X_n} = \overline{nX_1}$.

Corollary 5.2.5 *let X be a random variable. Assume that X has a continuous probability density function f. Then,* $\overline{nX}$ *has asymptotically the uniform distribution.*

Conclusion

We write $\overline{\alpha(n)S_n} = \overline{\beta(n)(X_1 + X_2 + ... + X_n)}$. Then, it seems that the XORLT holds under assumptions weaker than the CLT. Let us notice that it is even sometimes verified by the classic counterexample $X_n = Y_{n+1} - Y_n$ where Y_n is IID.

Moreover, generally, it holds with $\beta(n) = 1$,i.e. $\overline{X_1 + X_2 + ... + X_n} \xrightarrow{D} U$ where U has the uniform distribution. At last, under the hypotheses of our data, we did not find a single case where it is not verified.

5.3 Examples

In this section, we compare the limit distributions. In these examples we shall note the strength of the XORLT.

Let $S_n^2 \in \mathbb{R}^2$ such that $S_n^2 \xrightarrow{D} S_0^2$ where $S_0^2 \sim N_2(0, C)$ when C is a covariance matrix. One knows that $g(S_n^2) \xrightarrow{D} g(S_0^2)$ if g is continuous with $P_{S_0^2}$ probability 1 (cf [42] page 24). Then, $\overline{S_n^2} \xrightarrow{D} \overline{S_0^2}$. Moreover, we shall note that the dependence of S_0^2 does not exist any more for $\overline{S_0^2}$. We shall deduce the XORLT for $\sigma(n)\overline{S_n^2}$.

5.3.1 Example 1

In this section we study the following example

Example 5.3.1 *Let X and Y be two independent random variable with distribution N(0,1). Let* $Z = \frac{X+aY}{\sqrt{1+a^2}}$ *where* $a \in \mathbf{R}$.

Test of the linear correlation coefficient Under the previous hypotheses, Z has the N(0,1) distribution. Moreover the linear correlation coefficient of X and Z is $\rho = E\{XZ\} = (1+a^2)^{-1/2}E\{X(X+aY)\} = (1+a^2)^{-1/2}$. For example, $\rho = 0.701$ si a=1.

let ρ_n be the empirical linear correlation coefficient associated to a sample (X_s, Z_s). Let ρ_n^U be the empirical linear correlation coefficient associated to the sample $(\overline{X_s}, \overline{Z_s})$.

Then, ρ_n et ρ_n^U allow us to estimate the linear correlation coefficients of (X_0, Z_0) and $(\overline{X_0}, \overline{Z_0})$.

Let N be the size of the sample. The following results have been obtained

ρ	N	ρ_n	ρ_n^U	ρ_n	ρ_n^U
0.7071	1000	0.7063	-0.0607	0.6941	0.0367
0.4472	1000	0.4597	0.0017	0.4488	-0.0260
0.2425	1000	0.2472	0.0054	0.2167	-0.0252
0.7071	5000	0.7034	-0.0294	0.6996	-0.0012
0.4472	5000	0.4536	0.0002	0.4436	-0.0270
0.2425	5000	0.2351	0.0075	0.2290	0.0216
0.7071	10000	0.7108	0.0061	0.7107	0.0010
0.4472	10000	0.4469	-0.0020	0.4454	-0.0049
0.2425	10000	0.2675	0.0099	0.2478	0.0101
0.7071	100000	0.7074	-0.0011	0.7056	-0.0007
0.4472	100000	0.4433	-0.0013	0.4467	0.0002
0.2425	100000	0.2466	-0.0037	0.2445	-0.0015

Then ρ_n^U is smaller than ρ_n. As a matter of fact, if we do tests, we can even consider that ρ_n^U is the estimate of the correlation coefficient equal to 0.

Indeed, let ρ_n be a empirical linear correlation coefficient associated to an IID sample. Then, by the CLT, $P\left\{\sqrt{n}\rho_n \leq x\right\} \approx \Gamma(x)$ cf [10].

Chi squared independence test We test the independence of $\overline{X_n}$ and $\overline{Z_n}$ by the chi squared independence test.

We use a partition (15,15). The chi-squared statistics has asymptotically a normal distribution (cf proposition A.1.2) : $\sqrt{2\chi_2} - \sqrt{2d-1}$ where d is the degree of freedom : (15-1)(15-1).

Assume that the linear correlation coefficient is equal to 0.7071. Then, for some samples, the following result is obtained for $\sqrt{2\chi_2} - \sqrt{2d-1}$ (cf proposition A.1.2) .

1.1256	-0.1246	2.0030	-0.8977	-0.7952	0.6594	-0.7758
-0.3079	0.5618	-0.3380	-1.2630	-0.5369	-1.0617	0.9458
-1.6506	-0.3484	0.9821	-0.8853	-0.1215	-0.5373	1.0599

In fact $(\overline{X}, \overline{Z})$ is enough close to an independent vector.

Conclusion Under the previous hypotheses,

$$(\overline{X1 + ... + X_n}, \overline{Z_1 + ... + Z_n}) \to \overline{\sigma(n)\overline{(X, Z)}} \ .$$

Now $\overline{(X, Z)}$ is already close to an independent vector. Then, it will thus be even truer for $\overline{\sigma(n)\overline{(X, Z)}}$ because the multiplication by $\sigma(n)$ modulo 1 makes uniform the distribution as soon as $\sigma(n)$ is enough big.

In conclusion, the fact that (X, Z) is already almost independent shows the rate of the convergence of the XORLT.

5.3.2 Example 2

In this example we have similar results when the dimensions are larger than 2.

Let $(X_1, X_2, X_3, X_4, X_5)$ be a random vector which has a independent normal distribution. We are interested by

$U_1 = X_1$
$U_2 = X_1 + X_2$
$U_3 = X_1 + X_2 + X_3$
$U_4 = X_1 + X_2 + X_3 + X_4$
$U_5 = X_1 + X_2 + X_3 + X_4 + X_5$.

We use the chi squared independence test on $\mathbb{R}^5$. We use a statistics whose the distribution is close to the Gaussian one : $\sqrt{2\chi_2} - \sqrt{2d-1}$. We assume that we use hypercubes associated to a (5,5,5,5,5) partition and that the size of the sample is 100000. For some various samples, the following results have been obtained.

-0.5232	1.0150	0.6986	-1.8970	-0.7312	0.9638	0.0767
1.8270	0.1473	-0.3621	-1.1102	-0.7045	-1.1002	0.9371

These results shows that these random variables behave as if they were independent.

5.3.3 Example using datas of this report

We study an example using the datas of section 11.2.5.

We estimate by using histogram the probability density functions f_g and f_h associated to the sizes $G(j)$ and H(j) defined in section 11.2.5. We note that f_G has a graph in form of a bell : figure 5.1 . That corresponds to the CLT. Moreover, f_H has a graph close to that one of the uniform density : figure 5.2 : That corresponds to the XORLT

We note that the distances between the diverse values of the histogram are important. It occurs because we took samples of size 100 for a partition in 100 intervals. We so acted it to study closer possible the real probability in every point. If we take larger samples, we obtain a density very close to the uniform distribution.

As a matter of fact, when we studied numerically various examples using data of the type "text", "computer programs", "mathematical reports", etc., we always found that $\overline{X_1 + X_2 + ... + X_n}$ has asymptotically the uniform distribution.

We obtained results similar in several dimensions: for the data used in this report, we always found that $(\overline{X_{1,1} + X_{2,1} + ... + X_{n,1}}, \overline{X_{1,2} + X_{2,2} + ... + X_{n,2}})$ has asymptotically the uniform distribution on $[0,1]^p$ for p=2. We obtained similar results for p=3,4,5,6.

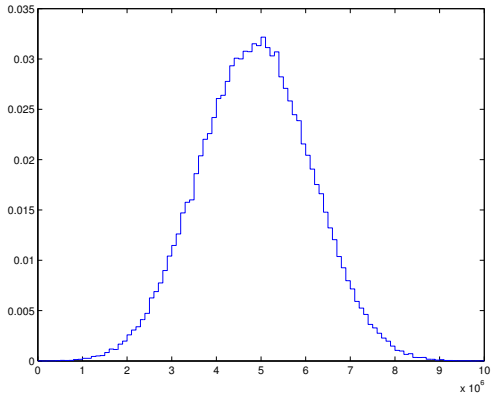

Figure 5.1: graph of f_g

5.3.4 Distributions close to the uniform distribution

This result is confirmed by considering the case of sum of 5 random variables $X_n \in \{0, 1, ..., q\}$ for example :

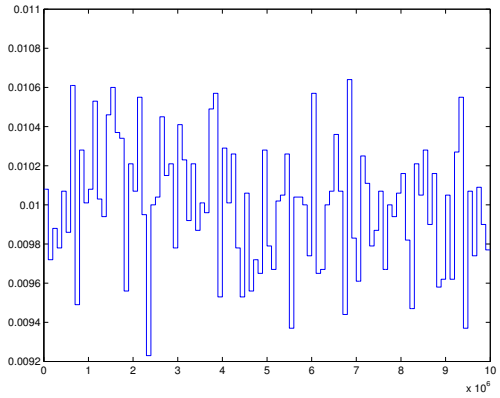

Figure 5.2: Graph of f_h

$$P\{\sum_{i=1}^{5} X_i = 0\}$$

$$= P\big\{\{X_1 = 0\} \cap \{X_2 = 0\} \cap \{X_3 = 0\} \cap \{X_4 = 0\} \cap \{X_5 = 0\}\big\}$$

$$P\{\sum_{i=1}^{5} X_i = 1\}$$

$$= P\big\{\{X_1 = 1\} \cap \{X_2 = 0\} \cap \{X_3 = 0\} \cap \{X_4 = 0\} \cap \{X_5 = 0\}\big\}$$

$$+P\big\{\{X_1 = 0\} \cap \{X_2 = 1\} \cap \{X_3 = 0\} \cap \{X_4 = 0\} \cap \{X_5 = 0\}\big\}$$

$$+.....$$

$$+P\big\{\{X_1 = 0\} \cap \{X_2 = 0\} \cap \{X_3 = 0\} \cap \{X_4 = 0\} \cap \{X_5 = 1\}\big\}$$

$$P\{\sum_{i=1}^{5} X_i = 2\}$$

$$= P\big\{\{X_1 = 2\} \cap \{X_2 = 0\} \cap \{X_3 = 0\} \cap \{X_4 = 0\} \cap \{X_5 = 0\}\big\}$$

$$
\begin{aligned}
&+.....\\
&+P\big\{\{X_1=0\}\cap\{X_2=0\}\cap\{X_3=0\}\cap\{X_4=0\}\cap\{X_5=2\}\big\}\\
&+P\big\{\{X_1=1\}\cap\{X_2=1\}\cap\{X_3=0\}\cap\{X_4=0\}\cap\{X_5=0\}\big\}\\
&+.....\\
&+P\big\{\{X_1=1\}\cap\{X_2=0\}\cap\{X_3=0\}\cap\{X_4=0\}\cap\{X_5=1\}\big\}\\
&+P\big\{\{X_1=0\}\cap\{X_2=1\}\cap\{X_3=1\}\cap\{X_4=0\}\cap\{X_5=0\}\big\}\\
&+.....\\
&+P\big\{\{X_1=0\}\cap\{X_2=1\}\cap\{X_3=0\}\cap\{X_4=0\}\cap\{X_5=1\}\big\}\ .
\end{aligned}
$$

And so on. Finally this type of equality shows well that the curve of the probability is smooth and has the form of bell.

More generally, let us take a sum of n terms. Let us take its values close of n/2, the middle of [0, n]: let us look at number $(x_1, ..., x_n)$ such as $x_1 + ... + x_n$ is close of n/2. There are much more of them than there are $(x_1, ..., x_n)$'s such as $x_1 + x_2 + ... + x_n$ is close to zero for example. We thus obtain a curve more and more in the form of bell which satisfies $|f_n(z) - f_n(z')| \leq K_n|x - x'|$ where $K_n \leq K_0$. It is thus one good curve (for the CLT).

It is the same if one uses the XORLT. Indeed, if we take the curve of $\overline{X_1 + X_2 + ... + X_n}$, this one smooth down and thus becomes uniform when $n \to \infty$. Then, the classic methods of integration implicate the convergence to the uniform distribution.

Another reason of convergence to the uniform distribution, it is because when we use independent random variables with uniform distribution, $\overline{X_1 + ... + X_n}$ has exactly the uniform distribution.

All this shows that if we are in an approached case, the density probability function of $\overline{X_1 + X_2 + ... + X_n}$ is very close to 1.

5.4 Numerical study

In this section, we study the rate of convergence de $X_1 + X_2 + ... + X_n$ and $\overline{X_1 + X_2 + ... + X_n}$ by numerical calculations.

For that purpose, we shall choose n varying between 3 and 20 and we suppose $X_s \in \{0, 1, ..., q\}$. In the following examples, the distributions of the $(X_1, ..., X_n)$'s are not independent, but chosen with dependences strong enough.

For these values, we can then notice that the graphs are about the ones of a normal distribution or a uniform distribution except when the probability are concentrated near a small number of points.

5.4.1 Case n=7 or n=8

On the following graphs we choose n=7 or n = 8 for various covariance matrices of $(X_1, X_2, ..., X_n)$.

Case general Generally we obtain curves close to those of the normal or uniform distributions : it is the case for the examples A1 and A2 : figures 5.3 , 5.4, 5.5 and 5.6 . It is not the case for the examples A3 whose some marginal probabilities are close to zero : $P\{X_s = 0\} = 0.35$, $P\{X_s = 1\} = 0$, $P\{X_s = 2\} = 0.25$, $P\{X_s = 3\} = 0.1$, $P\{X_s = 4\} = 0.3$: cf figures 5.7 and 5.8.

Indeed, the more n is big, the more curves are close to limit curves. In fact, the convergence is very fast. It confirms the results of the section 5.5.1.

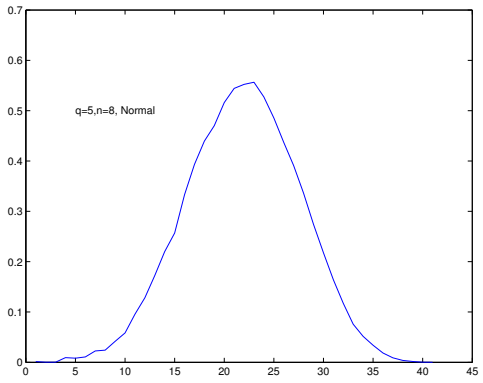

Figure 5.3: Example A1 : normal convergence

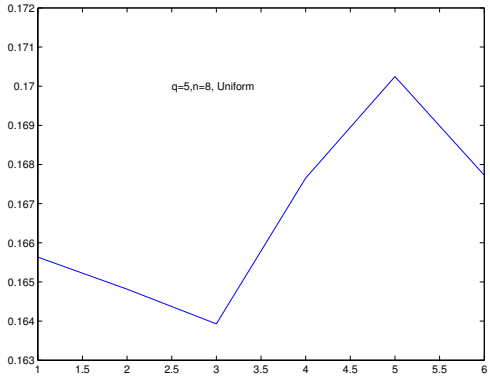

Figure 5.4: Example A1 : uniform convergence

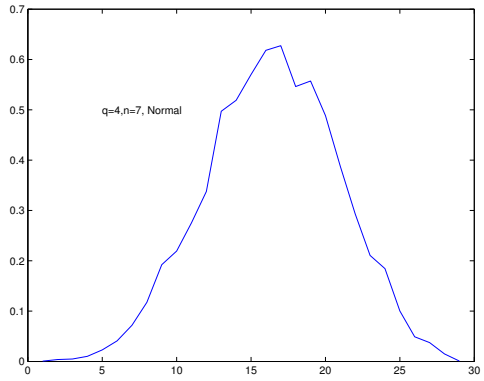

Figure 5.5: Example A2 : normal convergence

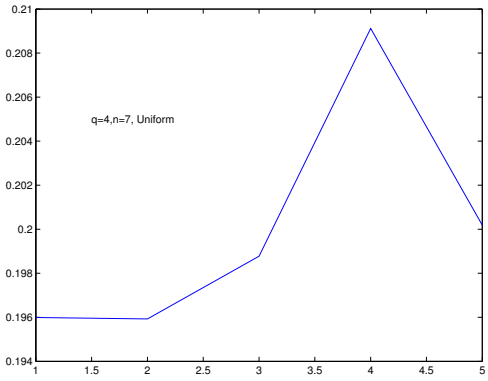

Figure 5.6: Example A2 : uniform convergence

Case q=3 In the figures 5.9 , 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16, we have choosen sums $X_1 + ... + X_7$ with $X_s \in \{0, 1, 2, 3\}$.

We understand that the curve of the probability is the most remote from a shape of bell when the marginal probabilities are concentrated near some points: example C3, probabilities 0.56, 0, 0.037, 0.07. This is normal: for example if the marginal probabilites is concentrated on even points, the probability of the sums is equal to 0 in the odd points.

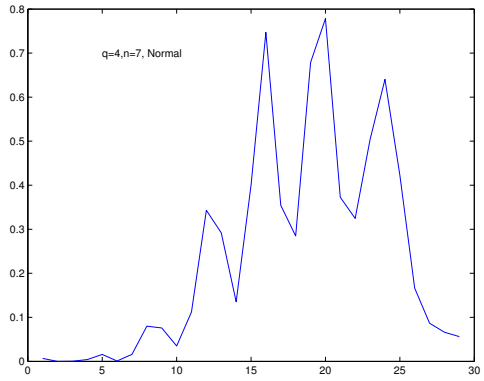

Figure 5.7: Example A3 : normal convergence

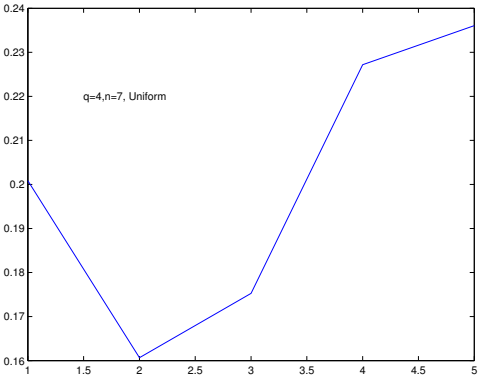

Figure 5.8: Example A3 : uniform convergence

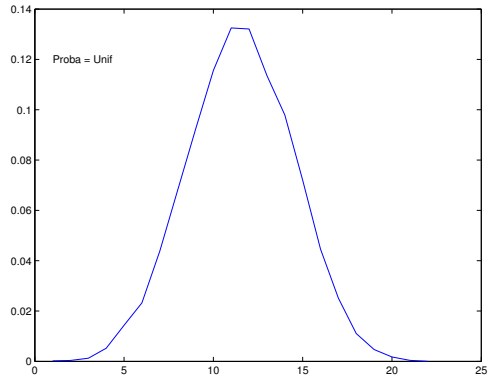

Figure 5.9: Example C1 : n=7, q=3

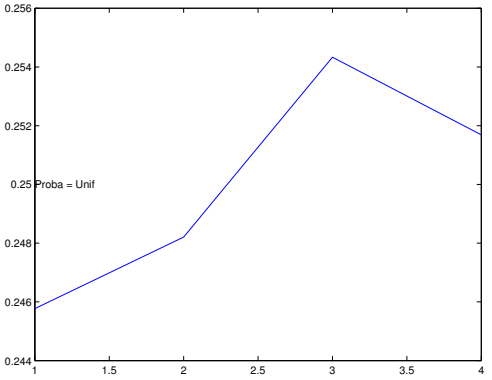

Figure 5.10: Example C1 : n=7, q=3

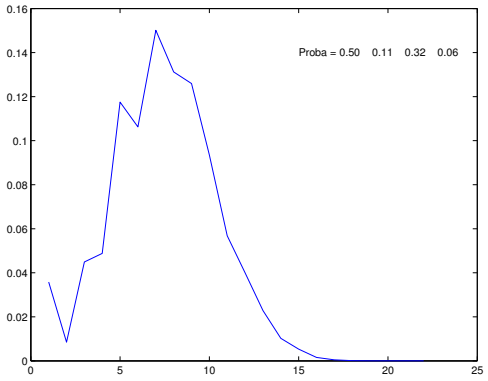

Figure 5.11: Example C2 : n=7, q=3

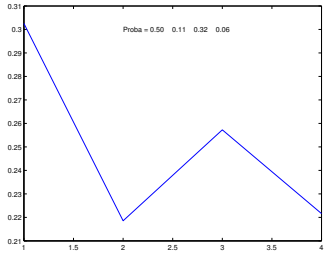

Figure 5.12: Example C2 : n=7, q=3

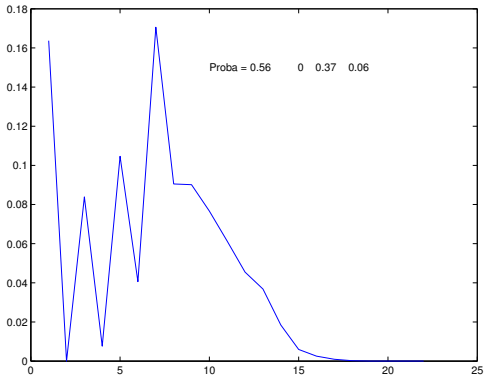

Figure 5.13: Example C3 : n=7, q=3

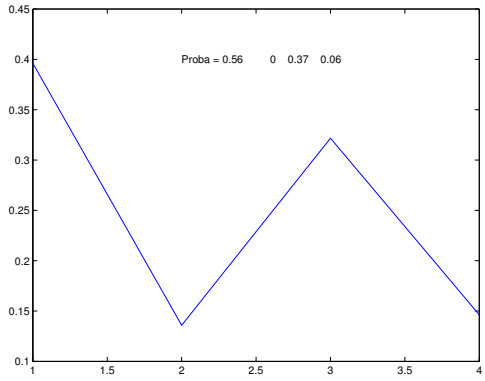

Figure 5.14: Example C3 : n=7, q=3

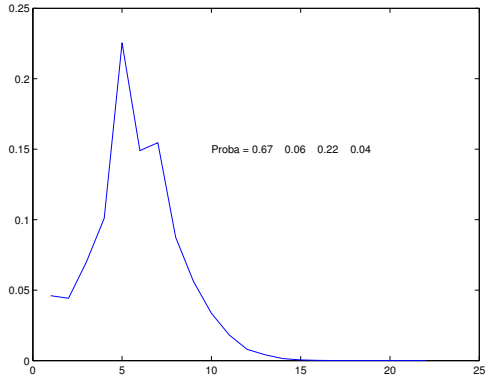

Figure 5.15: Example C4 : n=7, q=3

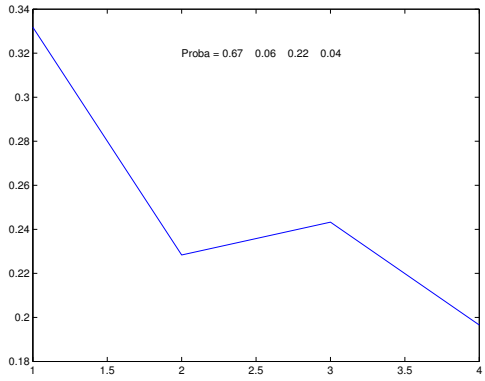

Figure 5.16: Example C4 : n=7, q=3

5.4.2 Case n=3

In this section, we calculate the exact probability that $X_1 + X_2 + X_3 = p$ when $X_s \in \{0, 1, ..., q\}$ for s=1,2,3. : cf figures 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23, 5.24, 5.25, 5.26.

It is naturally a case where we could doubt to have already a sufficient estimate of the limit distributions.

Nevertheless, we understand that if q is enough big, the curve of the probabilties of $X_1 + X_2 + X_3$ has a shape of bell : examples B2 B3, B4, B5. We also understand that it is enough that $P\{X_s = q\}$ have a distribution close to the uniform distribution to have this type of curve (example B2). But, it is not the case if the marginal distributions are very different from uniform distribution (example B1).

5.4.3 Calculation by estimate : variations of n

The previous results allow to verify the rate of the convergence to normal or to the uniform distributions.

This type of numerical calculations are possible only for n and q enough small. But, we can also verify the speed of this convergence by estimate, for example by using histograms: we give an example in figures 5.1 and 5.2. These results confirm our previous conclusions. As soon as $n \geq 8$, the densities have a curve which has a shape of bell for the most part of probabilities. It is even more true for the convergence to the uniform distribution.

5.4.4 Case where there is no convergence

The only case which we found where there is not convergence is the one where the probabilities are concentrated near a small number of points

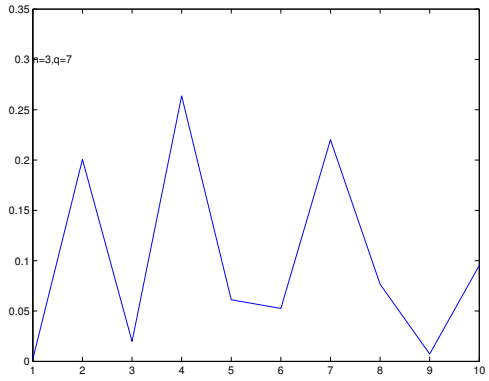

Figure 5.17: Example B1 : convergence to the normal distribution

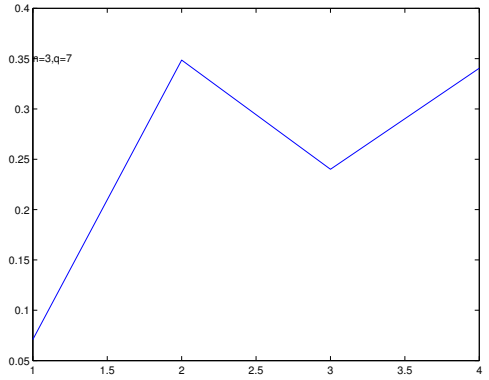

Figure 5.18: Example B1 : convergence to the uniform distribution

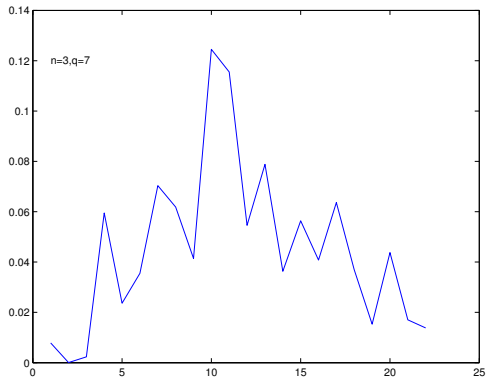

Figure 5.19: Example B2 : convergence to the normal distribution

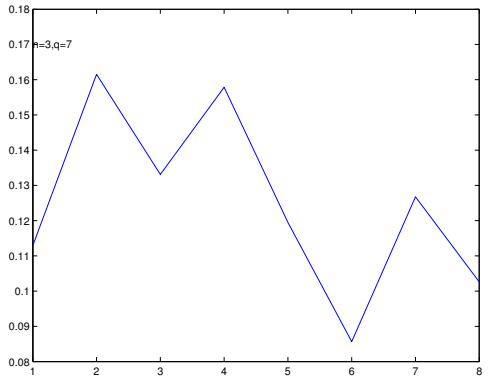

Figure 5.20: Example B2 : convergence to the uniform distribution

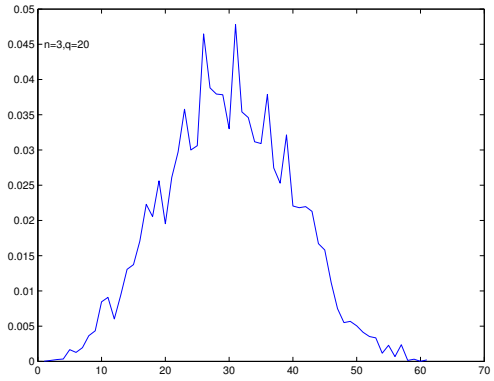

Figure 5.21: Example B3 : convergence to the normal distribution

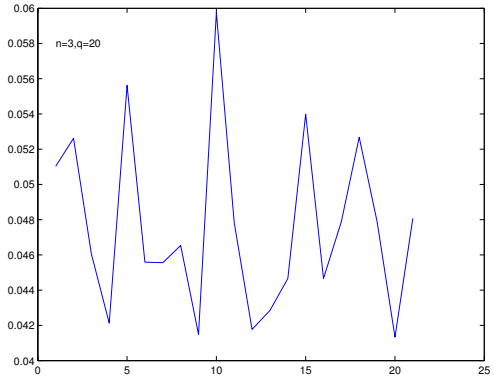

Figure 5.22: Example B3 : convergence to the uniform distribution

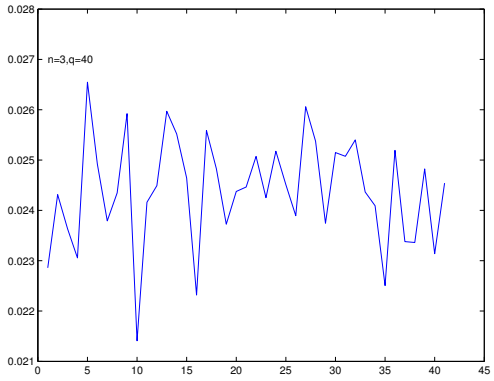

Figure 5.23: Example B4 : convergence to the normal distribution

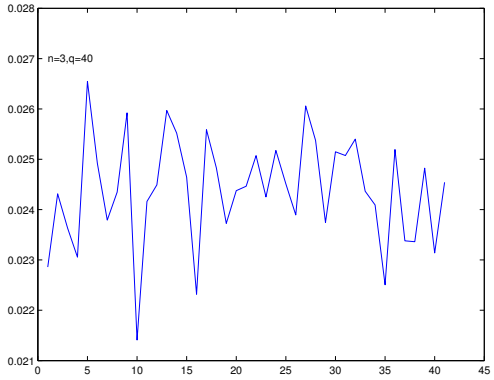

Figure 5.24: Example B4 :convergence to the uniform distribution

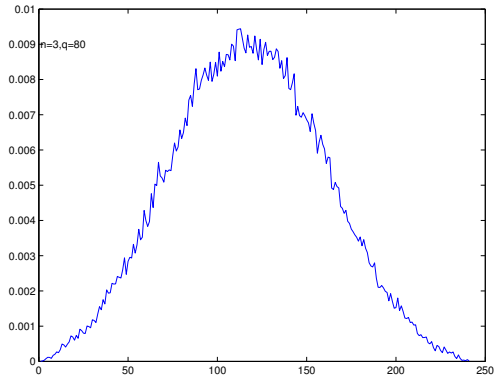

Figure 5.25: Example B5 : convergence to the normal distribution

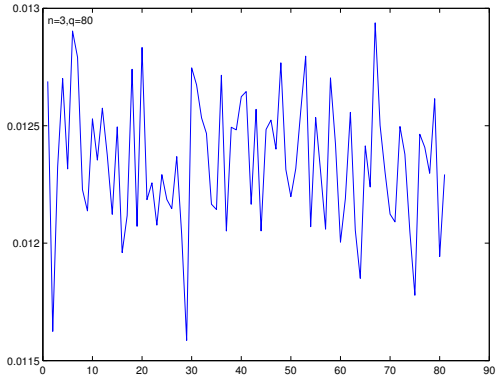

Figure 5.26: Example B5 : convergence to the uniform distribution

There is also very particular distributions as $X_n = Y_n - Y_{n-1}$ where Y_n is IID (cf [21]). But it is very special case where $\mathbb{E}\{(X_1 + ... + X_n)^2\}$ is bounded.

In that case, there is no convergence of $(X_1 + X_2 + ... + X_n)/\sqrt{n}$ to the normal distribution. On the other hand, often $\alpha(n)\overline{X_1 + X_2 + ... + X_n}$ converges nevertheless to the uniform distribution if $\alpha(n) \to \infty$.

In any case, it is easy to realize that the data which we choose in section 11.2 do not verify these hypotheses.

Because the only case where there is no convergence is the one where the probability are concentrated near a small number of points, a means of speed up this convergence is to standardize the marginal distributions. For it we use techniques defined in section 11.1.2 and in chapter 8.

5.4.5 Comparison beetween the XORLT and the CLT

We want to study numerically the rate of convergence of the XORLT when the CLT is satisfied. Then, we assume

$$Y = \frac{X_1 + ... + X_n}{\sigma\sqrt{n}} \sim N(0, 1) \ .$$

Then,

$$X_1 + ... + X_n = \sigma\sqrt{n}Y \sim N(0, n\sigma^2) \ .$$

Then, one can study the distribution of $\overline{X_1 + ... + X_n}$.

Here we study the distribution of $\overline{X_1 + ... + X_n}$ when n=10 with the following variances 1/15, 1/20, 1/50, 1/200 : cf figure 5.27, 5.28, 5.29, 5.30. We see that we are enough near of the uniform distribution if σ^2 is not too big. We remind that if $\sigma^2 = 1/12$ we are in the case of a uniform distribution : the density is thus identical to 1. If the variance is 1/200, it begins to have an important break of the uniform distribution. This one vanish enough fast if we increase n : cf figures 5.31 and 5.32.

5.4.6 Conclusion

All the previous results confirm the fast convergence of the curves of probability of $X_1 + X_2 + ... + X_n$ and $\overline{X_1 + X_2 + ... + X_n}$. The only case where there was no convergence enough fast is the one where the probability are concentrated near a small number of points.

For the data used in the construction of $b^1(n')$ in section 11.2, we can think that a sum of 10 terms is sufficient so that our hypotheses are satisfied. If we wanted to avoid every risk of error, it would be enough to choose S=15 or 20 (because of the rate of the convergence).

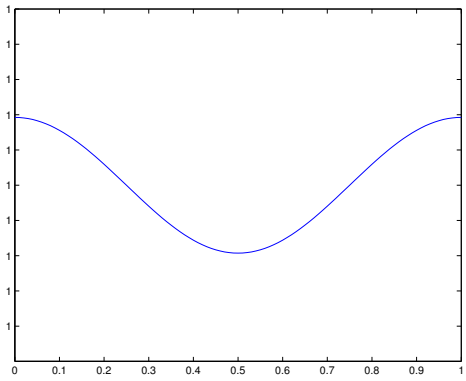

Figure 5.27: n=10, $\sigma^2 = 15$

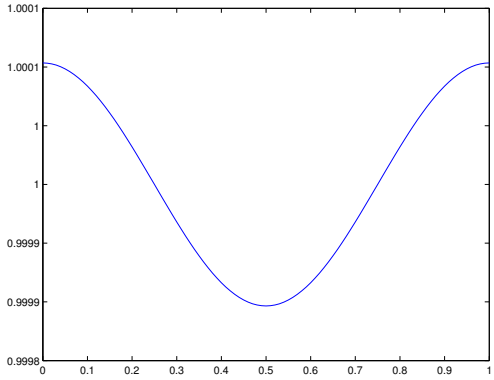

Figure 5.28: n=10, $\sigma^2 = 20$

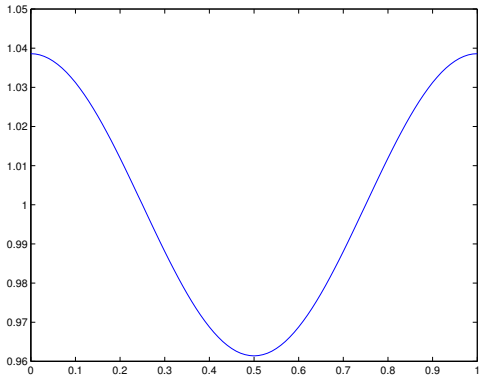

Figure 5.29: n=10, $\sigma^2 = 50$

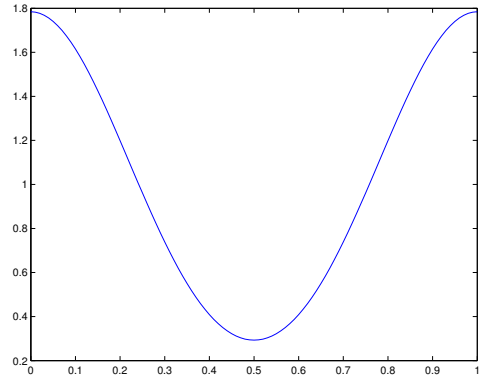

Figure 5.30: n=10, $\sigma^2 = 200$

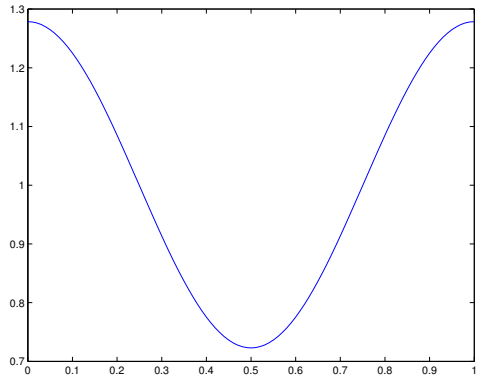

Figure 5.31: n=10, $\sigma^2 = 200$

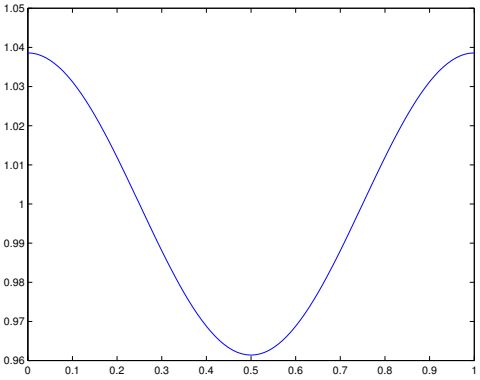

Figure 5.32: n=20, $\sigma^2 = 200$

5.5 Rate of convergence in the XORLT

In the CLT, it is known that the convergence to the normal distribution is very fast. It is the same in the XORLT for the convergence to the uniform distribution.

5.5.1 One-dimensional case

In this section, the following hypotheses are used.

Notations 5.5.1 *Let X_i, i=1,2,...,S, be a sequence of random variables with values in $\{0,1,...,N-1\}$. We set $p_{x_1,...,x_S} = P\{(X_1 = x_1) \cap ..... \cap (X_S = x_s)\}$.*

Hypothesis 5.5.1 *We suppose that $p_{x_1,...,x_S} = \frac{p'_{x_1,...,x_S}}{\sum_{x_1,...,x_S} p'_{x_1,...,x_S}}$.*

We suppose that the $p'_{x_1,...,x_S}$ are a sample of an IID sequence of random variables $P'_{x_1,...,x_S}$ defined on a probability space $(\Omega_4, A_4, Proba_4)$: $p'_{x_1,...,x_S} = P'_{x_1,...,x_S}(\omega_4)$. Let $E_{P'}$ and $\sigma^2_{P'}$ be, respectively, the associated expectation and the associated variance.

For example, suppose that the probabilities are choosen at random. In order to define mathematically this assumption, one can assume that $P'_{x_1,...,x_S}$ has the uniform distribution on $[0,1]$.

Let us notice that to consider the set of all the possible probabilities is a reasonable idea because the probabilities that we consider are those which correspond to a sample. There is thus a possible multitude of it.

Proposition 5.5.1 *Suppose that $\sqrt{N}$ is big and $\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^{S-1}}}$ is small. Then, with a probability in the order of $1 - 2\Gamma(b)$,*

$$\sum_{\overline{x_1+...+x_S}=y} p_{x_1,...,x_S} \approx 1/N + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{S+1}}} \ .$$

Proof : Let D be a subset of $\{0,1,...,N-1\}^S$. Let N(D) be the number of points of D. Asymptotically,

$$\frac{\sum_{(x_1,...,x_S)\in D}(P'_{x_1,...,x_S} - E_{P'})}{\sigma_{P'}\sqrt{N(D)}} \sim N(0,1) \ .$$

Let $b > 0$. Then, approximately,

$$Prob_4\left\{\frac{\left|\sum_{(x_1,\ldots,x_S)\in D}(P'_{x_1,\ldots,x_S} - E_{P'})\right|}{\sigma_{P'}\sqrt{N(D)}} \geq b\right\} = \Gamma(b) \ .$$

Then, with a probability in the order of $1 - \Gamma(b)$,

$$\frac{1}{N(D)}\Big[\sum_{(x_1,\ldots,x_S)\in D} p'_{x_1,\ldots,x_S} - N(D)E_{P'}\Big] = Ob(1)\frac{b.\sigma_{P'}\sqrt{N(D)}}{N(D)} \ .$$

Then, with a probability in the order of $1 - \Gamma(b)$,

$$\frac{1}{N^{S-1}}\sum_{\overline{x_1+\ldots+x_S}=y} p'_{x_1,\ldots,x_S} - E_{P'} = Ob(1)\frac{b.\sigma_{P'}}{\sqrt{N^{S-1}}} \ .$$

Therefore, with a probability in the order of $1 - \Gamma(b)$,

$$\frac{1}{N^S}\sum_{x_1,\ldots,x_S} p'_{x_1,\ldots,x_S} - E_{P'} = Ob(1)\frac{b.\sigma_{P'}}{\sqrt{N^S}} \ .$$

Then, if $\sqrt{N}$ is big, with a probability in the order of $1 - 2\Gamma(b)$,

$$\frac{(1/N^{S-1})\sum_{\overline{x_1+\ldots+x_S}=y} p'_{x_1,\ldots,x_S}}{(1/N^S)\sum_{x_1,\ldots,x_S} p'_{x_1,\ldots,x_S}} = \frac{E_{P'} + Ob(1)\frac{b.\sigma_{P'}}{\sqrt{N^{S-1}}}}{E_{P'} + Ob(1)\frac{b.\sigma_{P'}}{\sqrt{N^S}}}$$

$$= \frac{1 + Ob(1)\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^{S-1}}}}{1 + Ob(1)\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^S}}}$$

$$\approx 1 + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{S-1}}} + Ob(1)\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^S}}$$

$$\approx 1 + Ob(1)\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^{S-1}}} \ .$$

Therefore,

$$\frac{\sum_{\overline{x_1+\ldots+x_S}=y} p'_{x_1,\ldots,x_S}}{\sum_{x_1,\ldots,x_S} p'_{x_1,\ldots,x_S}}$$

$$= (1/N)\ \frac{(1/N^{S-1})\sum_{\overline{x_1+\ldots+x_S}=y} p'_{x_1,\ldots,x_S}}{(1/N^S)\sum_{x_1,\ldots,x_S} p'_{x_1,\ldots,x_S}}$$

$$\approx 1/N + Ob(1)\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^{S+1}}} \ .$$

Then, if $\sqrt{N}$ is big, with a probability in the order of $1 - 2\Gamma(b)$,

$$\sum_{\overline{x_1+...+x_S=y}} p_{x_1,...,x_S} \approx 1/N + Ob(1)\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^{S+1}}} \ . \blacksquare$$

It thus gives us an idea of the rate of convergence to the uniform distribution. For example, if S=11, N=1.000.000, b=100, if the $P'_{x_1,...,x_S}$'s have the uniform distribution on [0,1], with a probability bigger than $1 - 2e^{-10000/2}$,

$$\sum_{\overline{x_1+...+x_S=y}} p_{x_1,...,x_S} \approx 1/N + Ob(1)\frac{58}{\sqrt{N^{S+1}}} = \frac{1}{10^6}\left[1 + Ob(1)\frac{58}{10^{30}}\right] \ .$$

It is obviously very close to the uniform law. Thus we have a rate of convergence extremely fast .

5.5.2 Multidimensional case

In dimension p, we obtain results similar to the one-dimensional case for the case $(X_i^1, ..., X_i^p) \xrightarrow{D} (X^1, ..., X^p)$: the convergence to the p-dimensional uniform distribution holds. It means that the the X^t's are independent. It is thus an very useful result.

We generalize the notations of one-dimensional case by the following way.

Notations 5.5.2 *Let $(X_i^1, ..., X_i^p)$, i=1,2,...,S, be a sequence of random vectors with values in $\{0, 1, ..., N-1\}^p$. We set*

$$p_{x_1^1,..,x_S^1;..;x_1^p,..,x_S^p}$$
$$= P\{(X_1^1 = x_1^1) \cap ... \cap (X_S^1 = x_s^1) \cap .... \cap (X_1^p = x_1^p) \cap ... \cap (X_S^p = x_s^p)\}.$$

Hypothesis 5.5.2 *Suppose that*

$$p_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p} = \frac{p'_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p}}{\sum_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p} p'_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p}} \ .$$

We assume that the $p'_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p}$'s are a sample of a sequence of IID random variables $P'_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p}$ defined on a probability space $(\Omega_6, \mathcal{A}_6, Proba_6)$: $p'_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p} = P'_{x_1^1,...,x_S^1;...;x_1^p,...,x_S^p}(\omega_6)$. Let $E_{P'}$ and $\sigma^2_{P'}$ be the associated expectation and the associated variance, respectively .

Then the following proposition holds.

Proposition 5.5.2 *Suppose that* $\sqrt{N}$ *is big and* $\frac{b.\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}}$ *is small. Then, with a probability in the order of* $1-2\Gamma(b)$*,*

$$\sum_{\overline{x_1^1+....+x_S^1}=y_1;.....;\overline{x_1^p+....+x_S^p}=y_p} p_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} \approx 1/N^p\Big[1+\frac{Ob(1).b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}}\Big] \ .$$

Proof : Let D be a subset of $\{0,1,...,N-1\}^{Sp}$. Asymptoticaly,

$$\frac{\sum_{(x_1^1,...,x_S^1;....;x_1^p,...,x_S^p)\in D}(P'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}-E_{P'})}{\sigma_{P'}\sqrt{N(D)}} \sim N(0,1) \ .$$

Let $b>0$. Then,

$$Prob_6\left\{\frac{\big|\sum_{(x_1^1,...,x_S^1;....;x_1^p,...,x_S^p)\in D}(P'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}-E_{P'})\big|}{\sigma_{P'}\sqrt{N(D)}}\geq b\right\}=\Gamma(b) \ .$$

Then, with a probability in the order of $1-\Gamma(b)$,

$$\frac{1}{N(D)}\Big[\sum_{(x_1^1,...,x_S^1;....;x_1^p,...,x_S^p)\in D} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}-N(D)E_{P'}\Big]$$

$$=Ob(1)\frac{b.\sigma_{P'}\sqrt{N(D)}}{N(D)} \ .$$

Then, with a probability in the order of $1-\Gamma(b)$,

$$\frac{1}{N^{p(S-1)}}\sum_{\overline{x_1^1+....+x_S^1}=y_1;.....;\overline{x_1^p+....+x_S^p}=y_p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}-E_{P'}$$

$$=Ob(1)\frac{b\sigma_{P'}}{\sqrt{N^{p(S-1)}}} \ .$$

Therefore, with a probability in the order of $1-\Gamma(b)$,

$$\frac{1}{N^{pS}}\sum_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}-E_{P'}=Ob(1)\frac{b\sigma_{P'}}{\sqrt{N^{pS}}} \ .$$

Then, if $\sqrt{N}$ is big, with a probability in the order of $1-2\Gamma(b)$,

$$\frac{\frac{1}{N^{p(S-1)}}\sum_{\overline{x_1^1+....+x_S^1}=y_1;.....;\overline{x_1^p+....+x_S^p}=y_p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}}{\frac{1}{N^{pS}}\sum_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}}$$

$$= \frac{E_{P'} + Ob(1)\frac{b\sigma_{P'}}{\sqrt{N^{p(S-1)}}}}{E_{P'} + Ob(1)\frac{b\sigma_{P'}}{\sqrt{N^{pS}}}}$$

$$= \frac{1 + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}}}{1 + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{pS}}}}$$

$$\approx 1 + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}} + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{pS}}}$$

$$\approx 1 + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}} \ .$$

Therefore,

$$\frac{\sum_{\overline{x_1^1+....+x_S^1}=y_1;.....;\overline{x_1^p+....+x_S^p}=y_p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}}{\sum_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}}$$

$$= (1/N^p)\frac{\frac{1}{N^{p(S-1)}}\sum_{\overline{x_1^1+....+x_S^1}=y_1;.....;\overline{x_1^p+....+x_S^p}=y_p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}}{\frac{1}{N^{pS}}\sum_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} p'_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p}}$$

$$\approx 1/N^p + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S+1)}}} \ .$$

Then, if $\sqrt{N}$ is big, with a probability in the order of $1-2\Gamma(b)$,

$$\sum_{\overline{x_1^1+....+x_S^1}=y_1,.....,\overline{x_1^p+....+x_S^p}=y_p} p_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} \approx 1/N^p + Ob(1)\frac{b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S+1)}}}$$

$$= 1/N^p \Big[1 + \frac{Ob(1).b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}}\Big] \ . \ \blacksquare$$

The approximation of the uniform distribution remains valid even when p is big. So, in the maximum case, if p=N, with a probability of $1 - 2\Gamma(b)$,

$$\sum_{\overline{x_1^1+....+x_S^1}=y_1;.....;\overline{x_1^p+....+x_S^p}=y_p} p_{x_1^1,...,x_S^1;....;x_1^p,...,x_S^p} \approx \frac{1}{N^N}\Big[1 + \frac{Ob(1).b\sigma_{P'}}{E_{P'}\sqrt{N^{N(S-1)}}}\Big] \ .$$

It thus gives us an idea of the rate of convergence to the uniform distribution: we have a speed of convergence extremely fast.

Consequence Now, there is N^p possible $(y_1, ..., y_p)$. Therefore, with a probability bigger than $1 - 2N^p\Gamma(b)$, for all $(y_1, ..., y_p)$,

$$P\Big\{\{\overline{X_1^1 + X_2^1 + ... + X_S^1} = y_1\} \cap .... \cap \{\overline{X_1^p + X_2^p + ... + X_S^p} = y_p\}\Big\}$$

$$= (1/N^p)\Big[1 + Ob(1)\frac{b\sigma_{P'}}{E_{P'}N^{p(S-1)/2}}\Big] \ .$$

Then, we have a speed of convergence extremely fast. Nevertheless,

$$P\Big\{\{\overline{X_1^1 + X_2^1 + ... + X_S^1} = y_1\} \cap .... \cap \{\overline{X_1^p + X_2^p + ... + X_S^p} = y_p\}\Big\}$$

can be very different of $1/N^p$.

This case is always possible in the set of the probabilities. But it occurs with a very weak probability: e.g. $\Gamma(b) \leq 1/10^{197}$ if b=30. It occurs for example when the X_t 's are concentrated near a small number of points.

It occurs also when probabilities are not chosen randomly : it is an important case. Indeed it includes the continuous case : cf section 5.5.8.

5.5.3 Case of independence

Now, we understand that if we suppose that variables are independent, there is a probability stronger than $P\{\overline{X_1 + ... + X_S} = 1\} \neq 1/N$. This result could seem surprising. We are going to try to understand what it happens.

In this section we suppose that the following hypotheses hold.

Notations 5.5.3 *Let X_i, i=1,2,..,S, be S independent random variables with values in $\{0, 1, ..., N-1\}$. We set $P\{X_i = x_n\} = p^i_{x_n} = (1/N)[1 + v^i_{x_n}]$.*

Hypothesis 5.5.3 *We keep the same notations as in section 5.5.1. But we assume that* $p_{x_1,...,x_S} = p^1_{x_1}.....p^S_{x_S}$.

For all $i \in \{1, 2, ..., S\}$, *we assume that* $p^i_{x_n}$ *is a realization of the sequence of random variables* $P^i_{x_n}$ *defined on a probability space* $(\Omega_3, \mathcal{A}_3, Proba_3)$*:* $p^i_{x_n} = P^i_{x_n}(\omega_3)$ *and* $V^i_{x_n}$ *:* $v^i_{x_n} = V^i_{x_n}(\omega_3)$.

One can assume that the $V^i_{x_n}$'s have a mean equal to zero. Indeed the following lemma holds

Lemma 5.5.1 *Under the previous assumptions,* $\sum_{x_n=0}^{N-1} v^i_{x_n} = 0$.

Proof We have $1 = \sum_{x_1} P\{X_1 = x_1\} = \sum_{x_1}(1/N)(1 + v^1_{x_1}) = 1 + (1/N)\sum_{x_1} v^1_{x_1}$. Therefore, $\sum_{x_1} v^1_{x_1} = 0$. ■

Then, we study $p_y = P\{\overline{X_1 + X_2 + ... + X_S} = y\}$. We shall understand in section 5.5.4 that their behavior is determined by the sum

$$\sum_{\overline{x_1+....+x_S}=y} v^1_{x_1} v^2_{x_2}....v^S_{x_S}.$$

We know that, if the $v^1_{x_1} v^2_{x_2}....v^S_{x_S}$'s are all different, $\sum_{\overline{x_1+....+x_S}=y} v^1_{x_1} v^2_{x_2}....v^S_{x_S}$ can be considered as the sum of an IID sample of size $N^{(S-1)}$ (cf section 8.1.2).

Let σ^2_V be the variance of $V^1_{x_{i_1}}....V^S_{x_{i_S}}$. If one can apply the CLT, one has approximately :

$$Proba_3\Big\{\frac{\sum_{\overline{x_1+....+x_S}=y} V^1_{x_1}....V^S_{x_S}}{\sigma_V N^{(S-1)/2}} \geq b\Big\} \leq \Gamma(b) \ .$$

Because

$$P\{\overline{X_1 + X_2 + ... + X_S} = y\} = \sum_{\overline{x_1+....+x_S}=y} (1 + v^1_{x_1})....(1 + v^S_{x_S})/N^S \ ,$$

by section, 5.5.4, we have again the inequality

$$Proba_3\Big\{\frac{|P\{\overline{X_1 + X_2 + ... + X_S} = y\} - 1/N|}{\sigma_V} \geq \frac{b}{N^{(S+1)/2}}\Big\} \leq \Gamma(b) \ .$$

The point is to know if we can apply the CLT when we make add on samples of size N^{S-1} while we have NS variable $P^i_{x_{s_i}}$: It would be necessary that the number of variables is bigger than the size of samples. We shall study this problem in section 5.5.8.

5.5.4 Sum of $v^1_{x_1} v^2_{x_2} .... v^S_{x_S}$

Now we understand that the p_y's depend on sums of $v^1_{x_1} v^2_{x_2} .... v^S_{x_S}$.

Proposition 5.5.3 *With the previous notations,*
$P\{\overline{X_1 + ... + X_S} = y\} = (1/N) + (1/N^S) \sum_{\overline{x_1+...+x_S}=y} v^1_{x_1} v^2_{x_2} .... v^S_{x_S}$.

Proof First, $P\{\overline{X_1 + X_2 + ... + X_S} = y\}$
$= \sum_{\overline{x_1+....+x_S}=y} P\big\{\{X_1 = x_1\} \cap ..... \cap \{X_S = x_S\}\big\}$
$= \sum_{\overline{x_1+....+x_S}=y} P\{X_1 = x_1\}....P\{X_S = x_S\}$
$= (1/N^S) \sum_{\overline{x_1+...+x_S}=y} (1 + v^1_{x_1}).....(1 + v^S_{x_S})$.

Now,
$(1 + v^1_{x_1}).....(1 + v^S_{x_S})$

$= 1 + [v^1_{x_1} + ... + v^S_{x_S}] + \sum_{i_1<i_2} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} + \sum_{i_1<i_2<i_3} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}}$
$+$.....
$+ \sum_{i_1<i_2<....<i_q} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} .... v^{i_q}_{x_{i_q}}$
$+$.....
$+ v^1_{x_1} v^2_{x_2} .... v^S_{x_S}$.

Clearly, by the following lemma (5.5.2),
$P\{\overline{X_1 + ... + X_S} = y\} = (1/N) + (1/N^S) \sum_{\overline{x_1+...+x_S}=y} v^1_{x_1} v^2_{x_2} .... v^S_{x_S}$. ■

Lemma 5.5.2 *Suppose $q < S$. Then,*
$\sum_{\overline{x_1+....+x_S}=y} \big[\sum_{i_1<i_2<....<i_q} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}} .... v^{i_q}_{x_{i_q}}\big] = 0.$

Proof We have
$\sum_{\overline{x_1+....+x_S}=y} \big[\sum_{i_1<i_2<....<i_q} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}} .... v^{i_q}_{x_{i_q}}\big]$
$= \sum_{i_1<i_2<....<i_q} \big[\sum_{\overline{x_1+....+x_S}=y} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}} .... v^{i_q}_{x_{i_q}}\big]$.

For example, if $i_q < S$,
$\sum_{\overline{x_1+....+x_S}=y} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}} .... v^{i_q}_{x_{i_q}}$
$= \sum_{x_{i_1}} \sum_{x_{i_2}} .... \sum_{x_{i_{S-1}}} \sum_{x_S = \overline{y - x_1 + ... - x_{S-1}}} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}} .... v^{i_q}_{x_{i_q}}$
$= \sum_{x_{i_1}} \sum_{x_{i_2}} .... \sum_{x_{i_{S-1}}} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}} v^{i_3}_{x_{i_3}} .... v^{i_q}_{x_{i_q}}$
$= \Big[\sum_{x_{i_1}} v^{i_1}_{x_{i_1}}\Big].....\Big[\sum_{x_{i_q}} v^{i_q}_{x_{i_q}}\Big]$
$= 0$ because $\sum_{x_{i_1}} v^{i_1}_{x_{i_1}} = 0$. ■

5.5.5 Study of $\sum_{\overline{x_1+....+x_S}=y} v^1_{x_1} v^2_{x_2}....v^S_{x_S}$

Then, $\sum_{\overline{x_1+....+x_S}=y} \left[\sum_{i_1<i_2<....<i_S} v^{i_1}_{x_{i_1}} v^{i_2}_{x_{i_2}}....v^{i_S}_{x_{i_S}} \right]$ can be regarded as the sum of an IID sample of size $N^{(S-1)}$ of mean zero and of variance σ_V^2.

It means that, with a probability near 1, if the CLT holds, and if b is large enough,

$$P\{\overline{X_1 + X_2 + ... + X_S} = y\} = (1/N)\Big[1 + Ob(1)\frac{b\sigma_V}{N^{(S-1)/2}}\Big] .$$

But we understand now on examples that the CLT is not satisfied as well as in the general case. We shall study the case where $X_s \in \{0, 2, 4, ..., \}$.

5.5.6 Counterexamples

We obtain results similar at those of section 5.4.1 : if the probability is concentrated near a small number of points, $P\{\overline{X_1 + X_2 + ... + X_S} = y\} \neq 1/N$ is a possible result. The extreme case is the one where the probabilities $p^i_{x_s}$ are concentrated in a single point, for example $x_s = 1$. In this point, $p^i_1 = 1$ and $v^i_1 = N - 1$.

Now suppose that N is even. Suppose that the X_t's have the uniform distribution over the even numbers : $p^i_{2j} = 2/N$ and $p^i_{2j+1} = 0$. In this case, $v^i_j = (-1)^j = \pm 1$. Therefore, $v^1_{x_1} v^2_{x_2}....v^S_{x_S} = (-1)^{x_1+..+x_S}$.

For example, if y is even, $v^1_{x_1} v^2_{x_2}....v^S_{x_S} = 1$. Therefore,

$$\sum_{\overline{x_1+....+x_S}=y} v^1_{x_1} v^2_{x_2}....v^S_{x_S} = \sum_{\overline{x_1+....+x_S}=y} 1 = N^{S-1}.$$

Actually,

$$\sum_{\overline{x_1+....+x_S}=y} v^1_{x_1} v^2_{x_2}....v^S_{x_S} = (-1)^y N^{S-1}$$

and

$$P\{\overline{X_1 + X_2 + ... + X_S} = y\} = 0 \ or \ = 2/N.$$

Then $p_y = P\{\overline{X_1 + X_2 + ... + X_S} = y\}$ can be very different of 1/N. This case is always possible in the set of all probabilities. It is the case where

$$\frac{\sum_{\overline{x_1+....+x_S}=y}(p_y - 1/N)}{\sigma_V/N^{(S+1)/2}} \geq b .$$

But it occurs with a weak probability. Now, we need to know more this probability if we cannot use directly the CLT.

5.5.7 Probability concentrated near even numbers

We thus want to know which is the probability that $P\{\overline{X_1 + X_2 + ... + X_S} = y\}$ is different enough of 1/N in the case of independence.

For that purpose, in this section, we are in the particular case where $p_{x_s} = \frac{p_{x'_s}}{\sum_{x_s} p'_{x_s}}$, where $p'_{x_s} = P'_{x_s}(\omega_3)$ and where the P'_{x_s} 's are an IID sequence of random variables which have the uniform distribution over [0,1] and which are defined over a probability space $(\Omega_3, \mathcal{A}_3, Proba_3)$.

As a matter of fact, we regard a case close to the case where the probabilities are uniformly distributed on the even numbers, that is the case where $p^i_{2j} \approx 2/N$ and $p^i_{2j+1} \approx 0$.

Then, we have $p_{x_s} = p^1_{x_s} = (1/N)[1 + v_{x_s}]$. Therefore, $Np_{x_s} = 1 + v_{x_s}$. Then, $v_{x_s} = Np_{x_s} - 1$.

For this study the following lemma is needed.

Lemma 5.5.3 *Let $Z_{No} = (1/\sqrt{N}) \sum_{x_s} [P'_{x_s} - 1/2]$. Let I be an interval. Then the following equivalence holds*

$$v_{x_s} \in I \iff p'_{x_s} \in \left[1 + 2z_{No}/\sqrt{N}\right](I+1)/2 \ .$$

Proof At first, $z_{No}\sqrt{N} = \sum_{x_s} [p'_{x_s} - 1/2]$. Therefore, $\sum_{x_s} p'_{x_s} = N/2 + \sqrt{N} z_{No}$. Therefore,

$$v_{x_s} = Np_{x_s} - 1 = \frac{Np'_{x_s}}{\sum_{x_s} p'_{x_s}} - 1 = \frac{Np'_{x_s}}{N/2 + \sqrt{N} z_{No}} - 1 \ .$$

Then the following equivalence holds : $v_{x_s} \in I \iff \frac{Np'_{x_s}}{N/2+\sqrt{N}z_{No}} - 1 \in I \iff \frac{p'_{x_s}}{1/2+z_{No}/\sqrt{N}} - 1 \in I \iff \frac{2p'_{x_s}}{1+2z_{No}/\sqrt{N}} - 1 \in I \iff \frac{2p'_{x_s}}{1+2z_{No}/\sqrt{N}} \in I+1 \iff p'_{x_s} \in \left[1 + 2z_{No}/\sqrt{N}\right](I+1)/2$. ■

Remark that Z_{No} has almost the normal distribution with mean zero and variance 1/12.

The probability that $V_{x_s} \in [1-f, 1[$ where f is small - that is the probability that $P_{x_s} \approx 2/N$ - is equivalent to the probability that $P'_{x_s} \in (1+2Z_{No}/\sqrt{N})[1-f/2, 1[$. Then, it is almost equivalent to the probability that $P'_{x_s} \in [1-f/2, 1[$. Then, it is almost equal to (f/2).

The probability that $V_{x_s} \in [-1, -1+f[$ - that is the probability that $P_{x_s} \approx 0$ - is equivalent to the probability that $P'_{x_s} \in (1+2Z_{No}/\sqrt{N})[0, f/2[$. Then, it is almost equivalent to the probability that $P'_{x_s} \in [0, f/2[$. Then, it is almost equal to (f/2).

Then, the following property holds.

Property 5.5.4 *The probability that $p^i_{x_s} \approx 0$ for each odd numbers and that $p^i_{x_s} \approx 2/N$ for each even number is, for each row i, $(f/2)^N$, and for all the rows $(f/2)^{NS}$.*

Comparison with the general case We compare this result with the general result of section 5.5.1 :

$$Proba_4\Big\{\frac{|P\{\overline{X_1+X_2+....+X_S}=y\}-1/N|}{\sigma_{P'}} \geq \frac{b}{E_{P'}N^{(S+1)/2}}\Big\} \leq \Gamma(b) \ .$$

The following lemma is needed.

Lemma 5.5.5 *The following equivalences hold : $v_{x_s} \in [1-f, 1[\iff p_{x_s} \in [2/N - f/N, 2/N[$ and $v_{x_s} \in [-1, -1+f[\iff p_{x_s} \in [0, f/N[$.*

Proof We have $v_{x_s} \in [1-f, 1[\iff Np_{x_s} = v_{x_s}+1 \in [2-f, 2[\iff p_{x_s} = (1/N)[v_{x_s}+1] \in [2/N - f/N, 2/N[$. Moreover $v_{x_s} \in [-1, -1+f[\iff Np_{x_s} = v_{x_s}+1 \in [0, f[\iff p_{x_s} = (1/N)[v_{x_s}+1] \in [0, f/N[$. ■

Now, we use the following property.

Property 5.5.6 *Suppose that $p_{x_s} \in [0 \ , \ f/N[$ if N is odd and that $p_{x_s} \in [2/N - f/N, 2/N]$ if N is even.*

Suppose that y is odd. Let $\sigma^2_{P'} = 1/12$. Then,

$$\frac{|P\{\overline{X_1+X_2+....+X_S}=1\}-1/N|}{\sigma_{P'}} \geq \frac{(1/N)[1-2^{S-1}f]}{\sqrt{1/12}} \ .$$

Proof For example assume y=1. Then,

$$P\{\overline{X_1+X_2+....+X_S}=1\} = \sum_{\overline{x_1+....+x_S}=1} p^1_{x_1}....p^S_{x_S}.$$

In all these x_s, there is at least one x_s odd : $x_s = 2j_0 + 1$. Therefore, $p^i_{x_s} = p^i_{2j_0+1} \leq f/N$. Therefore,

$$p^1_{x_1}....p^S_{x_S} \leq (f/N)(2/N)^{S-1}.$$

Therefore,

$$\sum_{\overline{x_1+....+x_S}=1} p^1_{x_1}....p^S_{x_S} \leq N^{S-1}[(f/N)(2/N)^{S-1}] = 2^{S-1}f/N.$$

Therefore,

$$1/N - P\{\overline{X_1 + X_2 + ... + X_S} = 1\} \geq 1/N - 2^{S-1}f/N = 1/N[1 - 2^{S-1}f].$$

Then, with $(\sigma_{P'})^2 = 1/12$,

$$\frac{|P\{\overline{X_1 + X_2 + ... + X_S} = 1\} - 1/N|}{\sigma_{P'}} \geq \frac{(1/N)[1 - 2^{S-1}f]}{\sqrt{1/12}} \ . \ \blacksquare$$

Under the hypotheses of section 5.5.1 , with $(\sigma_{P'})^2 = 1/12$ and $E_{P'} = 1/2$, we know that

$$Proba_3\Big\{\frac{|P\{\overline{X_1 + ... + X_S} = 1\} - 1/N|}{\sigma_{P'}} \geq \frac{b}{E_{p'}N^{(S+1)/2}}\Big\} \leq \Gamma(b) \ .$$

Now in order that

$$\frac{|P\{\overline{X_1 + X_2 + ... + X_S} = 1\} - 1/N|}{\sigma_{P'}} \geq \frac{b}{E_{p'}N^{(S+1)/2}} \ ,$$

it is sufficient for example, that $\frac{2b}{N^{(S+1)/2}} = \frac{(1/N)[1-2^{S-1}f]}{\sqrt{1/12}}$, that is $b = \sqrt{3}[1 - 2^{S-1}f]N^{(S-1)/2}$.

Now we know [2] that, if b is big, $\Gamma(b) = \frac{\sqrt{2}e^{-b^2/2}}{\sqrt{\pi}b}$. Therefore,

$$\Gamma(b) \leq e^{-b^2/2}/b = \frac{e^{-1.5[1-2^{S-1}f]^2N^{(S-1)}}}{\sqrt{3}|1 - 2^{S-1}f|N^{(S-1)/2}} \leq e^{-1.5N^{S-1}}$$

[2] e.g. with the notations of Matlab, $\Gamma(b) = erfc(b/\sqrt{2})$, $erfcx(b/\sqrt{2}) = e^{b^2/2}erfc(b/\sqrt{2})$. If b is big $erfcx(b/\sqrt{2}) = \frac{\sqrt{2}}{\sqrt{\pi}b}$.

if f is small enough .

Now, in the case of independence, we know that it occurs with a probability in the order of $(f/2)^{NS}$: cf property 5.5.4 . Of course, it is not in the same order as $e^{-14N^{(S-1)}/10}$.

5.5.8 Consequences

Differences between the case of the independence and the general case The previous results mean that $\left\{ \frac{|P\{\overline{X_1+X_2+....+X_S}=1\}-1/N|}{\sigma_{P'}} \geq \frac{b}{E_{P'}N^{(S+1)/2}} \right\}$ has a probability bigger to come true in the set of the independent probabilities than in the set of all the probabilities.

This result could seem strange. Indeed, for the CLT, the convergence is admitted faster in the case of the independence. Now the XORLT is a transformation of the CLT.

However, in a way, it is normal. By imposing the hypothesis of independence, we lose parameters, thus probabilities : a priori there is less $p^1_{x_{s_1}}....p^S_{x_{s_S}}$ than possible $p_{x_1,....,X_S}$.

We have just understood that the probability that

$$\left\{ \frac{|P\{\overline{X_1+X_2+....+X_S}=1\}-1/N|}{\sigma_{P'}} \geq \frac{b}{E_{P'}N^{(S+1)/2}} \right\}$$

is much stronger in the case of the independence.

Why this difference? This difference thus results because there is less parameters. In the general case, there is N^S $(x_1, x_2, ..., x_S)$ S-uple which have to be considered. There is thus N^S possible $p'_{x_1,....,x_S}$. We make sums of N^{S-1} terms $p_{x_1,....,x_S}$. This is thus plausible.

In the case of independence, there is N $p^i_{x_s}$ for each row i. Then, all in all, there is NS $p^i_{x_s}$, i=1,..,S, $x_s \in F^*(N)$. But, we make always sums of N^{S-1} terms $p^1_{x_{s_1}}p^2_{x_{s_2}}....p^S_{x_{s_S}}$ which are determined by NS parameters. It is one of reasons of the problem.

Real behavior of probabilities In the case of the independence, the probability $Proba_3$ behaves as if we made a sum of NS terms. We should thus find a result of the type

$$Proba_3\left\{ \frac{|P\{\overline{X_1+X_2+....+X_S}=y\}-1/N|}{\sigma_{P'}} \geq \frac{b}{E_{P'}\sqrt{NS}} \right\} \leq \Gamma'(b) \ ,$$

where $\Gamma'(x) = O\big(\Gamma(x)\big)$.

But, in fact, we find $\frac{|P\{\overline{X_1+X_2+....+X_S}=1\}-1/N|}{\sigma_{P'}} \geq \frac{b}{E_{P'}N^{(S+1)/2}}$ only in case where the probability of X_s is concentrated near a small number of points. And we know that we have to eliminate this case : cf section 5.4.1.

In other cases, all our researches show that we find the same probability as in the general case

$$Proba_3\Big\{\frac{|P\{\overline{X_1+X_2+....+X_S}=y\}-1/N|}{\sigma_{P'}} \geq \frac{b}{E_{P'}N^{(S+1)/2}}\Big\} \leq \Gamma"(b) \ ,$$

where $\Gamma"(b) = O\big(\Gamma(b)\big)$.

We thus find the general case where the convergence is extremely fast with a probability infinitely close to 1. To convince itself, we can refer for example to the elementary study made in section 5.3.4.

We understand in section 8.2.2 how eliminating the case where the X_t 's are concentrated near a small number of points.

One of the reason of the problem

In fact the problem comes as well from the fact as we suppose that the x_n's have to be different some of the others in order that a sequence of real x_n behaves as an IID sequence, : cf section 8.1.2.

Of course, it is not the case when the probability is concentrated uniformly on the even numbers: there are only 2 possible values for $p_{x_1,....,x_S}$: 0 or $2^S/N^s$.

Now if the probability is concentrated on the even numbers in a random way, we meet ourselves in the same study but we replace N by N/2. It means that the probability that $\frac{|P\{\overline{X_1+X_2+....+X_S}=y\}-1/N|}{\sigma_{P'}} \geq b'$ can occur but with a less weak probability.

Now, if the x_n's - and thus the p_{x_t}'s - are not concentrated near a small number of points, the number of the possible values for $p_{x_1,....,x_S}$ is much bigger and well distributed. We shall thus have a fast convergence with a probability infinitely close to 1.

Anyway, we understand well the difference: the sum $\sum p^1_{x_{s_1}}....p^S_{x_{s_S}}$ can be considered as the sum of a sample IID as soon as the p_{x_t}'s are different enough. The probability that such a sample satisfies $\frac{|P\{\overline{X_1+X_2+....+X_S}=y\}-1/N|}{\sigma_{P'}} \geq b$ becomes again very weak. This probability increases as soon as p_{x_s} is concentrated near a small number of points.

Thus if the p_{x_t}'s are different, the sums $\sum p_{x^1_{s_1}}....p_{x^S_{s_S}}$ behave as sums of numbers chosen at random.

It stops if the p_{x_t}'s are concentrated near a small number of points (as in the case of uniform concentration near the even numbers). In that case, it is normal that $\frac{P\{\overline{X_1+X_2+....+X_S}=y\}-1/N)}{\sigma_{P'}} \geq b'$ is bigger.

Thus if the p_{x_t}'s are different enough, the sums $\sum p_{x^1_{s_1}}....p_{x^S_{s_S}}$ behave as sums of numbers chosen at random. We find the probability of the general case.

Let us notice that all these results are coherent.

We understand that the problem comes from the fact that the x_t's are concentrated near a small number of points. Otherwise, we have often a fast convergence with a probability infinitely close to 1.

It is an important result : it means that, under our hypotheses, there is a break of the XORLT if the probabilities are concentrated near a small number of points.

We thus know which problem, it is necessary to avoid to have a fast convergence. We shall use differents way for it : cf section 8.2.2. It will allow us to admit the hypotheses that we chose for our various type of construction : cf chapter 11 and 12

Problem in some cases

Proposition 5.5.1 is only a mathematical theorem with a measure on the set of the probabilities chosen a priori. This measure is not thus inevitably adapted to certain assumptions.

Continuous case It is not difficult to understand that proposition 5.5.1 gives absurd results in the case of continuous density.

Indeed, let us consider random variables with values in $F(m)$. We denote by f_Y the densisty of the sum with respect to μ_m :

$$f_Y(y) = m \sum_{\overline{x_1+...+x_S}=y} p_{x_1,...,x_S} \ .$$

Then, by proposition 5.5.1,

$$f_Y(y) = m \sum_{\overline{x_1+...+x_S}=y} p_{x_1,...,x_S} = 1 + O\Big(\frac{1}{m^{(S-1)/2}}\Big) \ .$$

Now, for probabilities estimated starting from sample of small size, one often admits that the X_s have a continuous density: it is an assumption a priori realistic and usual. Now, if X_t has values in [0,1] with a continuous density, according to proposition 5.5.1, with a probability equal to 1,

$$f_Y(y) = 1 \ .$$

It is understood that this fact is false. Thus let us choose S=3, and suppose that, for s=1,2, the densities of X_s with respect to μ_m is $f_{X_s}(x) = \eta_m 6x^5$ where

$\eta_m \to 1$ as $m \to \infty$. Suppose that the densities of X_3 is $f_{X_3}(x) = \eta_m 6(1-x)^5$ Then,

$$\sum_{\overline{x_1+x_2+x_3}=y} p_{x_1}p_{x_2}p_{x_3} = \sum_{(x_1,x_2)} p_{x_1}p_{x_2}p_{\overline{y-x_1-x_2}}$$

$$= \frac{1}{m^3} \sum_{(x_1,x_2)} f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(\overline{y-x_1-x_2})$$

$$= \frac{\eta_m^3 6^3}{m^3} \sum_{(x_1,x_2)} (x_1)^5 (x_2)^5 \big(1 - \overline{y-x_1-x_2}\big)^5 .$$

Therefore, if y=1 and if x_1 and x_2 are close to 1,

$$\sum_{\overline{x_1+x_2+x_3}=y} p_{x_1}p_{x_2}p_{x_3} = \frac{\eta_m^3 6^3}{m^3} \sum_{(x_1,x_2)} (x_1)^5 (x_2)^5 \big(x_1+x_2-1\big)^5 .$$

It is not difficult to understand that

$$m \sum_{\overline{x_1+x_2+x_3}=1} p_{x_1}p_{x_2}p_{x_3} \neq m \sum_{\overline{x_1+x_2+x_3}=7/8} p_{x_1}p_{x_2}p_{x_3}.$$

Therefore, the distribution of $\overline{X_1+X_2+X_3}$ is different from the uniform distribution.

However, this case is that where the marginal probabilities are concentrated each one nearly one only point: $(1,1,0)$. One thus finds the case which should a priori be excluded. But is really necessary it to exclude this case?

There is indeed concentration around a point. However this concentration is not so strong. In spite of that, one will not obtain $\sum_{\overline{x_1+x_2+x_3}=1} p_{x_1}p_{x_2}p_{x_3} \approx 1/m$.

Therefore, the result of proposition 5.5.1 does not seem to be representative in certain cases, e.g. when the X_n's have continuous densities.

A solution To avoid this problem, one can transform the random variables. Indeed, the previous result will change completely if one multiplies each X_t by a real number α_t modulo 1: $X'_t = \overline{\alpha_t X_t}$ and if one uses $\overline{X'_1+X'_2+X'_3}$. For example, one can choose $\alpha_1 = 9875465559300025458$ in the example of section 11.2.5 with $X_s = F(s,1)$. One can also choose $\alpha_1 = 9875465559300025458 * m/10$.

Indeed, if there are continuous peaks, i.e points of concentrations, this multiplication will remove them by distributing the probabilities in a enough random way. One will be able to thus apply proposition 5.5.1 which uses probabilities taken randomly.

For example suppose that $\alpha_t = 1$ for s=1,2,...,S-1, that the $X_s \in F^*(m)$ have the same distribution and that $\alpha_S = a = fi_{n_0-1}$ and $m = fi_{n_0}$: that is $T(x) \equiv ax$ modulo m is a Fibonacci congruence. Then, $p'_x = P\{X'_S = x\} = P\{\overline{T}(X_S) = x\} = P\{X_S = \overline{T}^{-1}(x)\} = p_{\overline{T}^{-1}(x)}$.

Then, $\overline{X_1 + ... + X_{S-1} + X'_S}$ has a probability close to the uniform distribution. Indeed,

$$\sum_{\overline{x_1+x_2+...+x_S}=y} p_{x_1}p_{x_2}.....p'_{x_S} = \sum_{\overline{x_1+x_2+...+x_S}=y} p_{x_1}p_{x_2}.....p_{\overline{T}^{-1}(x_S)}$$

$$= \sum_{(x_1,....,x_{S-1})} p_{x_1}p_{x_2}.....p_{x_{S-1}}p_{\overline{T}^{-1}(y-x_1-x_2-....-x_{S-1})}$$

$$= \sum_{t=0}^{m-1} \sum_{(x_1,....,x_{S-1}) \,:\, \overline{x_1+...+x_{S-1}}=t} p_{x_1}p_{x_2}.....p_{x_{S-1}}p_{\overline{T}^{-1}(y-t)}$$

$$= \sum_{t=0}^{m-1} \Pi_t p_{\overline{T}^{-1}(y-t)} \ ,$$

where

$$\Pi_t = \sum_{(x_1,....,x_{S-1}) \,:\, \overline{x_1+...+x_{S-1}}=t} p_{x_1}p_{x_2}.....p_{x_{S-1}} \ .$$

By the properties of Fibonacci congruences, (cf Chapter 7), the $\overline{T^{-1}(y-t)}$ behave as if they were chosen randomly compared to the continuous case. Then, one can regard that $p_{\overline{T}^{-1}(y-t)}$ and Π_t are independent.

Let us remind that in the case of sequences of N random variable X_t and Y_t independent,

$$1/N \sum_{t=0}^{N} X_t Y_t - \Big(1/N \sum_{t=0}^{N} X_t\Big)\Big(1/N \sum_{t=0}^{N} Y_t\Big) \xrightarrow{P} 0 \ \ as \ N \to \infty \ .$$

Therefore,

$$(1/m) \sum_{t=0}^{m-1} \Pi_t p_{\overline{T}^{-1}(y-t)} - \Big((1/m) \sum_{t=0}^{m-1} \Pi_t\Big)\Big((1/m) \sum_{t=0}^{m-1} p_{\overline{T}^{-1}(y-t)}\Big) \approx 0 \ .$$

Now

$$(1/m) \sum_{t=0}^{m-1} \Pi_t = (1/m) \sum_{(x_1,....,x_{S-1})} p_{x_1}p_{x_2}.....p_{x_{S-1}}$$

$$= (1/m)\Big(\sum_{x_1} p_{x_1}\Big).....\Big(\sum_{x_{S-1}} p_{x_{S-1}}\Big) = 1/m \ .$$

and

$$(1/m)\sum_{t=0}^{m-1} p_{\overline{T}^{-1}(y-t)} = 1/m \ .$$

Therefore,

$$(1/m)\sum_{t=0}^{m-1} \Pi_t p_{\overline{T}^{-1}(y-t)} \approx \Big((1/m)\sum_{t=0}^{m-1} \Pi_t\Big)\Big((1/m)\sum_{t=0}^{m-1} p_{\overline{T}^{-1}(y-t)}\Big) = 1/m^2 \ .$$

Therefore,

$$\sum_{\overline{x_1+x_2+...+x_S}=y} p_{x_1}p_{x_2}.....p_{x_S} \approx 1/m \ .$$

Therefore $\overline{X_1 + ... + X_{S-1} + X'_S}$ has a distribution close to the uniform distribution.

As a matter of fact, the multiplication by α_t modulo 1 defines a permutation if α_t is suitably selected.

But in this case, one has again the problem of the choice of the permutations: the permutations too simple are not appropriate. Is this case here? This problem is not so simple. On the one hand, Knuth ([1]) explains why one cannot use permutations built by algorithm (cf also section 2.1.1).

On the other hand, one understands in chapter 7 that the multiplication corresponding to a Fibonacci congruence is a good permutation. Because of this result, one can always regard that the multiplication by α_S modulo 1 corresponds to a good permutation if α_S is suitably selected.

But what interests us here be the sums $\sum_{\overline{x_1+...+x_S}=y} p_{x_1}p_{x_2}.....p_{x_{S-1}}p_{T^{-1}(x_S)}$. Now, in the numerical studies that we have made, the $p_{x_1}p_{x_2}.....p_{x_{S-1}}p_{T^{-1}(x_S)}$ are distributed in a enough random way so that one can considers that one is under the assumptions of proposition 5.5.1.

Return to the continuous case Let us check if we solved the case studied at the beginning of this subsection by multiplying the X_t by coefficients α_t.

This result seems correct: if one multiplies X_t by a real number α_0 chosen randomly, one multiplies normally by an irrational number. Now, this irrational number defines himself a random sequence (for example cf page 158, theorem F [1]).

Model associated to a sample Let us remind that to consider the set of all the possible probabilities is normal because there is a multitude of models associated with a sample and a priori, one does not know which is the best model which one can choose. However, the model which corresponds to probabilities chosen randomly is natural if one takes samples resulting from texts for example.

Of course, if $N << m$, one can also choose a model with continuous density. But it is a very particular model and a priori it is not logically appropriate if text is used.

5.5.9 Secund type of assumptions

In this section, we suppose again p=1 and that the X_s are independent. But we study by another way the numbers $v^i_{x_n}$ defined in notations 5.5.3.

Notations 5.5.4 *Let X_i, i=1,2,...,S, be a sequence of independent random variables with values in $\{0,1,...,N-1\}$. For all $s \in \{1,2,...,S\}$, we set $p"^s_{x^s_n} = P\{X_s = x^s_n\}$.*

Hypotheses

Then, we assume that $p"^i_{x_n} = P"^i_{x_n}(\omega_7)$ where $p"^i_{x_n} = (1/N)[1 + r^i_N(v^i_{x_n} - v^i_N)]$ and where $v^i_{x_n} = V^i_{x_n}(\omega_7)$ is a realization of an IID sequence defined by the following way.

Hypothesis 5.5.4 *For all $i \in \{1,2,...,S\}$, we assume that $v^i_{x_n}$ is a realization of an IID sequence of random variables $V^i_{x_n}$ defined on a probability space $(\Omega_7, \mathcal{A}_7, Proba_7)$ such that $-1 \leq V^i_{x_n} \leq N-1$ and $\mathbb{E}\{V^i_{x_n}\} = 0$.*

Then, we set $v^i_N = (1/N)\sum_{x_s} v^i_{x_s}$ and $V^i_N = (1/N)\sum_{x_s} V^i_{x_s}$.

Then, the following results holds

Lemma 5.5.7 *There exists a sequence of random variables $0 < R^i_N \leq 1$ such that $-1 \leq R^i_N(V^i_{x_n} - V^i_n) \leq N-1$ and $R^i_N \xrightarrow{P} 1$ as $N \to \infty$.*

Proof : Because $-1 \leq V^i_{x_n} \leq N-1$, one can write $-1-e \leq V^i_{x_n} - V^i_N \leq N-1+e$ where $e > 0$. Then, one can write $-1 \leq R^i_N(V^i_{x_n} - V^i_n) \leq N-1$ where $0 < R^i_N \leq 1$. By the CLT, $V^i_N \xrightarrow{P} 0$. Therefore, $R^i_N \xrightarrow{P} 1$. ■

Then, we can define probabilities over $F^*(N)$.

Proposition 5.5.4 *For all $x_n \in F^*(N)$, we set $P"^i_{x_n} = (1/N)[1 + R^i_N(V^i_{x_n} - V^i_N)]$. Then, $0 \leq P"^i_{x_n} \leq 1$ and $\sum_{x_n} P"^i_{x_n} = 1$*

Proof : We have $0 \leq P"^i_{x_n} \leq 1$. Moreover, $\sum_{x_n} P"^i_{x_n} = \sum_{x_n}(1/N)[1 + R^i_N(V^i_{x_n} - V^i_N)] = \sum_{x_n}(1/N) + (R^i_N/N)\sum_{x_n}(V^i_{x_n} - V^i_N)] = 1$. ■

Then, we assume that the following hypothesis holds.

Hypothesis 5.5.5 *For all $i \in \{1,2,...,S\}$, we assume that $p"^i_{x_n}$ is a realization of the sequence of random variables $P"^i_{x_n}$ defined over $(\Omega_7, \mathcal{A}_7, Proba_7)$: $p"^i_{x_n} = P"^i_{x_n}(\omega_7)$ where $P"^i_{x_n} = (1/N)[1 + R^i_N(V^i_{x_n} - V^i_N)]$.*

Then, we have the following theorem.

Theorem 6 *Assume that, for all $s \in \{1, 2, ..., S\}$, the variance of V_1^s is $\sigma_{V_s}^2$. Then, with a probability greater than $1 - \Gamma(b)$ approximately,*

$$P\{\overline{X_1 + ... + X_S} = y\} = \frac{1}{N}\Big[1 + \frac{b.Ob(1)\sigma_{V_1}....\sigma_{V_S}}{\sqrt{N^{S-1}}}\Big] .$$

Remark 5.5.8 *If P_x^i has a distribution similar to that of $P_x^i = \frac{P_x'^i}{\sum_{t=1}^N P_t'^i}$ when $P_x'^i$ has the uniform distribution, then $\sigma_V^2 = O(1)$.*

Indeed, $NP_{x_n}^i - 1 = V_{x_n}^i$ and $p_x^i \approx \frac{p_x'^i}{(N/2)[1+O(1)/\sqrt{N}]} = O(1/N)$ with a probability very close to 1. Moreover, $\mathbb{E}\{V_{x_n}^i\} = 0$. For example, one can choose $\sigma_V^2 \leq 1$.

Proof of theorem 6

At first, the following proposition holds.

Proposition 5.5.5 *The following equality holds :*
$P\{\overline{X_1 + ... + X_S} = y\} = \frac{1}{N} + \frac{r_N^1.....r_N^S}{N^S}\sum_{\overline{x_1+...+x_S}=y}(v_{x_1}^1 - v_N^1).....(v_{x_S}^S - v_N^S)$.

Proof At first, $P\{\overline{X_1 + X_2 + ... + X_S} = y\} = \sum_{\overline{x_1+....+x_S}=y} p"_{x_1}....p"_{x_S}$
$= (1/N^S)\sum_{\overline{x_1+...+x_S}=y}[1 + r_N^1(v_{x_n}^1 - v_N^1)].....[1 + r_N^S(v_{x_n}^S - v_N^S)]$.

Now, $[1 + r_N^1(v_{x_1}^1 - v_N^1)].....[1 + r_N^S(v_{x_S}^S - v_N^S)]$

$= 1 + \big[r_N^1(v_{x_1}^1 - v_N^1) + ... + r_N^S(v_{x_n}^S - v_N^S)\big]$
$+.....$
$+ \sum_{i_1<i_2<....<i_q} r_N^{i_1}(v_{x_{i_1}}^{i_1} - v_N^{i_1})r_N^{i_2}(v_{x_{i_2}}^{i_2} - v_N^{i_2}).....r_N^{i_q}(v_{x_{i_q}}^{i_q} - v_N^{i_q})$
$+.....$
$+r_N^1(v_{x_1}^1 - v_N^1)r_N^2(v_{x_2}^2 - v_N^2).....r_N^S(v_{x_S}^S - v_N^S)$.

We deduce the proposition by using the following lemma (5.5.9). ■

Lemma 5.5.9 *Suppose $q < S$. Then,*
$\sum_{\overline{x_1+....+x_S}=y}\big[\sum_{i_1<i_2<....<i_q} r_N^{i_1}(v_{x_{i_1}}^{i_1} - v_N^{i_1}).....r_N^{i_q}(v_{x_{i_q}}^{i_q} - v_N^{i_q})\big] = 0$.

Proof We have
$\sum_{\overline{x_1+....+x_S}=y}\big[\sum_{i_1<i_2<....<i_q} r_N^{i_1}(v_{x_{i_1}}^{i_1} - v_N^{i_1}).....r_N^{i_q}(v_{x_{i_q}}^{i_q} - v_N^{i_q})\big] = 0$
$= \sum_{i_1<i_2<....<i_q}\big[\sum_{\overline{x_1+....+x_S}=y} r_N^{i_1}(v_{x_{i_1}}^{i_1} - v_N^{i_1}).....r_N^{i_q}(v_{x_{i_q}}^{i_q} - v_N^{i_q})\big]$.

For example, if $i_q < S$,

$\sum_{\overline{x_1+....+x_S}=y}(v^{i_1}_{x_{i_1}}-v^{i_1}_N).....(v^{i_q}_{x_{i_q}}-v^{i_q}_N)$
$=\sum_{x_{i_1}}\sum_{x_{i_2}}....\sum_{x_{i_{S-1}}}\sum_{x_S=\overline{y-x_1+....-x_{S-1}}}(v^{i_1}_{x_{i_1}}-v^{i_1}_N).....(v^{i_q}_{x_{i_q}}-v^{i_q}_N)$
$=\sum_{x_{i_1}}\sum_{x_{i_2}}....\sum_{x_{i_{S-1}}}(v^{i_1}_{x_{i_1}}-v^{i_1}_N).....(v^{i_q}_{x_{i_q}}-v^{i_q}_N)$
$=\sum\Big[\ \Big[\sum_{x_{i_1}}(v^{i_1}_{x_{i_1}}-v^{i_1}_N)\Big].....\Big[\sum_{x_{i_q}}(v^{i_q}_{x_{i_q}}-v^{i_q}_N)\Big]\ \Big]$
$=0$ because $\sum_{x_{i_1}}v^{i_1}_{x_{i_1}}=Nv^{i_1}_N$. ■

Then, one can prove that $\sum_{\overline{x_1+....+x_S}=y}\big[\sum_{i_1<i_2<....<i_S}V^{i_1}_{x_{i_1}}V^{i_2}_{x_{i_2}}....V^{i_S}_{x_{i_S}}\big]$ has asymptotically a normal distribution by using theorem 2.

Proposition 5.5.6 *Under the hypothesis 5.5.4,* $\frac{\sum_{\overline{x_{i_1}+....+x_{i_S}}=y}V^1_{x_{i_1}}V^2_{x_{i_2}}....V^S_{x_{i_S}}}{N^{(S-1)/2}}$ *has asymptotically a distribution* $N(0,\sigma^2_{V_1}.....\sigma^2_{V_S})$.

Proof By theorem 2, it is sufficient that

$$\frac{\sum_{r\neq s}\big[\mathbb{E}\{(X_s)^2(X_r)^2\}-\mathbb{E}\{(X_s)^2\}\mathbb{E}\{(X_r)^2\}\big]}{(N_0)^2}\to 0$$

$$\frac{\sum_{t_1<t_2<.....<t_p}\mathbb{E}\{X_{t_1}X_{t_2}.....X_{t_p}\}}{(N_0)^{p/2}}\to 0\ .$$

Then, we apply this theorem with $X_{t_s}=V^1_{i_1}....V^{S-1}_{i_{S-1}}V^S_{\overline{y-i_1-.....-i_{S-1}}}$ and $N_0=N^{S-1}$.

The first relation is obvious. For example, if S=3, this relation is equivalent to the convergence of $(1/N^4)\sum_{r\neq s}\big[\mathbb{E}\{(X_s)^2(X_r)^2\}-\mathbb{E}\{(X_s)^2\}\mathbb{E}\{(X_r)^2\}\big]$, which is equivalent to the convergence of

$$\frac{\Big|\sum_{(i,j)\neq(i',j')}\big[\mathbb{E}\{(V^1_iV^2_jV^3_{\overline{y-i-j}})^2(V^1_{i'}V^2_{j'}V^3_{\overline{y-i'-j'}})^2\}}{N^4}.$$
$$\frac{-\mathbb{E}\{(V^1_iV^2_jV^3_{\overline{y-i-j}})^2\}\mathbb{E}\{(V^1_{i'}V^2_{j'}V^3_{\overline{y-i'-j'}})^2\}\big]\Big|}{N^4}.$$

Now, in order that

$$\mathbb{E}\{(V^1_iV^2_jV^3_{\overline{y-i-j}})^2(V^1_{i'}V^2_{j'}V^3_{\overline{y-i'-j'}})^2\}\neq\mathbb{E}\{(V^1_iV^2_jV^3_{\overline{y-i-j}})^2\}\mathbb{E}\{(V^1_{i'}V^2_{j'}V^3_{\overline{y-i'-j'}})^2\}\ ,$$

it is necessary that $i=i'$ or $j=j'$. Therefore, at the maximum, there is $2N^3$ such $V^1_iV^2_jV^3_{\overline{y-i-j}}V^1_{i'}V^2_{j'}V^3_{\overline{y-i'-j'}}$. Then, there exists a constant C^2_3 such that

$$\frac{\Big|\sum_{(i,j)\neq(i',j')}\big[\mathbb{E}\{(V^1_iV^2_jV^3_{\overline{y-i-j}})^2(V^1_{i'}V^2_{j'}V^3_{\overline{y-i'-j'}})^2\}}{N^4}$$

$$\frac{-\mathbb{E}\{((V_i^1 V_j^2 V^3_{\overline{y-i-j}}))^2\}\mathbb{E}\{(V_{i'}^1 V_{j'}^2 V^3_{\overline{y-i'-j'}})^2\}\big]\Big|}{N^4} \leq \frac{2C_3^2}{N} \ .$$

In the general case, there is $(S-1)N * N^{2(S-2)}$ such

$$V_{i_1}^1 .... V_{i_{S-1}}^{S-1} V^S_{\overline{y-i_1-.....-i_{S-1}}} V_{i'_1}^1 .... V_{i'_{S-1}}^{S-1} V^S_{\overline{y-i'_1-.....-i'_{S-1}}}$$

at the maximum. Therefore, there exists a constant C_S^2 such that

$$\frac{\sum_{r\neq s} \left|\mathbb{E}\{(X_s)^2(X_r)^2\} - \mathbb{E}\{(X_s)^2\}\mathbb{E}\{(X_r)^2\}\right|}{N^{2(S-1)}} \leq \frac{SC_S^2}{N} \to 0 \ .$$

Now we study the condition $p!$ $\frac{\sum_{t_1<t_2<.....<t_p} \mathbb{E}\{X_{t_1} X_{t_2} ..... X_{t_p}\}}{(N^{S-1})^{p/2}} \to (N_2)^p \mu_p$.

First, assume S=2 : in this case, $X_{t_1} = V^1_{x_n^1} V^2_{\overline{y-x_n^1}}$. Then, $\mathbb{E}\{X_{t_1}.....X_{t_p}\} = \mathbb{E}\{V^1_{x^1_{n_1}} V^2_{\overline{y-x^1_{n_1}}}.....V^1_{x^1_{n_p}} V^2_{\overline{y-x^1_{n_p}}}\} = \mathbb{E}\{V^1_{x^1_{n_1}}\}.....\mathbb{E}\{V^1_{x^1_{n_p}}\}\mathbb{E}\{V^2_{\overline{y-x^1_{n_1}}}.....V^2_{\overline{y-x^1_{n_p}}}\} = 0$ because the x_n^1 are all dissimilar.

Assume S=3 : in this case, $X_{t_1} = V^1_{x_n^1} V^2_{x^2_{n_2}} V^3_{\overline{y-x^1_{n_1}-x^2_{n_2}}}$. Then,

$$\mathbb{E}\{X_{t_1} X_{t_2} ..... X_{t_p}\}$$

$$= \mathbb{E}\{V^1_{x_{n_1}} V^2_{x'_{n_1}} V^3_{\overline{y-x_{n_1}-x'_{n_1}}} V^1_{x_{n_2}} V^2_{x'_{n_2}} V^3_{\overline{y-x_{n_2}-x'_{n_2}}} ..... V^1_{x_{n_p}} V^2_{x'_{n_p}} V^3_{\overline{y-x_{n_p}-x'_{n_p}}}\}$$

$$= \mathbb{E}\{V^1_{x_{n_1}} .....V^1_{x_{n_p}}\}\mathbb{E}\{V^2_{x'_{n_1}} .....V^2_{x'_{n_p}}\}\mathbb{E}\{V^3_{\overline{y-x_{n_1}-x'_{n_1}}} .....V^3_{\overline{y-x_{n_p}-x'_{n_p}}}\} \ .$$

If p=2, $\mathbb{E}\{X_{t_1} X_{t_2}\} = \mathbb{E}\{V^1_{x_{n_1}} V^1_{x_{n_2}}\}\mathbb{E}\{V^2_{x'_{n_1}} V^2_{x'_{n_2}}\}\mathbb{E}\{V^3_{\overline{y-x_{n_1}-x'_{n_1}}} V^3_{\overline{y-x_{n_p}-x'_{n_p}}}\}$. Because $t_1 < t_2$, $x_{n_1} \neq x_{n_2}$ or $x'_{n_1} \neq x'_{n_2}$. Then, $\mathbb{E}\{X_{t_1} X_{t_2}\} = 0$.

If p=3,

$$\mathbb{E}\{X_{t_1} X_{t_2} X_{t_3}\}$$

$$= \mathbb{E}\{V^1_{x_{n_1}} V^1_{x_{n_2}} V^1_{x_{n_3}}\}\mathbb{E}\{V^2_{x'_{n_1}} V^2_{x'_{n_2}} V^2_{x'_{n_3}}\}\mathbb{E}\{V^3_{\overline{y-x_{n_1}-x'_{n_1}}} V^3_{\overline{y-x_{n_2}-x'_{n_2}}} V^3_{\overline{y-x_{n_3}-x'_{n_3}}}\} \ .$$

In order that $\mathbb{E}\{X_{t_1} X_{t_2} X_{t_3}\} \neq 0$, it is necessary that $x_{n_1} = x_{n_2} = x_{n_3}$ and $x'_{n_1} = x'_{n_2} = x'_{n_3}$. Then, $X_{t_1} = X_{t_2} = X_{t_3}$. It is impossible.

If p=4,

$$\mathbb{E}\{X_{t_1}X_{t_2}X_{t_3}X_{t_4}\}$$
$$= \mathbb{E}\{V^1_{x_{n_1}}V^1_{x_{n_2}}V^1_{x_{n_3}}V^1_{x_{n_4}}\}\mathbb{E}\{V^2_{x'_{n_1}}V^2_{x'_{n_2}}V^2_{x'_{n_3}}V^2_{x'_{n_4}}\}$$
$$*\mathbb{E}\{V^3_{\overline{y-x_{n_1}-x'_{n_1}}}V^3_{\overline{y-x_{n_2}-x'_{n_2}}}V^3_{\overline{y-x_{n_3}-x'_{n_3}}}V^3_{\overline{y-x_{n_4}-x'_{n_4}}}\}.$$

In order that $\mathbb{E}\{X_{t_1}X_{t_2}X_{t_3}X_{t_4}\} \neq 0$, it is necessary that, or $x_{n_1} = x_{n_2}$ and $x_{n_3} = x_{n_4}$, or $x_{n_1} = x_{n_3}$ and $x_{n_2} = x_{n_4}$, or $x_{n_1} = x_{n_4}$ and $x_{n_3} = x_{n_2}$

For example, assume $x_{n_1} = x_{n_2}$, $x_{n_3} = x_{n_4}$. Then, we had to assume also $x'_{n_1} = x'_{n_3}$ and $x'_{n_2} = x'_{n_4}$. *Now, we assume that these relations hold.*

Then, in order that

$$\mathbb{E}\{V^1_{x_{n_1}}V^1_{x_{n_2}}V^1_{x_{n_3}}V^1_{x_{n_4}}\}\mathbb{E}\{V^2_{x'_{n_1}}V^2_{x'_{n_2}}V^2_{x'_{n_3}}V^2_{x'_{n_4}}\}$$
$$\mathbb{E}\{V^3_{\overline{y-x_{n_1}-x'_{n_1}}}V^3_{\overline{y-x_{n_2}-x'_{n_2}}}V^3_{\overline{y-x_{n_3}-x'_{n_3}}}V^3_{\overline{y-x_{n_4}-x'_{n_4}}}\} \neq 0\ ,$$

it is necessary that

OR 1) $\overline{y-x_{n_1}-x'_{n_1}} = \overline{y-x_{n_2}-x'_{n_2}}$ and $\overline{y-x_{n_3}-x'_{n_3}} = \overline{y-x_{n_4}-x'_{n_4}}$,
OR 2) $\overline{y-x_{n_1}-x'_{n_1}} = \overline{y-x_{n_3}-x'_{n_3}}$ and $\overline{y-x_{n_2}-x'_{n_2}} = \overline{y-x_{n_4}-x'_{n_4}}$,
OR 3) $\overline{y-x_{n_1}-x'_{n_1}} = \overline{y-x_{n_4}-x'_{n_4}}$ and $\overline{y-x_{n_2}-x'_{n_2}} = \overline{y-x_{n_3}-x'_{n_3}}$.

If 1) holds, $y - x'_{n_1} \equiv y - x'_{n_2}$. Then, $x'_{n_1} \equiv x'_{n_2}$. Therefore, $x'_{n_1} = x'_{n_2}$. Then, $X_{t_1} = V^1_{x_{n_1}}V^2_{x'_{n_1}}V^3_{\overline{y-x_{n_1}-x'_{n_1}}} = V^1_{x_{n_2}}V^2_{x'_{n_2}}V^3_{\overline{y-x_{n_2}-x'_{n_2}}} = X_{t_2}$: it is impossible.

If 2) holds, $x_{n_1} \equiv x_{n_3}$. Then, $x_{n_1} = x_{n_3}$. Then, $V^1_{x_{n_1}}V^2_{x'_{n_1}}V^3_{\overline{y-x_{n_1}-x'_{n_1}}} = V^1_{x_{n_3}}V^2_{x'_{n_3}}V^3_{\overline{y-x_{n_3}-x'_{n_3}}}$: it is impossible.

If 3) holds, $x_{n_1} + x'_{n_1} \equiv x_{n_3} + x'_{n_2}$ and $x_{n_1} + x'_{n_2} \equiv x_{n_3} + x'_{n_1}$. Then, $x'_{n_1} - x'_{n_2} \equiv x_{n_3} - x_{n_1}$ and $x_{n_1} - x_{n_3} \equiv x'_{n_1} - x'_{n_2}$. Therefore, $2(x'_{n_1} - x'_{n_2}) \equiv 0$ and $2(x_{n_1} - x_{n_3}) \equiv 0$. If N is odd, $x'_{n_1} = x'_{n_2}$ and $x_{n_1} = x_{n_3}$: it is impossible.

If N is even, $x'_{n_1} - x'_{n_2} = \delta_1(N/2)$ and $x_{n_1} - x_{n_3} = \delta_2(N/2)$ where $\delta_s = 0$, -1 or 1.

Therefore, there are $\frac{C'_0N^4}{N^2}$ possible variables $X_{t_1} = V^1_{x_{n_1}}V^2_{x'_{n_1}}V^3_{\overline{y-x_{n_1}-x'_{n_1}}}$, $X_{t_2} = V^1_{x_{n_2}}V^2_{x'_{n_2}}V^3_{\overline{y-x_{n_2}-x'_{n_2}}}$, $X_{t_3} = V^1_{x_{n_3}}V^2_{x'_{n_3}}V^3_{\overline{y-x_{n_3}-x'_{n_3}}}$, $X_{t_4} = V^1_{x_{n_4}}V^2_{x'_{n_4}}V^3_{\overline{y-x_{n_4}-x'_{n_4}}}$ such that $\mathbb{E}\{X_{t_1}X_{t_2}X_{t_3}X_{t_4}\} \neq 0$. Therefore,

$$\frac{\sum_{t_1<t_2<t_3<t_4}\mathbb{E}\{X_{t_1}X_{t_2}X_{t_3}X_{t_4}\}}{(N_0)^{4/2}} \leq \frac{\sum_{t_1<t_2<t_3<t_4}\mathbb{E}\{X_{t_1}X_{t_2}X_{t_3}X_{t_4}\}}{(N^2)^2} < \frac{C'_0N^2}{N^4} \to 0\ .$$

In the general case, $X_{t_r} = V^1_{x^1_{n_r}} V^2_{x^2_{n_r}} .....V^{S-1}_{x^{S-1}_{n_r}} V^S_{\overline{y-x^1_{n_r}-....-x^{S-1}_{n_r}}}$. Then,

$$\mathbb{E}\{X_{t_1}X_{t_2}.....X_{t_p}\}$$

$$= \mathbb{E}\Big\{V^1_{x^1_{n_1}}....V^1_{x^1_{n_p}}\Big\}.....\mathbb{E}\Big\{V^{S-1}_{x^{S-1}_{n_1}}.....V^{S-1}_{x^{S-1}_{n_p}}\Big\}\mathbb{E}\Big\{V^S_{\overline{y-x^1_{n_1}-...-x^{S-1}_{n_1}}}.....V^S_{\overline{y-x^1_{n_{S-1}}-...-x^{S-1}_{n_{S-1}}}}\Big\}.$$

In order that $\mathbb{E}\{X_{t_1}X_{t_2}.....X_{t_p}\} \neq 0$, it is necessary that, for all t=1,2,...,S-1, and all n_r, there is one equation $x^t_{n_r} = x^t_{n_{r'}}$. At the maximum, there are $C_0(N^{S-1})^{p/2}$ possible variables $V^1_{x^1_{n_r}} V^2_{x^2_{n_r}} .....V^{S-1}_{x^{S-1}_{n_r}}$, r=1,2,.....,p, such that $\mathbb{E}\Big\{V^1_{x^1_{n_1}}....V^1_{x^1_{n_p}}\Big\}.....\mathbb{E}\Big\{V^{S-1}_{x^{S-1}_{n_1}}.....V^{S-1}_{x^{S-1}_{n_p}}\Big\} \neq 0$.
Moreover, there are$\lfloor p/2 \rfloor$ equations $\overline{y - x^1_{n_r} - .... - x^{S-1}_{n_r}} = \overline{y - x^1_{n_{r'}} - .... - x^{S-1}_{n_{r'}}}$: at the maximum there are $\frac{C'_0(N^{S-1})^{p/2}}{N}$ possible variables $V^1_{x^1_{n_r}} V^2_{x^2_{n_r}} .....V^{S-1}_{x^{S-1}_{n_r}}$ $V^S_{\overline{y-x^1_{n_r}-....-x^{S-1}_{n_r}}}$. Therefore

$$\frac{\sum_{t_1<t_2<.....<t_p}\mathbb{E}\{X_{t_1}X_{t_2}.....X_{t_p}\}}{(N_0)^{p/2}} \leq \frac{\sum_{t_1<t_2<.....<t_p}\mathbb{E}\{X_{t_1}X_{t_2}.....X_{t_p}\}}{(N^{S-1})^{p/2}} < \frac{C'_0}{N} .$$

Then all conditions of theorem 2 hold. Then, $\frac{\sum_{i=1}^{n_0} X_n}{\sqrt{N_0}} \xrightarrow{D} N(0, \sigma^2_{V_1}....\sigma^2_{V_S})$ because $\mathbb{E}\{X^2_{t_r}\} = \mathbb{E}\Big\{\Big(V^1_{x^1_{n_r}} V^2_{x^2_{n_r}} ...V^{S-1}_{x^{S-1}_{n_r}} V^S_{\overline{y-x^1_{n_r}-...-x^{S-1}_{n_r}}}\Big)^2\Big\} = \prod_{r=1}^S \mathbb{E}\Big\{(V^r_1)^2\Big\}$.
■

Now, one can assume that $\sum_{x^s_n} v^s_{x^s_n} = 0$.

Proposition 5.5.7 *Under the previous assumptions,*

$$\frac{1}{\sqrt{N^{S-1}}} \sum_{x^1_n, x^2_n, ..., x^{S-1}_n} \Big[\prod_{t=1}^{S-1} R^t_N(V^t_{x^t_n} - V^t_N)\Big] R^S_N(V^S_{\overline{y-x^1_n-....-x^{S-1}_n}} - V^S_N)$$

has asymptotically the distribution $N(0, \sigma^2_{V_1}.....\sigma^2_{V_S})$.

Proof Assume S=2. Then,

$$\frac{1}{\sqrt{N}} \sum_{x^1_n} R^1_N(V^1_{x^1_n} - V^1_N) R^2_N(V^2_{\overline{y-x^1_n}} - V^2_N)$$

$$= \frac{R_N^1 R_N^2}{\sqrt{N}} \sum_{x_n^1} V_{x_n^1}^1 [V_{\overline{y-x_n^1}}^2 - V_N^2] - \frac{R_N^1 R_N^2 V_N^1}{\sqrt{N}} \sum_{x_n^2} [V_{x_n^2}^2 - V_N^2]$$

$$= \frac{R_N^1 R_N^2}{\sqrt{N}} \sum_{x_n^1} V_{x_n^1}^1 [V_{\overline{y-x_n^1}}^2 - V_N^2]$$

$$= \frac{R_N^1 R_N^2}{\sqrt{N}} \sum_{x_n^1} V_{x_n^1}^1 V_{\overline{y-x_n^1}}^2 - \frac{R_N^1 R_N^2}{\sqrt{N}} V_N^2 \sum_{x_n^1} V_{x_n^1}^1 \ ,$$

where $\frac{R_N^1 R_N^2}{\sqrt{N}} \sum_{x_n^1} V_{x_n^1}^1 V_{\overline{y-x_n^1}}^2$ and $\frac{1}{\sqrt{N}} \sum_{x_n^1} V_{x_n^1}^1$ have asymptotically a normal distribution (cf proposition 5.5.6). Moreover, V_N^2 converges in probability to 0.

Then, $\frac{1}{\sqrt{N}} \sum_{x_n^1} R_N^1 (V_{x_n^1}^1 - V_N^1) R_N^2 (V_{\overline{y-x_n^1}}^2 - V_N^2)$ has asymptotically the distribution $N(0, \sigma_{V_1}^2 \sigma_{V_2}^2)$.

Assume S=3. Then,

$$\frac{1}{N} \sum_{x_n^1, x_n^2} R_N^1 (V_{x_n^1}^1 - V_N^1) R_N^2 (V_{x_n^2}^2 - V_N^2) R_N^3 (V_{\overline{y-x_n^1-x_n^2}}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 (V_{x_n^2}^2 - V_N^2)(V_{\overline{y-x_n^1-x_n^2}}^3 - V_N^3)$$

$$- \frac{R_N^1 R_N^2 R_N^3 V_N^1}{N} \sum_{x_n^1, x_n^2} (V_{x_n^2}^2 - V_N^2)(V_{\overline{y-x_n^1-x_n^2}}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 (V_{x_n^2}^2 - V_N^2)(V_{\overline{y-x_n^1-x_n^2}}^3 - V_N^3)$$

$$- \frac{R_N^1 R_N^2 R_N^3 V_n^1}{N} \sum_{x_n^2, x_n^3} (V_{x_n^2}^2 - V_N^2)(V_{x_n^3}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 (V_{x_n^2}^2 - V_N^2)(V_{\overline{y-x_n^1-x_n^2}}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2 (V_{\overline{y - x_n^1 - x_n^2}}^3 - V_N^3)$$
$$- \frac{R_N^1 R_N^2 R_N^3 V_N^2}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 (V_{\overline{y - x_n^1 - x_n^2}}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2 (V_{\overline{y - x_n^1 - x_n^2}}^3 - V_N^3)$$
$$- \frac{R_N^1 R_N^2 R_N^3 V_N^2}{N} \sum_{x_n^1, x"_n^3} V_{x_n^1}^1 (V_{x"_n^3}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2 (V_{\overline{y - x_n^1 - x_n^2}}^3 - V_N^3)$$

$$= \frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2 V_{\overline{y - x_n^1 - x_n^2}}^3 - \frac{R_N^1 R_N^2 R_N^3 V_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2$$

where $\frac{R_N^1 R_N^2 R_N^3}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2 V_{\overline{y - x_n^1 - x_n^2}}^3$ and $\frac{1}{N} \sum_{x_n^1, x_n^2} V_{x_n^1}^1 V_{x_n^2}^2$ have asymptotically a normal and a chi squared distributions. Moreover, V_N^3 converges in probability to 0.

Then, $\frac{1}{N} \sum_{x_n^1, x_n^2} R_N^1 (V_{x_n^1}^1 - V_N^1) R_N^2 (V_{x_n^2}^2 - V_N^2) R_N^3 (V_{\overline{y - x_n^1 - x_n^2}}^3 - V_n^3)$ has asymptotically the distribution $N(0, \sigma_{V_1}^2 \sigma_{V_2}^2 \sigma_{V_3}^2)$.

In the general case, we prove this proposition by the same way . ■

Proof 5.5.10 *Now we prove theorem 6*

By proposition 5.5.5, we know that

$$\left| P\{\overline{X_1 + ... + X_S} = y\} - \frac{1}{N} \right| \leq \frac{\left| r_N^{i_1} ... r_N^{i_S} \sum_{\overline{x_1 + ... + x_S} = y} (v_{x_1}^1 - v_N^1) ..... (v_{x_S}^S - v_N^S) \right|}{N^S},$$

where by proposition 5.5.7,

$$\frac{1}{\sqrt{N^{S-1}}} \sum_{x_n^1, x_n^2, ..., x_n^{S-1}} \Big[\prod_{t=1}^{S-1} R_N^t (V_{x_n^t}^t - V_n^t) \Big] R_N^S (V_{\overline{y - x_n^1 - .... - x_n^{S-1}}}^S - V_n^S)$$

has asymptotically the distribution $N(0, \sigma^2_{V_1}.....\sigma^2_{V_S})$. Then,

$$R^1_N.....R^S_N \ \frac{\sum_{\overline{x_1+...+x_S}=y}(V^1_{x_1}-V^1_N).....(V^S_{x_S}-V^S_N)}{\sqrt{N^{S-1}}}$$

has asymptotically the distribution $N(0, \sigma^2_{V_1}.....\sigma^2_{V_S})$. Then, with a probability greater than $1-\Gamma(b)$ approximately,

$$r^{i_1}_N.....r^{i_S}_N \ \frac{\sum_{\overline{x_1+...+x_S}=y}(v^1_{x_1}-v^1_N).....(v^S_{x_S}-v^S_N)}{N^S} = \frac{b.Ob(1)\sigma_{V_1}.....\sigma_{V_S}}{\sqrt{N^{S+1}}} \ . \ \blacksquare$$

5.5.10 Conclusion

By applying all the results of this chapter, we can admit that the convergence to the uniform distribution is very fast as soon as the X_t's are not concentrated near a small number of points and that probabilities are chosen randomly.

It remains to prove it completely. For that purpose, it would be necessary to make another study which can possibly turn out complex.

But for our conclusions, it is not very useful. There are indeed two solutions

1) We admit that the result is true in the set of probabilities with a probability infinitely close to 1 : cf chapter 8.

2) We admit that the hypotheses which we chose in section 7.2 are verified. In that case, it is a guess which remains to prove mathematically.

However everywhere where we were able to make numerical studies, in particular when q is small (cf figures of section 5.4.1) , we found this result without any possible error.

Moreover the comparison made with the sum of an IID sample involve consequences which seem clear: cf section 5.5.8.

5.6 Theoretical study of density

In this section, we confirm that it is more practical to use the XORLT than the CLT by another theoretical study : we compare the densities of the functions G(j) and H(j) defined in section 11.1.2.

For that purpose, the following lemmas are needed.

Lemma 5.6.1 *Let $X_m \in F(m)$ be a sequence of random variables. Assume that f_n is the probability density function of X_m with respect to μ_m. Suppose that $f_n \to f$ uniformly as $n \to \infty$ where f is a continuous function bounded on [0,1[.*

Then $X_m \xrightarrow{D} X$ where X has the probability density function f with respect to μ.

Proof Let $t \in [0,1[$. Then, $P\{X_m \in [0,t[\} = \int_0^t f(x).\mu_m(dx) + \int_0^t (f_m(x) - f).\mu_m(dx) = \int_{x=0}^t f(x).\mu_m(dx) + \epsilon(m)$ where $\epsilon(m) \to 0$ as $m \to \infty$.

Then, by theorem A, page 16 of [42], $P\{X_m \in [0,t[\} \to \int_{x=0}^t f(x).\mu(dx) = P\{X \in [0,t[\}$. ■

Lemma 5.6.2 *Let $X \in F(m)$ be a random variable. Let f_X be the probability density function of X with respect to μ_m. Then, $P\{X = x\} = f_X(x)\ \frac{1}{m}$.*

Of course $\int_{[0,1]} f_X(x).\mu_m(dx) = 1$.

Moreover, $|f_X(x) - f_X(x')| \leq K_0|x - x'|$ if and only if

$$|P\{X = x\} - P\{X = x'\}| \leq K_0|x - x'|/m.$$

Lemma 5.6.3 *Let x=k/m, x'=k'/m where $k \in \mathbb{N}$ and $k' \in \mathbb{N}$. Suppose that $|P\{mX = k\} - P\{mX = k+1\}| \leq K_1$. Then,*
1) $|P\{X = x\} - P\{X = x'\}| \leq mK_1|x - x'|$,
2) $|f_X(x) - f_X(x')| \leq m^2K_1|x - x'|$.

Proof We have $|P\{X = k/m\} - P\{X = k/m + 1/m\}| \leq K_1 = (mK_1)/m$.
Then, $|P\{X = k/m\} - P\{X = k'/m\}| \leq (mK_1)\frac{|k-k'|}{m}$.
Then, $|f_X(x) - f_X(x')| = m|P\{X = k/m\} - P\{X = k'/m\}| \leq m^2K_1|x - x'|$. ■

We apply these lemmas to the conditional densities of the sequences G(j) and H(j) (cf section 11.1.2).

Notations 5.6.1 *We keep the notations of the chapter 11 with $m_S = m$: $G(j) = \sum_{i=1}^S F(i,j) \in F^*(Sm)$, $H(j) = \overline{G(j)} \in F^*(m)$. We set $H_j = H(j)$ and $G_j = G(j)$.*

Let $P_{g_2,g_3,..}\{G_i = g\}$ be the conditional probability $P_{g_2,g_3,..}\{G_i = g\} = P\Big\{G_{i+j_1} = g \Big| G_{i+j_s} = g_s, s = 2,3,...\Big\}$.

Suppose that
$|P_{h_2,h_3,..}\{H_i = h\} - P_{h_2,h_3,..}\{H_i = h+1\}| \leq K_H$,
$|P_{g_2,g_3,..}\{G_i = g\} - P_{g_2,g_3,..}\{G_i = g+1\}| \leq K_G$.
Let $f_{H/m}$ and $f_{G/(Sm)}$ be the probability densities functions associated to the conditional probabilities of H_j/m and $G_j/(Sm)$, respectively. Let $K_{f_{H/m}}$ and $K_{f_{G/(Sm)}}$ the associated constant of Lipschitz. Then
$|f_{H/m}(x) - f_{H/m}(x')| \leq m^2K_H|x - x'| = K_{f_{H/m}}|x - x'|$.
$|f_{G/(Sm)}(x) - f_{G/(Sm)}(x')| \leq (m^2S^2K_G)|x - x'| = K_{f_{G/(Sm)}}|x - x'|$.
That is $K_{f_{H/m}} = m^2K_H$ and $K_{f_{G/(Sm)}} = m^2S^2K_G$.

Suppose that one can apply the proposition 4.1.1 with congruences T_H and T_G to H(j) and G(j). By this theorem, there exists constants $C_H = O\big(6.K_{f_{H/m}}\big)$ and $C_G = O\big(6.K_{f_{G/(Sm)}}\big)$ such that
$P\{T_H(H/m) \in I\} = L(I) + Ob(1)C_H/\sqrt{m}$,

$P\{T_G(G/(Sm)) \in I\} = L(I) + Ob(1)C_G/\sqrt{Sm}.$

Then, we have
$P\{T_H(H/m) \in I\} \approx L(I) + Ob(1)6.K_H m^{3/2}$,
$P\{T_G(G/(Sm)) \in I\} \approx L(I) + Ob(1)6.K_G(mS)^{3/2}$.

Now, in the numerical results which we obtain (cf section 5.4) K_H has the same order of size as K_G. It means that it be better to use the functions T_q (cf definition 1.3.5) with the sequence H(j) than with the sequence G(j) if we use the proposition 4.1.1. This conclusion is confirmed by the following result.

Proposition 5.6.1 *Let $f_{g_2,g_3,..}$ be the probability density function of $P_{g_2,g_3,..}$ with respect to μ^*_{Sm}. Let $f_{h_2,h_3,..}$ be the probability density function of $P_{h_2,h_3,..}$ with respect to μ^*_m. Let K^G such that*

$$\left|f_{g_2,g_3,..}(g) - f_{g_2,g_3,..}(g')\right| \leq K^G|g - g'| \ .$$

Then,

$$\left|f_{h_2,h_3,..}(h) - f_{h_2,h_3,..}(h')\right| \leq K^G|h - h'| \ .$$

Proof Clearly $\left|f_{g_2,g_3,..}\{g\} - f_{g_2,g_3,..}\{g'\}\right| \leq K^G|g - g'|$ involves that

$$mS.\left|P_{g_2,g_3,..}\{G_i = g\} - P_{g_2,g_3,..}\{G_i = g'\}\right| \leq K^G|g - g'| \ .$$

If $H_{i+j_t} = h_t$, then $G_{i+j_t} \in \cup_{s_t} g_t^{s_t}$ where $g_t^{s_t} = h_t + s_t m$, $s_t \in \{0, 1, ..., S-1\}$. We set $g_s = g_1^s$. Then, the following equalities hold

$$P_{h_2,h_3,..}\{H_i = h\} = \frac{P\Big\{\{H_{i+j_1} = h\} \cap \{H_{i+j_2} = h_2\} \cap \{H_{i+j_3} = h_3\} \cap ....\}}{P\Big\{\{H_{i+j_2} = h_2\} \cap \{H_{i+j_3} = h_3\} \cap ....\}}$$

$$= \frac{P\Big\{\{G_{i+j_1} \in \cup_s g_s\} \cap \{G_{i+j_2} \in \cup_{s_2} g_2^{s_2}\} \cap \{G_{i+j_3} \in \cup_{s_3} g_3^{s_3}\} \cap ....\Big\}}{P\Big\{\{G_{i+j_2} \in \cup_{s_2} g_2^{s_2}\} \cap \{G_{i+j_3} \in \cup_{s_3} g_3^{s_3}\} \cap ....\Big\}}$$

$$= \sum_{s=0}^{S-1} \frac{\sum_{s_2,s_3,...} P\Big\{\{G_{i+j_1} = g_s\} \cap \{G_{i+j_2} = g_2^{s_2}\} \cap \{G_{i+j_3} = g_3^{s_3}\} \cap ....\Big\}}{\sum_{s_2,s_3,...} P\Big\{\{G_{i+j_2} = g_2^{s_2}\} \cap \{G_{i+j_3} = g_3^{s_3}\} \cap ....\Big\}}$$

$$= \sum_{s=0}^{S-1} \sum_{s,s_2,s_3,\ldots} \frac{P\Big\{\{G_{i+j_1} = g_s\} \cap \{G_{i+j_2} = g_2^{s_2}\} \cap \ldots\Big\}}{\sum_{s_2,s_3,\ldots} P\Big\{\{G_{i+j_2} = g_2^{s_2}\} \cap \ldots\Big\}} \frac{P\Big\{\{G_{i+j_2} = g_2^{s_2}\} \cap \ldots.\Big\}}{P\Big\{\{G_{i+j_2} = g_2^{s_2}\} \cap \ldots.\Big\}}$$

$$= \sum_{s=0}^{S-1} \sum_{s_2,s_3,\ldots} \eta_{s_2,s_3,\ldots} P\Big\{G_{i+j_1} = g_s \Big|\ G_{i+j_2} = g_2^{s_2} \ldots.\Big\}$$

$$= \sum_{s=0}^{S-1} \sum_{s_2,s_3,\ldots} \eta_{s_2,s_3,\ldots} P_{g_2^{s_2},g_3^{s_3},..}\{G_{i+j_1} = g_s\}\ ,$$

where $\sum_{s_2,s_3,\ldots} \eta_{s_2,s_3,\ldots} = 1$.

Therefore,

$$m|P_{h_2,h_3,..}\{H_i = h\} - P_{h_2,h_3,..}\{H_i = h'\}|$$

$$= m.\Big| \sum_{s=0}^{S-1} \sum_{s_2,s_3,\ldots} P_{g_2^{s_2},g_3^{s_3},..}\{G_i = g_s\}\, \eta_{s_2,s_3,\ldots} - \sum_{s=1}^{S} \sum_{s_2,s_3,\ldots} P_{g_2^{s_2},g_3^{s_3},..}\{G_i = g'_s\}\, \eta_{s_2,s_3,\ldots}\Big|$$

$$\leq (1/S) \sum_{s=0}^{S-1} \sum_{s_2,s_3,\ldots} \eta_{s_2,s_3,\ldots}(mS)\Big|P_{g_2^{s_2},g_3^{s_3},..}\{G_i = g_s\} - P_{g_2^{s_2},g_3^{s_3},..}\{G_i = g'_s\}\Big|$$

$$\leq (1/S) \sum_{s=0}^{S-1} \sum_{s_2,s_3,\ldots} \eta_{s_2,s_3,\ldots} K^G(|g_s - g'_s|)$$

$$\leq (1/S) \sum_{s=0}^{S-1} \sum_{s_2,s_3,\ldots} \eta_{s_2,s_3,\ldots} K^G(|h - h'|)$$

$$\leq K^G|h - h'|\ .$$

Therefore,

$$\big|f_{h_2,h_3,..}(h) - f_{h_2,h_3,..}(h')\big| \leq K^G|h - h'|\ \ .\ \blacksquare$$

Now, we want to know what this result means. Then, we use the following corollary.

Corollary 5.6.2 *Let $f^*_{g_2,g_3,..}(g/(mS))$ be the conditional density of $G_i/(mS)$ given $G_{i+j_s} = g_s$ and let $f^*_{h_2,h_3,..}\{h/m\}$ be the conditional density of H_i/m given $H_{i+j_s} = h_s$.*

Let K^H be the Lipschitz coefficient associated to $f_{h_2,h_3,..}$. Let $K_{f_{G/(Sm)}}$ and $K_{f_{H/m}}$ be the Lipschitz coefficients associated to $f^*_{g_2,g_3,..}$ and $f^*_{h_2,h_3,..}$.

Then,

$$K_{f_{H/m}} \leq \frac{K_{f_{G/(Sm)}}}{S} \ .$$

Proof At first, recall that $f_{X/m}(x_0)/m = \int_{x_0-\epsilon}^{x_0+\epsilon} f_{X/m}(u).\mu_m(du) = P\{X/m = x_0\}$ and that $f_X(mx_0)/m = \int_{mx_0-1/2}^{mx_0+1/2} f_X(v).\mu^*_m(dv) = P\{X = mx_0\}$ where $X \in F(m)$ is a random variable, $x_0 = k/m$, $k \in \mathbb{N}$, and ϵ is small.

Then, $f_{X/m}(x_0)/m = P\{X/m = x_0\} = P\{X = mx_0\} = f_X(mx_0)/m$.

Then, $f^*_{g_2,g_3,..}(g/(mS)) = f_{g_2,g_3,..}(g)$ and $f^*_{h_2,h_3,..}(h/m) = f_{h_2,h_3,..}(h)$.

Now,

$$\left|f_{g_2,g_3,..}(g) - f_{g_2,g_3,..}(g')\right| \leq K^G|g - g'|$$
$$\left|f_{h_2,h_3,..}(h) - f_{h_2,h_3,..}(h')\right| \leq K^G|h - h'| \ .$$

Then,

$$\left|f^*_{g_2,g_3,..}(g/(mS)) - f^*_{g_2,g_3,..}(g'/(mS))\right| \leq K^G|g - g'|$$
$$\left|f^*_{h_2,h_3,..}(h/m) - f^*_{h_2,h_3,..}(h'/m)\right| \leq K^G|h - h'| \ .$$

Then,

$$\left|f^*_{g_2,g_3,..}(g/(mS)) - f^*_{g_2,g_3,..}(g'/(mS))\right| \leq (mS)K^G|g - g'|/(mS)$$
$$\left|f^*_{h_2,h_3,..}(h/m) - f^*_{h_2,h_3,..}(h'/m)\right| \leq m.K^H|h - h'|/m \ .$$

Then, $K_{f_{G/(Sm)}} = (mS)K^G$ and $K_{f_{H/m}} = mK^H$.

Therefore, $K_{f_{H/m}}/m = K^H \leq K^G = K_{f_{G/(Sm)}}/(mS)$.

Therefore, $K_{f_{H/m}} \leq K_{f_{G/(Sm)}}/S$. ■

Then, by proposition 4.1.1 ,

$$P\{T(H/m) \in I\} = L(I) + \frac{Ob(1)6.K_{f_{H/m}}}{\sqrt{m}} \approx L(I) + \frac{Ob(1)6.K_{f_{G/(Sm)}}}{\sqrt{m}S} \ ,$$

$$P\{T(G/(Sm)) \in I\} = L(I) + \frac{Ob(1)6.K_{f_{G/(Sm)}}}{\sqrt{Sm}} \ .$$

There also, it seems better to use T with H(j) than with G(j), that is with the XORLT rather than the CLT.

We obtain the same conclusion when we make numerical studies. These results are one of reasons for which we choose to build x(j) by applying T_q to h(j) and not to g(j) : cf section 11.

5.7 Limit theorems for conditional probabilities

We study the case where we add the random variables F(i,j) ($G(j) = \sum_{i=1}^{S} F(i,j)$) where rows F (i,.) are independent : cf section 11.2.4. In that case, the distribution of the sums admitting for probabilities the conditional probabilities is the one of a sum of independent variables. Indeed the following proposition holds.

Proposition 5.7.1 *Let $X_{i,j}$, i=1,...,I, j=1,...,p, be a sequence of random variables. We assume that the rows $X_{i,.} \in F(m)^p$ are independent. For all i=1,2,...,I, let $Q_i^{\{x^*_{i,j}\}}$ be the conditional distribution of $X_{i,1}$ given $X_{i,2} = x^*_{i,2}$, ...,$X_{i,J} = x^*_{i,J}$.*

Then, for all Borel set Bo,

$$P\big\{X_{1,1} + ... + X_{I,1} \in Bo \ \big| \ X_{i,j} = x^*_{i,j}, i = 1, ..., I, j = 2, ..., p\big\}$$

$$= \int_{Bo} \big(x_{1,1} + ... + x_{I,1}\big) Q_1^{\{x^*_{1,j}\}}(dx_{1,j}).....Q_I^{\{x^*_{I,j}\}}(dx_{I,j}) \ .$$

Proof Let f be the probability density function of

$$\big[\big(X_{1,1}, ..., X_{I,1}\big) \ , \ \big(X_{1,2}, ..., X_{I,2}, X_{1,3}, ..., X_{I,3}.....X_{1,p}, ..., X_{I,p}\big)\big]$$

with respect to $\nu \otimes \nu'$ where ν is the distribution of $\big(X_{1,1}, ..., X_{I,1}\big)$ and where ν' is the distribution of $\big(X_{1,2}, ..., X_{I,2}, X_{1,3}, ..., X_{I,3}.....X_{1,p}, ..., X_{I,p}\big)$.

This probability density function exists always because we use random vectors with values in the finite sets $F(m)^p$.

Then, we write it in function of the vectors $\{x_{i,1}\} = \{x_{i,1}\}_{i=1,....,I}$ and $\{x^*_{i,j}\} = \{x^*_{i,j}\}_{i=1,...,I;j=2,...,p}$: $f\big(\{x_{i,1}\}, \{x^*_{i,j}\}\big)$.

Then $\big(X_{1,1}, ..., X_{I,1}\big)$ given $X_{i,j} = x^*_{i,j}, \ i = 1, ..., I, \ j = 2, ..., p$, has a probability density function with respect to ν equal to $f\big(\{x_{i,1}\}, \{x^*_{i,j}\}\big)$.

Now, $\nu = \nu_1 \otimes \nu_2 \otimes ..... \otimes \nu_I$ where the ν_t's are the distributions of $X_{t,1}$, t=1,2,....,I.

Now, $\nu' = \nu'_1 \otimes \nu'_2 \otimes .... \otimes \nu'_I$ where the ν'_t's are the distributions of $\big(X_{t,2}, ..., X_{t,p}\big)$, t=1,2,....,p.

Let g be the probability density function of $\{X_{i,j}\}$, i=1,..,I, j=1,...,p, with respect to the uniform measure $\mu_m \otimes ..... \otimes \mu_m$.

Then, $g(\{x_{i,j}\}) = g_1(\{x_{1,j}\}) .... g_I(\{x_{I,j}\})$ where $g_s(\{x_{i,j}\})$ is the probability density function of $\big(X_{s,1}, ..., X_{s,p}\big)$ with respect to $\mu_m \otimes ..... \otimes \mu_m$.

Let f_{ν_i} and $f'_{\nu'_i}$ be the probability density functions of ν_i and ν'_i with respect to μ_m. Then,
$g(\{x_{i,j}\}) = f\big(\{x_{i,1}\}, \{x^*_{i,j}\}\big) f_{\nu_1}(\{x_{1,1}\}) .... f_{\nu_I}(\{x_{I,1}\}) f'_{\nu'_1}(\{x^*_{1,j}\}) ..... f'_{\nu'_I}(\{x^*_{I,j}\})$.

Therefore,

$$f\big(\{x_{i,1}\}, \{x^*_{i,j}\}\big)$$

$$= \frac{g_1(\{x_{1,j}\}) .... g_I(\{x_{I,j}\})}{f_{\nu_1}(\{x_{1,1}\}) .... f_{\nu_I}(\{x_{I,1}\}) f'_{\nu'_1}(\{x^*_{1,j}\}) ..... f'_{\nu'_I}(\{x^*_{I,j}\})}$$

$$= \frac{g_1(\{x_{1,j}\})}{f_{\nu_1}(\{x_{1,1}\}) . f'_{\nu'_1}(\{x^*_{1,j}\})} ..... \frac{g_I(\{x_{I,j}\})}{f_{\nu_I}(\{x_{I,1}\}) . f'_{\nu'_I}(\{x^*_{I,j}\})} \ .$$

It means that $f\big(\{x_{i,1}\}, \{x^*_{i,j}\}\big) = f_1\big(\{x_{1,1}\}, \{x^*_{1,j}\}\big) ..... f_I\big(\{x_{I,1}\}, \{x^*_{I,j}\}\big)$, where

$$f_i\big(\{x_{i,1}\}, \{x^*_{i,j}\}\big) = \frac{g_i(\{x_{i,j}\})}{f_{\nu_i}(\{x_{i1}\}) . f'_{\nu'_i}(\{x^*_{i,j}\})}$$

is the probability density function of $\big(X_{i,1}, X_{i,2}, ..., X_{i,p}\big)$ with respect to the marginal distributions ν_i and ν'_i.

Therefore,

$$P\big\{X_{1,1} + ... + X_{I,1} \in Bo \ \big| \ X_{i,j} = x^*_{i,j}, i = 1, ..., I, j = 2, ..., p\big\}$$

$$= \int_{Bo} \Big(x_{1,1} + ... + x_{I,1}\Big) f_1\Big(\{x_{1,1}\}, \{x^*_{1,j}\}\Big) ... f_I\Big(\{x_{I,1}\}, \{x^*_{I,j}\}\Big) \nu_1(\{dx_{1,1}\}) ..... \nu_I(\{x_{I,1}\})$$

$$= \int_{Bo} \Big(x_{1,1} + ... + x_{I,1}\Big) \Big[f_1\Big(\{x_{1,1}\}, \{x^*_{1,j}\}\Big) \nu_1(\{dx_{1,1}\})\Big] ..... \Big[f_I\Big(\{x_{I,1}\}, \{x^*_{I,j}\}\Big) \nu_I(\{x_{I,1}\}\Big]$$

$$= \int_{Bo} \Big(x_{1,1} + ... + x_{I,1} \Big) Q_1^{\{x_{1,j}^*\}}(dx_{1,j}) ..... Q_I^{\{x_{I,j}^*\}}(dx_{I,j}) \ . \ \blacksquare$$

Then, we study the conditional probabilities of the sums given the value of the sums. That is we shall use the following lemma

Lemma 5.7.1 *Let* $Y_j = X_{1,j} + ... + X_{I,j}$. *We set*

$$\{Y_j = y_j\} = \cup_{x_{1,j}+...+x_{I,j}=y_j} \Big\{ \{X_{1,j} = x_{1,j}\} \cap .... \cap \{X_{I,j} = x_{I,j}\} \Big\} \ .$$

Then,

$$P\{Y_1 \in Bo \mid Y_2 = y_2, ..., Y_p = y_p\}$$

$$= \sum_{x_{i,j}:\forall j, \ x_{1,j}+...+x_{I,j}=y_j} \eta_{\{x_{i,j}\}} P\big\{ Y_1 \in Bo \ \big| \ X_{i,j} = x_{i,j}, i = 1, ..., I, j = 2, ..., p \big\} \ ,$$

where

$$\eta_{\{x_{i,j}\}} = \frac{P\Big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \Big\}}{\sum_{x_{i,j}:\forall j, \ x_{1,j}+...+x_{I,j}=y_j} P\Big\{ \big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \big\} \Big\}} \ .$$

Proof We have

$$P\{Y_1 \in Bo \mid Y_2 = y_2, ..., Y_p = y_p\}$$

$$= \frac{P\big\{ (Y_1 \in Bo) \cap \big(Y_2 = y_2 \big) \cap ... \cap \big(Y_p = y_p \big) \big\}}{P\big\{ \big(Y_2 = y_2 \big) \cap ... \cap \big(Y_p = y_p \big) \big\}}$$

$$= \frac{P\Big\{ \{Y_1 \in Bo\} \cap \Big[\cup_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} \Big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \Big\} \Big] \Big\}}{P\Big\{ \cup_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} \Big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \Big\} \Big\}}$$

$$= \frac{P\Big\{ \cup_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} \{Y_1 \in Bo\} \cap \Big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \Big\} \Big\}}{P\Big\{ \cup_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} \Big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \Big\} \Big\}}$$

$$= \frac{\sum_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} P\Big\{ \{Y_1 \in Bo\} \cap \Big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \Big\} \Big\}}{\sum_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} P\Big\{ \big\{ \cap_i \{X_{i,2} = x_{i,2}\} ... \cap_i \{X_{i,p} = x_{i,p}\} \big\} \Big\}}$$

$$= \sum_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} \eta_{\{x_{i,j}\}} \frac{P\Big\{\{Y_1 \in Bo\} \cap \Big\{ \cap_i \{X_{i,2} = x_{i,2}\}... \cap_i \{X_{i,p} = x_{i,p}\}\Big\}\Big\}}{P\Big\{ \cap_i \{X_{i,2} = x_{i,2}\}... \cap_i \{X_{i,p} = x_{i,p}\}\Big\}} \ .$$

Now,

$$\frac{P\Big\{\{Y_1 \in Bo\} \cap \Big\{ \cap_i \{X_{i,2} = x_{i,2}\}... \cap_i \{X_{i,p} = x_{i,p}\}\Big\}\Big\}}{P\Big\{ \cap_i \{X_{i,2} = x_{i,2}\}... \cap_i \{X_{i,p} = x_{i,p}\}\Big\}}$$

$$= P\big\{Y_1 \in Bo \ \big| \ X_{i,j} = x_{i,j}, i = 1, ..., I, j = 2, ..., p\big\} \ . \ \blacksquare$$

Remark that $\sum_{x_{i,j}:\forall j, x_{1,j}+...+x_{I,j}=y_j} \eta_{\{x_{i,j}\}} = 1$.

Then, the distribution of conditional probabilities

$$P\{Y_1 \in Bo \mid Y_2 = y_2, ..., Y_p = y_p\}$$

is always the one of a sum of independent variables. It involves a faster convergence.

The same results holds with the XORLT.

Proposition 5.7.2 *We keep the notations of lemma 5.7.1. Then,*

$$P\{\overline{Y_1} \in Bo \mid \overline{Y_2} = y_2, ..., \overline{Y_p} = y_p\}$$

$$= \sum_{x_{i,j}:\forall j, \overline{x_{1,j}+...+x_{I,j}}=y_j} \eta'_{\{x_{i,j}\}} P\big\{\overline{Y_1} \in Bo \ \big| \ X_{i,j} = x_{i,j}, i = 1, ..., I, j = 2, ..., p\big\} \ ,$$

where

$$\eta'_{\{x_{i,j}\}} = \frac{P\Big\{ \cap_i \{X_{i,2} = x_{i,2}\}... \cap_i \{X_{i,p} = x_{i,p}\}\Big\}}{\sum_{x_{i,j}:\forall j, \overline{x_{1,j}+...+x_{I,j}}=y_j} P\Big\{\big\{ \cap_i \{X_{i,2} = x_{i,2}\}... \cap_i \{X_{i,p} = x_{i,p}\}\big\}\Big\}} \ .$$

The proof is the same that that of the lemma 5.7.1 with sums modulo 1.

Then, these results show that, in many cases,

$$P\{\overline{Y_1} \in I \mid \overline{Y_2} = y_2, ..., \overline{Y_p} = y_p\} \to L(I) \; as \; n \to \infty \; .$$

Results obtained in section 5.3.3 show that this limit is checked for all the data used to build the random sequences $b^1(n')$.

Another way of understanding matters it is to apply the proposition 5.5.2 : with a probability bigger than $1 - 2\Gamma(b)$,

$$P\Big\{\{\overline{Y_1} = y_1\} \cap .... \cap \{\overline{Y_p} = y_p\}\Big\} \approx 1/N^p \Big[1 + \frac{Ob(1).b\sigma_{P'}}{E_{P'}\sqrt{N^{p(S-1)}}}\Big] \; .$$

It means that

$$P\{\overline{Y_1} = y_1 \mid \overline{Y_2} = y_2, ..., \overline{Y_p} = y_p\} \approx P\{\overline{Y_1} = y_1\} \; ,$$

if S is big enough.

Indeed,

$$P\{\overline{Y_1} \in I \mid \overline{Y_2} = y_2, ..., \overline{Y_p} = y_p\}$$

$$\approx \frac{P\{\overline{Y_1} \in I\} P\Big\{\{\overline{Y_2} = y_2\} \cap .... \cap \{\overline{Y_p} = y_p\}\Big\}}{P\Big\{\{\overline{Y_2} = y_2\} \cap .... \cap \{\overline{Y_p} = y_p\}\Big\}}$$

$$\approx P\{\overline{Y_1} \in I\} \approx L(I) \; .$$

Then, under the hypotheses of our data, we can admit the following hypothesis.

Hypothesis 5.7.1 *In the space of probabilities with the measure defined in proposition 5.5.2, the following approximations hold with a probability approximately bigger than* $1 - 2\Gamma(b)$*,*

$$P\{\overline{Y_1} = y_1 \mid \overline{Y_2} = y_2, ..., \overline{Y_p} = y_p\} \approx P\{\overline{Y_1} \in I\} \approx L(I) \; .$$

Chapter 6

Dependence induced by linear congruences

We study in this chapter the dependence induced by $\big(T^n(x_0), T^{n+1}(x_0)\big)$.

Notations 6.0.1 *In this section one denotes by T a congruence $T(x) \equiv ax + c$ mod(m) where $0 < a < m$ and where a, m et c are fixed.*

6.1 Theoretical study

This study will enable us to note that in fact congruences of Fibonacci induce the weakest dependences.

6.1.1 Notations

To study dependence amounts studying the distribution of the points $(\ell, \overline{T(\ell)})$. The following notation will thus be used.

Notations 6.1.1 *We set $E_2 = \big\{\ell, \overline{T(\ell)} \mid \ell \in \{0, 1, ..., m-1\}\big\}$.*

We will see that this dependence depends on the continued fraction $\frac{m}{a}$, i.e. it depends on sequences r_n and h_n defined in the following way.

Notations 6.1.2 *Let $r_0 = m$, $r_1 = a$. One denotes by r_n the sequence defined by $r_n = h_{n+1}r_{n+1} + r_{n+2}$ the Euclidean division of r_n by r_{n+1} when $r_{n+1} \neq 0$.*

One denotes by d the smallest integer such as $r_{d+1} = 0$. One sets $r_{d+2} = 0$ and $h_{d+1} = \infty$.

Therefore, $h_n \geq 1$ for all n=1,2,...,d and $r_{d-1} = h_d r_d + r_{d+1} = h_d r_d + 0 = h_d r_d$, $r_d = h_{d+1}r_{d+1} + r_{d+2} = 0 * h_{d+1} + 0 = 0 * \infty$.

The full sequence r_n is thus the sequence $r_0 = m$, $r_1 = a$, ..., $r_{d+1} = 0$, $r_{d+2} = 0$. Then, it is easy to prove the following result.

Proposition 6.1.1 *The congruence $T(x) \equiv a$ mod(m) is a Fibonacci conguence if $h_n = 1$ for n=1,2,...d-1, $h_d = 2$ and $r_d = 1$*

In this case, r_n is the Fibonacci sequence fi_n, except for the last terms. If one wants to be able to have all the Fibonacci sequence, one can adopt the following notations.

Notations 6.1.3 *One sets $r'_0 = m$, $r'_1 = a$. Let $r'_n = h'_{n+1}r'_{n+1} + r'_{n+2}$ the Euclidean division of r'_n by r'_{n+1} except if there exists d' such that $r'_{d'-2} = 3$ and $r'_{d'-1} = 2$. In this case, one sets $r'_{d'} = r'_{d'+1} = 1$ and $h'_{d'} = 1$.*

In addition, one considers also the following sequences.

Notations 6.1.4 *One sets $k_0 = 0$, $k_1 = 1$ and $k_{n+2} = h_{n+1}k_{n+1} + k_n$ if $n + 1 \leq d$. Lastly, one sets $k_{d+2} = \infty$.*

Then, we have the sequence $k_0 = 0$, $k_1 = 1$, ..., k_{d+1}, $k_{d+2} = \infty$.

Remark that if $h_n = 1$ for n=1,2,...,d-1, k_n is also the Fibonacci sequence for n=1,2,...,d.

6.1.2 Theorems

One will understand that dependence depends on the h_i: more they are small, more the dependence is weak. As $h_i \geq 1$, the best congruence will satisfy $h'_i = 1$. It will be thus the congruence of Fibonacci.

First Theorem

To understand that, first one considers the rectangles $[0, k_n] \otimes [0, r_{n-2}]$. Thus, if n is even, the set $\left\{(k_{n-1}\ell, \overline{k_{n-1}\ell})\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}$ is concentrated on the line $x \mapsto (r_{n-1}/k_{n-1})x$. More generally, we have the following theorem (the proof is in Proof 6.2.9).

Theorem 7 *One supposes c=0. Let $n \in \{2, 3...., d\}$. Then*

If n is even , $E_2 \cap \left\{[0, k_n[\otimes [0, r_{n-2}[\right\} = \left\{(k_{n-1}\ell\ ,\ r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}$. Moreover the points $(k_{n-1}\ell\ ,\ r_{n-1}\ell)$ are lined up.

If n is odd,
$E_2 \cap \left\{]0, k_n] \otimes]0, r_{n-2}]\right\} = \left\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}$. *Moreover, the points $(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)$ are lined up.*

That means that the rectangle $\left\{[0, k_n/2] \otimes [r_{n-2}/2, r_{n-2}[\right.$ does not contain points of E_2 : $E_2 \cap \left\{[0, k_n/2] \otimes [r_{n-2}/2, r_{n-2}[\right\} = \emptyset$.

If h_{n-1} is large, that will mean that an important rectangle of $\mathbb{R}^2$ is empty of points of E_2: that will mark a breakdown of independence.

For example suppose n=2. First, $m = r_0$, $r_1 = a$, $k_1 = 1$ and $k_2 = h_1 = \lfloor m/a \rfloor$ ou $\lfloor x \rfloor$ means the integer part of x. In this case, one regards the rectangles $[0, k_2[\otimes [0, m[$. One thus finds a traditional technique for n=2. Indeed when one makes chi-squared tests, one can use rectangles of type $[0, m/(2a)[\otimes [0, m/2[$.

Thus if "a" is not large enough compared to "m", there is rupture of independence. The rectangle $Rect_2 = [0, m/(2a)] \otimes [m/2, m[$ will not contain any point of E_2. However, this rectangle has its surface equal to $m^2/(4a)$: if the points of E_2 are distributed in a uniform way, one has about $m/(2a) = h_1/2 + r_2/(2a)$ points of E_2 (and not 0). Thus if "a" is not sufficiently large, i.e if h_1 is too large, there is breakdown of independence.

For example, choose the congruence $T(x) = 10^3 x$ modulo $10^6 - 1$: $Rect_1 = [0, m/(2a)] \otimes [0, m/2[$ contains 500 points of E_2 roughly and $Rect_2$ contains 0 points. Neither the chi-squared test that one could make on such rectangles nor the definitions 2.1.5 or 2.1.6 are satisfied.

Now choose m=99, a=5, $k_2 = 19$: cf figure 6.1: $Rect_1$ roughly contains 10 points of E_2 for a total sample of 99.

Choose m=99, a=10, $k_2 = 9$: cf figure 6.2: $Rect_2$ roughly contains 5 points of E_2 : the breakdown is less clear.

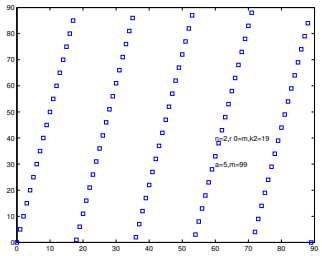

Figure 6.1: Points in rectangles a=5

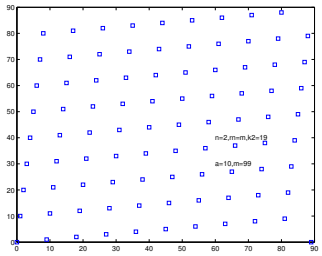

Figure 6.2: Points in rectangles a= 10

In these examples, we studied rectangles $[0, k_2/2] \otimes [0, m/2[\big\}$. In the general case, it is necessary that $[0, k_n/2] \otimes [r_{n-2}/2, r_{n-2}[\big\}$ has not a too big size. It is necessary thus that all the h_{n-1} are small.

Now, one can extend this result to other rectangles of E_2.

Corollary 6.1.2 *Let $(x_0, y_0) \in E_2$. Let $R^0 = \big\{[x_0, x_0 + k_n] \otimes [y_0, y_0 + r_{n-2}[\big\}$ and let $R_0 = \overline{R^0}$, be the rectangle R^0 modulo m. Then*

If n is even, $E_2 \cap R_0 = \big\{(\overline{x_0 + k_{n-1}\ell}\ ,\ \overline{y_0 + r_{n-1}\ell})\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$. Moreover the points $(\overline{x_0 + k_{n-1}\ell}\ ,\ \overline{y_0 + r_{n-1}\ell})$ are lined up modulo m .

If n is odd,
$E_2 \cap R_0 = \big\{(\overline{x_0 + k_{n-2} + k_{n-1}\ell}\ ,\ \overline{y_0 + r_{n-2} - r_{n-1}\ell})\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$. *Moreover, the points $(\overline{x_0 + k_{n-2} + k_{n-1}\ell}\ ,\ \overline{y_0 + r_{n-2} - r_{n-1}\ell})$ are lined up modulo m.*

Remark 6.1.1 *In general, it is only on the border that the rectangle modulo m R_0 satisfies $R_0 \neq R^0$. If not, R_0 is a normal rectangle.*

This result shows this fact : so that there is not breakdown of independence it is necessary that the h_i are small.

Principal theorem

Now, one will be interested in the number of points included in rectangles other than $[0, k_n] \otimes [0, r_{n-2}[$.

Notations 6.1.5 *Let $x_n \in \mathbb{R}^q$ be an unspecified sample. Let Bo be a Borel set of R^q. One denotes by $N(Bo)$ the number of x_n which belongs to Bo.*

Here one takes in account the number of points of E_2 contained in rectangles of the type $R_{ect} = [x, x + L[\otimes[y, y + L'[$.

To know the behavior of $N(R_{ect})$, one uses the following theorem which is based on the proof given in [13] page 47-131.

Theorem 8 *It is supposed that T is invertible. Let R_{ect} be a rectangle of $F^*(m)^2$, length $L_{on} \geq 1$, width $L_{ar} \geq 1$. Let $N(R_{ect})$ be the number of points of E_2 which belong to R_{ect} and let $S_{R_{ect}}$ be its surface. One denotes by Log the Neperian logarithm : Log(e)=1. Then,*

$$\big|N(R_{ect}) - \frac{S_{R_{ect}}}{m}\big| \leq (p^o + 1)(sup(h_i) + 1)\ ,$$

where p^o is a function of (L_{on}, L_{ar}) satisfying

$$2.0782.Log(m_{in}) + 2.00005 \geq p^o\ ,$$

where $m_{in} = Min(L_{on}, L_{ar})$.

Proof According the definition of lemma 7-6 page 115 of [13], let p' be the number of $\lambda_i \neq 0$ in the factorization :

$$L = L_{on} = \sum_{i=1}^{T} \lambda_i k_{i+c} + \epsilon \ ,$$

where $\lambda_i \in \mathbb{N}$, $\lambda_i \leq h_{c+i}$, $\lambda_1 < h_{c+1}$, $\lambda_{i-1} = 0$ if $\lambda_i = h_{c+i}$ and $-k_c \leq \epsilon < k_{c+1}$. We suppose $c + T \leq d + 1$.

Write differently this factorization of L : we suppress the writing of the λ_i's such that $\lambda_i = 0$, i.e. we write $L = \sum_{i=1}^{p'} \lambda'_i k'_{c+i} + \epsilon$, where $\lambda'_i = \lambda_{i+d(i)} > 0$, $k'_{c+i} = k_{c+d(i)+i}$, $d(i) \geq 0$, $d(i)$ increasing.

Now, remark that the sequence k_n used in the definitions 1-3 page 49 of [13] is the sequence k_n defined in this report in notations 6.1.4. Then, $k'_n \geq k_n$.

Let fi_n be the Fibonacci sequence. Then $k_n \geq fi_n$. Assume that $p' \geq 2$.

Because $\epsilon \geq -k_c$, $L_{on} = \sum_{i=1}^{T} \lambda_i k_{c+i} + \epsilon = \sum_{i=1}^{p'} \lambda'_i k'_{c+i} - k_c = \lambda'_{p'} k_{c+p'+d(p')} + \sum_{i=1}^{p'-1} \lambda'_i k_{c+i+d(i)} - k_c \geq \lambda'_{p'} k_{c+p'+d(p')} \geq fi_{c+p'+d(p')} \geq fi_{p'}$.

Assume $p' = 1$. Then, $L_{on} \geq 1 = fi_1 = fi_{p'}$.

Now let p" be the number of $\mu_i \neq 0$ in the factorization :

$$L_{ar} = \sum_{i=1}^{T} \mu_i r_{c'-i} + \epsilon_2 \ ,$$

where $\mu_i \in \mathbb{N}$, $\mu_i \leq h_{c'-i}$, $\mu_1 < h_{c'-1}$, $\mu_{i-1} = 0$ if $\mu_i = h_{c'-i}$ and $-r_{c'} \leq \epsilon_2 < r_{c'-1}$ [1].

Let $p^0 = Min(p', p")$. Then, $L_{on} \geq fi_{p'} \geq fi_{p^0}$. By the same way, $L_{ar} \geq fi_{p^0}$. Then, $min \geq fi_{p^0}$.

By lemma 7-6 page 115 of [13], we know that

$$\big|N(R_{ect}) - \frac{S_{R_{ect}}}{m}\big| \leq (p^o + 1)(sup(h_i) + 1) \ .$$

Now, it is known that

$$fi_n = \frac{1}{\sqrt{5}}\Big[\Big(\frac{1+\sqrt{5}}{2}\Big)^n - \Big(\frac{1-\sqrt{5}}{2}\Big)^n\Big] \ .$$

[1] The correct definition is page 112, notations 7-2 of [13]. But it is not necessary to use it : it is enough that $L_{ar} \geq fi_{p"}$.

Because $m_{in} \geq fi_{p^o}$,

$$Log(m_{in}) + Log(\sqrt{5}) - Log\Big(1 - \Big(\frac{1-\sqrt{5}}{1+\sqrt{5}}\Big)^{p^o}\Big) \geq p^o.Log\Big(\frac{1+\sqrt{5}}{2}\Big) \ .$$

Therefore, because $1 - \sqrt{5} < 0$, the weakest increase holds for $p^0 = 2$:

$$Log(m_{in}) + 0.80472 + 0.15771 \geq 0.4812.p^o \ .$$

Therefore,

$$2.0782.Log(m_{in}) + 2.00005 \geq p^o \ . \ \blacksquare$$

Remark 6.1.2 *In [13] page 48, we define r_n by the following way : $r_0 = m$, $r_1 = Min(a, m-a)$ and $r_n = h_{n+1}r_{n+1} + r_{n+2}$ the Euclidean division of r_n by r_{n+1}. It is a very small difference : results are identical.*

We deduce the following corollary.

Corollary 6.1.3 *Let $P_e(R_{ect}) = N(R_{ect})/m$ be the empirical probability of R_{ect}/m associated to the sample E_2/m and let $P_U(R_{ect}) = L(R_{ect}/m)$ be the uniform probability of R_{ect}/m in $[0,1]^2$. Then, in $F(m)^2$,*

$$\big|P_e(R_{ect}) - P_U(R_{ect})\big| \leq \frac{(2.0782.Log(m_{in}) + 3.00005)(sup(h_i)+1)}{m} \ . \quad (6.1)$$

Therefore, to the maximum, $p^o = d$, $m_{in} = m$ and

$$\big|P_e(R_{ect}) - P_U(R_{ect})\big| \leq \frac{(2.0782.Log(m) + 3.00005)(sup(h_i)+1)}{m} \ .$$

Now let us consider the approximately normal distribution associated with

$$\frac{\sqrt{m}\big|P_e(R_{ect}) - P_U(R_{ect})\big|}{\sigma_{R_{ect}}} \ ,$$

where $\sigma^2_{R_{ect}}$ is the variance associated to $\mathbf{1}_{R_{ect}/m}(U_n)$ when U_n is an IID sample with uniform distribution in $[0,1]^2$: $\sigma^2_{R_{ect}} = L(R_{ect}/m)[1 - L(R_{ect}/m)]$.

Under the IID assumption, one checks with a probability of 99 percent:

$$\frac{\sqrt{m}\big|P_e(R_{ect}) - P_U(R_{ect})\big|}{\sigma_{R_{ect}}} \leq 2.57 \ .$$

Proposition 6.1.4 *The test of independence with levels of significance equal to 1 percent and which has the statistics*

$$\frac{\sqrt{m}\big|P_e(R_{ect}) - P_U(R_{ect})\big|}{\sqrt{L(R_{ect}/m)[1 - L(R_{ect}/m)]}} \leq 2.57$$

is checked by the sample E_2 except for some rectangles R_{ect}/m, for example those such that

$$\frac{0.6539.Log(m)^2(sup(h_i)+1)^2}{m} \geq L(R_{ect}/m) \ .$$

Proof. Indeed the test is checked if

$$\frac{\sqrt{m}\big|P_e(R_{ect}) - P_U(R_{ect})\big|}{2.57\sqrt{1 - L(R_{ect}/m)}} \leq \sqrt{L(R_{ect}/m)} \ .$$

However,

$$\big|P_e(R_{ect}) - P_U(R_{ect})\big| \leq \frac{(2.0782.Log(m) + 3.00005)(sup(h_i)+1)}{m} \ .$$

Therefore, the test is checked as soon as

$$\frac{(2.0782.Log(m) + 3.00005)(sup(h_i)+1)}{2.57\sqrt{1 - L(R_{ect}/m)}\sqrt{m}} \leq \sqrt{L(R_{ect}/m)} \ .$$

That could be not the case if

$$\frac{(2.0782.Log(m) + 3.00005)(sup(h_i)+1)}{2.57\sqrt{1 - L(R_{ect}/m)}\sqrt{m}} \geq \sqrt{L(R_{ect}/m)} \ .$$

In particular, it is true if

$$\frac{(2.0782.Log(m) + 3.00005)(sup(h_i)+1)}{2.57\sqrt{m}} \geq \sqrt{L(R_{ect})/m} \ .$$

In particular, it is true if

$$\frac{2.0782.Log(m)(sup(h_i)+1)}{2.57\sqrt{m}} \geq \sqrt{L(R_{ect}/m)} \ ,$$

that is to say,

$$\frac{0.65389.Log(m)^2(sup(h_i)+1)^2}{m} \geq L(R_{ect}/m) \ . \ \blacksquare$$

Study of the principal theorem

The previous results show that if there exist i_0 such that h_{i_0} is large, one will have a breakdown of independence. That, one knew it by corollary 6.1.2.

It remains to be understood that the conditions of independence are well carried out if the h_i are small.

Then, suppose that the h_i's are small.

The proposition 6.1.4 means that it is necessary that $L(R_{ect}/m)$ is small so that the tests of independence are not checked for some rectangles R_{ect}.

But in this case, the increase of the proposition 6.1.4 is too weak. Indeed, $m_{in} \leq \sqrt{m^2 L(R_{ect}/m)}$. One can thus improve it. For that, let us take again the increase of the equation 6.1 :

$$\big|P_e(R_{ect}) - P_U(R_{ect})\big|$$

$$\leq \frac{(2.0782.Log(m) + (1/2)Log\big(L(R_{ect}/m)\big) + 3.00005(sup(h_i) + 1)}{m} \ ,$$

where $Log\big(L(R_{ect}/m)\big) < 0$.

To improve this result, one takes again previous calculation with the new parameters and one obtains a new equation of the type of that of proposition 6.1.4. But for the same reasons, this result will have to still be improved. One will act again in the same way, and so on as long as the increase of $L(R_{ect}/m)$ can be improved. One thus obtains rectangles R_{ect}/m of size $L(R_{ect}/m)$ increasingly small.

Moreover, if h_i is small, the results of [13] page 47-13 can be improved : for example for some rectangle $R_{ect} = R^0$, theorem 6-14, page 106-107 :

$$\big|N(R^0) - \frac{S_{R^0}}{m}\big| \leq 1 \ .$$

Then one can again improve the increases of theorem 8 :

$$\big|N(R_{ect}) - \frac{S_{R_{ect}}}{m}\big| \leq (p^o + 1)(sup(h_i) + 1) \ ,$$

$$2.0782.Log(m_{in}) + 2.00005 \geq p^o \ .$$

Finally, if h_i is small, the only rectangles where there is maybe breakdown of independence are the rectangles of the type $R^0 = [x, x + k_n[\otimes[y, y + r_{n-2}[$. Besides, these rectangles do not contain enough points to make tests if h_i is small: if $y = \overline{T}(x)$ the breakdown with independence is proved by corollary 6.1.2 : there is $h_{n-1} + 1$ lined up points. If h_i is small, it is easy to understand that it is not important.

Thus, in the case of the Fibonacci sequence, there is at the most 2 points lined up in these rectangles! It is thus a correct result if there is independence.

Now, if $y \neq \overline{T}(x)$, by the theorem 6-14, page 106 of [13], $N(R^0) = k_n r_{n-2}/m + e$ where $|e| \leq 1$. In the case $x = \overline{T}(y)$, $N(R^0) = h_{n-1} + 1$. Therefore, $k_n r_{n-2}/m = h_{n-1}+1+e'$ where $|e'| \leq 1$. In the case $y = \overline{T}(x-1)$, it is easy to understand that generally $N(R^0) = h_{n-1}$. Therefore, $k_n r_{n-2}/m = h_{n-1}+e"$ where $|e"| \leq 1$. Therefore, $h_{n-1} \leq k_n r_{n-2}/m \leq h_{n-1}+1$. Then, $N(R^0) = h_{n-1}$ or $N(R^0) = h_{n-1} + 1$.

Thus, in the case of the Fibonacci sequence, there is 1 or 2 points in these rectangles! It is also thus a correct result if there is independence.

Thus in the case of the Fibonacci sequence, all rectangles satisfy the test of normality. In fact, it is even statistically too. It is not important. We do not make use of it like sample of independent couple.

Numerical examples

We confirm by graphs that it is necessary and sufficient that $sup(h_i)$ is small so that there is independence.

In these graph, one takes m=21.

If a = 13, we have a Fibonacci congruence. The points are well distributed in the square : cf figure 6.3.

In the figures 6.4 et 6.5 breakdowns of independence appear.

If one chooses a=10, $sup(h_i) = 20$: cf figure 6.4 .

If one chooses a=5, $sup(h_i) = 5$: cf figure 6.5.

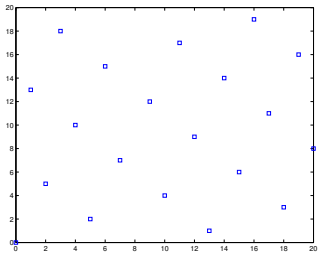

Figure 6.3: $sup(h_i) = 1$

Conclusion

To avoid any dependence, it is necessary that $sup(h_i)$ is small. In the case of the Fibonacci sequence, $sup(h'_i) = 1$ and independence is checked on all rectangles R_{ect}.

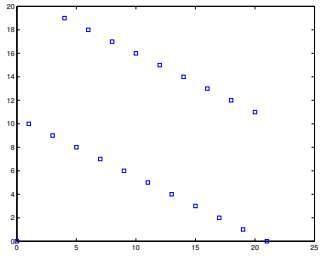

Figure 6.4: $sup(h_i) = 20$

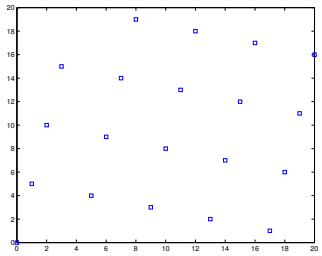

Figure 6.5: $sup(h_i) = 5$ Fig

Remark We studied independence only between two successive elements. It does not mean there is independence for the triplets $(\ell, T(\ell), T^2(\ell))$.

Thus, for the Fibonacci congruence $T^2 = \pm Id$ where Id is the identity [2]. Therefore there is main dependence between the elements $\ell_n = \overline{T^n}(x_0)$ and $\ell_{n+2} = \pm\ell_n$ modulo m. One cannot thus apply it to create a pseudo-random sequence, but only to make possible that a number is independent of another.

It is what we do in section 8.1. We use only independence between two successive elements.

6.2 Proof of theorem 7

In this section, congruences are conguences modulo m : $t \equiv s$ means $t \equiv s$ modulo m. In order to prove theorem 7, few lemmas are needed. The first one is obvious.

[2] With the notations of [13] page 99, $k_d a \equiv \pm 1$ (lemma 6-3). Moreover, if T is the Fibonacci congruence, by lemma 6-5 of [13], $m = h_d k_d + k_{d-1} = 2k_d + k_{d-1} = 2fi_d + fi_{d-1} = fi_{d+2}$ and $m = a + r_2 = r_1 + r_2 = fi_{d+1} + fi_d$. Then, $k_d = r_2$. Then, $m = r_1 + r_2 = a \pm \overline{a^{-1}}$. Then, $\overline{a^{-1}} = a$ or $\overline{a^{-1}} = m - a$

Lemma 6.2.1 *For n=3,4,...,d+1,* $k_{n+1} > k_n > k_{n-1}$. *Moreover* $k_{n+2} = h_{n+1}k_{n+1} + k_n$ *is the Euclidean division of* k_{n+2} *by* k_{n+1}.

Now, we prove the following results.

Lemma 6.2.2 *Let n=0,1,2,...,d. If n is even,* $\overline{k_n a} = m - r_n$. *If n is odd,* $\overline{k_n a} = r_n$.

Proof : We prove this lemma by recurrence.
For n=0, $\overline{k_n a} = \overline{0} = 0 = m - m = m - r_0$. For n=1, $\overline{k_n a} = \overline{a} = a = r_1$.

We suppose that it is true for n.
One supposes n even. Then, $k_{n+1}a \equiv ah_n k_n + ak_{n-1} \equiv -h_n r_n + r_{n-1} = r_{n+1}$.
One supposes n odd. Then, $k_{n+1}a \equiv ah_n k_n + ak_{n-1} \equiv h_n r_n - r_{n-1} = -r_{n+1} \equiv m - r_{n+1}$. Therefore, $\overline{k_{n+1}a} = m - r_{n+1}$. ■

Lemma 6.2.3 *Let n=2,3,...,d+1. Let* $t \in \{1, 2, ..., k_n - 1\}$. *If* $n \geq 2$ *is even,* $r_{n-1} \leq \overline{at} < m - r_n$. *If* $n \geq 3$ *is odd,* $m - r_{n-1} \geq \overline{at} > r_n$.
Moreover, if $n \geq 2$ *is even,* $\overline{k_n a} = m - r_n$. *If* $n \geq 3$ *is odd,* $\overline{k_n a} = r_n$.

Proof : The second assertion is lemma 6.2.2. Now, we prove the first assertion by recurrence.
One supposes n=2. Then, $m = r_0 = h_1 r_1 + r_2 = h_1 a + r_2$. Moreover, $k_2 = h_1$. If $1 \leq t < h_1 = k_2$, $r_1 = a \leq at < h_1 a = m - r_2$.

One supposes that the first assertion is true for n where $2 \leq n \leq d$.

Let $0 < t' < k_{n+1}$. Let $t' = fk_n + e$ be the Euclidean division of t' by k_n : $e < k_n$.
Then, $f \leq h_n$. If not, $t' \geq (h_n + 1)k_n + e \geq h_n k_n + k_{n-1} = k_{n+1}$.

One supposes n even.
In this case, $r_{n-1} \leq \overline{at} < m - r_n$ for $t \in \{1, 2, ..., k_n - 1\}$.
Moreover, $at' \equiv fak_n + ae \equiv f(m - r_n) + ae \equiv -fr_n + ae$.

First, one supposes $e = 0$. Then, $f \geq 1$.
Moreover, because $n \geq 2$, $m - r_n \geq m - fr_n \geq m - h_n r_n = m - (r_{n-1} - r_{n+1}) = r_0 - r_{n-1} + r_{n+1} \geq r_0 - r_1 + r_{n+1} > r_{n+1}$.
Therefore, because $at' \equiv -fr_n$, $\overline{at'} = m - fr_n$.
Therefore, $m - r_n \geq \overline{at'} > r_{n+1}$.

Now, one supposes $f < h_n$ and $e > 0$.
By recurrence, $m - r_n \geq \overline{ae} \geq \overline{ae} - fr_n \geq r_{n-1} - fr_n \geq r_{n-1} - (h_n - 1)r_n =$

$r_n + r_{n+1} > r_{n+1}$.
Therefore, because $at' \equiv -fr_n + ae$, $\overline{at'} = \overline{ae} - fr_n$.
Therefore, $m - r_n \geq \overline{at'} > r_{n+1}$.

One supposes $f = h_n$, $e \neq k_{n-1}$ and $e > 0$.
If $e \neq k_{n-1}$, $\overline{ae} \neq \overline{k_{n-1}a}$. Indeed, if not, $\overline{a(e - k_{n-1})} = 0$. For example, if $e - k_{n-1} > 0$, $k_n > e - k_{n-1} > 0$. Then, because our recurence, $\overline{a(e - k_{n-1})} > r_{n-1} > 0$: it is impossible.

Now, if $n = 2$, $\overline{k_{n-1}a} = \overline{k_1 a} = \overline{a} = r_1 = r_{n-1}$.
Moreover, if $n > 2$, $n \geq 4$. Then, by recurence $\overline{k_{n-1}a} = r_{n-1}$.
Then, if $e \neq k_{n-1}$, $\overline{ae} \neq \overline{k_{n-1}a} = r_{n-1}$. Then, $\overline{ae} > r_{n-1}$.

Moreover, $m - r_n \geq \overline{ae} \geq \overline{ae} - fr_n > r_{n-1} - fr_n \geq r_{n-1} - h_n r_n = r_{n+1}$.
Therefore, because $at' \equiv -fr_n + ae$, $\overline{at'} = \overline{ae} - fr_n$.
Therefore, $m - r_n \geq \overline{at'} > r_{n+1}$.

One supposes $f = h_n$ and $e = k_{n-1}$. Then, $t' = h_n k_n + k_{n-1} = k_{n+1}$. It is oppositite to the assumption.

Then, in all the cases, for $t' \in \{1, 2, ..., k_{n+1} - 1\}$, $m - r_n \geq \overline{at'} > r_{n+1}$.
Therefore, the lemma is true for n+1 if n is even. Then, it is also true for n+1=3.

One supposes n odd with $n \geq 3$. In this case, $r_n < \overline{at} \leq m - r_{n-1}$ for $t \in \{1, 2, ..., k_n - 1\}$.
Moreover, $\overline{ak_n} = r_n$. Therefore, $at' \equiv fak_n + ae \equiv fr_n + ae$.

Assume $e = 0$. Then, $f \geq 1$.
Then, $r_n \leq fr_n \leq h_n r_n = r_{n-1} - r_{n+1} < m - r_{n+1}$.
Then, because $at' \equiv fr_n$, $r_n \leq \overline{at'} = fr_n < m - r_{n+1}$.

Assume $e > 0$ and $f \leq h_n - 1$.
By recurrence, $r_n < \overline{ae} + fr_n \leq m - r_{n-1} + fr_n \leq m - r_{n-1} + (h_n - 1)r_n = m - (r_{n-1} - h_n r_n) - r_n = m - r_{n+1} - r_n < m - r_{n+1}$.
Then, because $at' \equiv fr_n + ae$, $r_n < \overline{at'} = \overline{ae} + fr_n < m - r_{n+1}$

Assume $e > 0$, $e \neq k_{n-1}$ and $f = h_n$.
Because, $e \neq k_{n-1}$, $\overline{ae} \neq m - r_{n-1}$. If not, $\overline{ae} = \overline{ak_{n-1}} = m - r_{n-1}$. For example, if $e > k_{n-1}$, $\overline{a(e - k_{n-1}) = 0}$ where $0 < e - k_{n-1} < k_n$. Then, by the assumption of recurrence, $\overline{a(e - k_{n-1})} > 0$. It is impossible.
Then, $\overline{ae} < m - r_{n-1}$.
Then, by recurrence, $r_n \leq \overline{ae} + h_n r_n < m - r_{n-1} + h_n r_n = m - r_{n+1}$.
Then, because $at' \equiv h_n r_n + ae$, $r_n \leq \overline{at'} = \overline{ae} + h_n r_n < m - r_{n+1}$

One supposes $f = h_n$ and $e = k_{n-1}$. Then, $t' = h_n k_n + k_{n-1} = k_{n+1}$. It is oppositite to the assumption.

Then the lemma is true for n+1. ■

Lemma 6.2.4 *The following inequalities holds :* $k_{d+1} \leq m$.

Proof If $t \in \{1, 2, ..., k_{d+1} - 1\}$, by lemma 6.2.3, $r_d \leq \overline{at} < m - r_{d+1}$ or $m - r_d \geq \overline{at} > r_{d+1}$, i.e. $r_d \leq \overline{at} < m$ or $m - r_d \geq \overline{at} > 0$ where $r_d > 0$. Then, $0 < \overline{at} < m$ or $m > \overline{at} > 0$.

Then, if $k_{d+1} > m$, there exists $t_0 \in \{1, 2, ..., k_{d+1} - 1\}$ such that $t_0 = m$, i.e. $\overline{at_0} = \overline{am} = 0$. It is impossible. ■

Lemma 6.2.5 *Let* $t,\ t' \in \{1, 2, ..., k_{d+1} - 1\}$ *such that* $\overline{at} = \overline{at'}$. *Then, t=t'.*

Proof Suppose $t > t'$. Then, $a(t - t') \equiv 0$ and $\overline{a(t - t')} = 0$. Then, by lemma 6.2.3, $r_d \leq \overline{a(t - t')} < m - r_{d+1}$ or $m - r_d \geq \overline{a(t - t')} > r_{d+1} = 0$ where $r_d > 0$. Then, $0 < a(t - t')$. It is a contradiction. ■

Lemma 6.2.6 *Let n=1,2,...,d. Let* $H_n = h_1k_1 + h_2k_2 + h_3k_3 + ... + h_nk_n$. *Then,* $H_n = k_{n+1} + k_n - 1$.

Proof We have $H_n = h_1k_1 + h_2k_2 + h_3k_3 + ... + h_{n-1}k_{n-1} + h_nk_n$

$= k_2 - k_0 + k_3 - k_1 + k_4 - k_2 + k_5 - k_3 + k_6 - k_4 + k_7 - k_5 + ... + k_n - k_{n-2} + k_{n+1} - k_{n-1}$.

Therefore, if n=2m,

$$H_n =$$

$$k_2 - k_0 + k_3 - k_1 + k_4 - k_2 + k_5 - k_3 + k_6 - k_4 + ... + k_{2m} - k_{2m-2} + k_{2m+1} - k_{2m-1}$$

$$= k_2 - k_0 + k_4 - k_2 + k_6 - k_4 + ... + k_{2m} - k_{2m-2}$$

$$+ k_3 - k_1 + k_5 - k_3 + k_7 - k_5 + ... + k_{2m+1} - k_{2m-1}$$

$$= k_{2m} - k_0 + k_{2m+1} - k_1 = k_{n+1} + k_n - k_1 - k_0 = k_{n+1} + k_n - 1.$$

If n=2m+1

$$H_n =$$

$$k_2 - k_0 + k_3 - k_1 + k_4 - k_2 + k_5 - k_3 + k_6 - k_4 + ... + k_{2m+1} - k_{2m-1} + k_{2m+2} - k_{2m}$$

$$= k_2 - k_0 + k_4 - k_2 + k_6 - k_4 + ... + k_{2m+2} - k_{2m}$$

$$+ k_3 - k_1 + k_5 - k_3 + k_7 - k_5 + ... + k_{2m+1} - k_{2m-1}$$

$$= k_{2m+2} - k_0 + k_{2m+1} - k_1 = k_{n+1} + k_n - 1\ .$$ ■

Lemma 6.2.7 *Let n=1,2,3,...,d-1 . Let* $L_n = \{t | t = 0, 1, 2, ..., H_n\}$. *Then, for all* $n \geq 1$, $L_{n+1} = \{t = l + gk_{n+1} | l \in L_n, g \leq h_{n+1}\}$.

Proof Let $l \in L_n$, $l \leq H_n$. Let $g \leq h_{n+1}$.
Therefore, if $t = l + gk_{n+1}$, $t \leq H_n + h_{n+1}k_{n+1} = H_{n+1}$.
Therefore, $\{t = l + gk_{n+1} | l \in L_n, g \leq h_{n+1}\} \subset L_{n+1}$.

Reciprocally, let $t \in L_{n+1}$ and let $t = fk_{n+1} + e$, $e < k_{n+1}$ be the Euclidean division of t by k_{n+1}.

We know that $H_n = k_{n+1} + k_n - 1 \geq k_{n+1}$. Therefore, $e \leq H_n$. Therefore, $e \in L_n$.

Therefore, if $f \leq h_{n+1}$, $t = fk_{n+1} + e \in \{t = l + gk_{n+1} | l \in L_n, g \leq h_{n+1}\}$.

Moreover, if $f > h_{n+1} + 1$, $t = fk_{n+1} + e \geq (h_{n+1} + 2)k_{n+1} + e \geq h_{n+1}k_{n+1} + 2k_{n+1} = H_{n+1} - H_n + 2k_{n+1} = H_{n+1} - k_{n+1} - k_n + 1 + 2k_{n+1} = H_{n+1} + k_{n+1} - k_n + 1 \geq H_{n+1} + 1$. Therefore, $t \notin L_{n+1}$.

Then, suppose $f = h_{n+1} + 1$. Then, $t = fk_{n+1} + e = (h_{n+1} + 1)k_{n+1} + e = h_{n+1}k_{n+1} + k_{n+1} + e = H_{n+1} - H_n + k_{n+1} + e = H_{n+1} - k_{n+1} - k_n + 1 + k_{n+1} + e = H_{n+1} - k_n + 1 + e$.
Because $t \in L_{n+1}$ and $t = H_{n+1} - k_n + 1 + e$, $e + 1 - k_n \leq 0$. Therefore, $e \leq k_n - 1$.
Therefore, $t = fk_{n+1} + e = h_{n+1}k_{n+1} + k_{n+1} + e$,
where $k_{n+1} + e \leq k_{n+1} + k_n - 1 = H_n$
Therefore, $t = h_{n+1}k_{n+1} + e'$ where $e' \leq H_n$.
Therefore, $t \in \{t = l + gk_{n+1} | l \in L_n, g \leq h_{n+1}\}$.
Therefore, $L_{n+1} \subset \{t = l + gk_{n+1} | l \in L_n, g \leq h_{n+1}\}$.

Therefore, $L_{n+1} = \{t = l + gk_{n+1} | l \in L_n, g \leq h_{n+1}\}$. ■.

Lemma 6.2.8 *Let* $F_n = \{\overline{at} | t = 0, 1, 2, ..., H_n\}$.
Let $E_n = \{\overline{at} + km | t = 0, 1, 2, ..., H_n, k \in \mathbb{Z}\}$. *We set* $E_n = \{o_s^n | s \in \mathbb{Z}\}$ *where* $o_0^n = 0$ *et* $o_{s+1}^n > o_s^n$ *for all* $s \in \mathbb{Z}$.
Then, for all $s \in \mathbb{Z}$, $o_{s+1}^n - o_s^n = r_n$ *or* $o_{s+1}^n - o_s^n = r_{n+1}$.

Proof We prove this lemma by recurrence.
Suppose n=1. Then, $r_1 = a$, $H_1 = h_1k_1 = k_2 = h_1$. Therefore,
$F_1 = \{\overline{at} | t = 0, 1, 2, ..., h_1\} = \{0, a, 2a, ..., h_1a\} = \{0, r_1, 2r_1, ..., h_1r_1 = m - r_2\}$.
Therefore, the lemma is true for n=1.

Suppose that the lemma is true for n.

Then, $E_{n+1} = \{\overline{at} + km | t = 0, 1, 2, ..., H_{n+1}, k \in \mathbb{Z}\}$,
where $H_{n+1} = h_1k_1 + h_2k_2 + h_3k_3 + ... + h_{n+1}k_{n+1} = H_n + h_{n+1}k_{n+1}$.

Because $t \in \{0, 1, 2, ..., H_{n+1}\}$, $t \in L_{n+1}$. By lemma 6.2.7, si $t \in L_{n+1}$, $t = l + gk_{n+1}$ where $g \leq h_{n+1}$. By lemma 6.2.2, $\overline{at} \equiv \overline{a(l + gk_{n+1})} \equiv \overline{al} + (-1)^{n+2}gr_{n+1} \equiv \overline{al} + (-1)^n gr_{n+1}$.

Therefore,

$$\begin{aligned} E_{n+1} &= \{\overline{at} + km | t \in L_{n+1}, k \in \mathbb{Z}\} \\ &= \{\overline{at} + km | t = l + gk_{n+1}, l \in L_n, g \leq h_{n+1}, k \in \mathbb{Z}\} \\ &= \{\overline{al} + (-1)^n gr_{n+1} + km | l \in L_n, g \leq h_{n+1}, k \in \mathbb{Z}\} \\ &= \{f + (-1)^n gr_{n+1} + km | f \in F_n, g \leq h_{n+1}, k \in \mathbb{Z}\} \\ &= \{o_s^n + (-1)^n gr_{n+1} + km | s \in Z, g \leq h_{n+1}, k \in \mathbb{Z}\}. \end{aligned}$$

Suppose that n is even.
Then, $o_s^n + (-1)^n gr_{n+1} = o_s^n + gr_{n+1} \leq o_s^n + r_n - r_{n+2}$ because $gr_{n+1} \leq h_{n+1}r_{n+1} = r_n - r_{n+2}$.

Use the recurrence. Suppose $o_{s+1}^n - o_s^n = r_n$. Then, $o_s^n + (-1)^n gr_{n+1} \leq o_s^n + r_n - r_{n+2} = o_{s+1}^n - r_{n+2}$.
Therefore,
$\{o_t^{n+1} \mid o_s^n \leq o_t^{n+1} < o_{s+1}^n\} = \{o_s^n < o_s^n + r_{n+1} < .... < o_s^n + h_{n+1}r_{n+1} < o_{s+1}^n\}$.
Therefore, $o_{t+1}^{n+1} - o_t^{n+1} = r_{n+1}$ or r_{n+2} if $o_s^n \leq o_t^{n+1} < o_{t+1}^{n+1} \leq o_{s+1}^n$.

Suppose $o_{s+1}^n - o_s^n = r_{n+1}$. Then, s is fixed .
Let $T = min\{t = 0, 1, ..., |o_{s+t+1}^n - o_{s+t}^n = r_n\}$. Therefore, $o_{s+T+1}^n - o_{s+T}^n = r_n$.
Let $O = \cup_{t=0}^{T}\{o_{s+t}^n + gr_{n+1} \mid 0 \leq g \leq h_{n+1}\}$.
Then, $O = \{o_s^n, o_{s+1}^n, ..., o_{s+T-1}^n\} \cup \{o_{s+T}^n + gr_{n+1} | \ 0 \leq g \leq h_{n+1}\}$.
Therefore, $O = \{o'_s, o'_{s+1}, ..., o'_{s+K}\}$ where $o'_{s'+1} - o'_{s'} = r_{n+1}$. Moreover, $o_{s+T+1}^n - o'_{s+K} = r_n - h_{n+1}r_{n+1} = r_{n+2}$.
Therefore, if $o_{t'}^{n+1}$ and $o_{t'+1}^{n+1} \in \{o_t^{n+1} \mid o_s^n \leq o_t^{n+1} \leq o_{s+T+1}^n\}$, $o_{t'+1}^{n+1} - o_{t'}^{n+1} = r_{n+1}$ or r_{n+2}.

Suppose that n is odd.
Then, $o_s^n + (-1)^n gr_{n+1} = o_s^n - gr_{n+1} \geq o_s^n - r_n + r_{n+2}$ because $gr_{n+1} \leq h_{n+1}r_{n+1} = r_n - r_{n+2}$.

Suppose $o_s^n - o_{s-1}^n = r_n$. Then, $o_s^n + (-1)^n gr_{n+1} \geq o_s^n - r_n + r_{n+2} = o_{s-1}^n - r_{n+2}$.
Therefore,
$\{o_t^{n+1} \mid o_s^n \geq o_t^{n+1} > o_{s-1}^n\} = \{o_s^n > o_s^n - r_{n+1} > ..... > o_s^n - h_{n+1}r_{n+1} > o_{s-1}^n\}$.
Therefore, $o_t^{n+1} - o_{t-1}^{n+1} = r_{n+1}$ or r_{n+2} if $o_s^n \geq o_t^{n+1} > o_{t-1}^{n+1} \geq o_{s-1}^n$.

Suppose $o^n_s - o^n_{s-1} = r_{n+1}$. Let $T = min\{t = 0, 1, ..., \ |o^n_{s-t} - o^n_{s-t-1} = r_n\}$.
Therefore, $o^n_{s-T} - o^n_{s-T-1} = r_n$
Let $O = \cup^T_{t=0}\{o^n_{s-t} - gr_{n+1} \mid 0 \leq g \leq h_{n+1}\}$.
Then, $O = \{o^n_s, o^n_{s-1}, ..., o^n_{s-T+1}\} \cup \{o^n_{s-T} - gr_{n+1}| \ 0 \leq g \leq h_{n+1}\}$.
Therefore, $O = \{o'_s, o'_{s-1}, ..., o'_{s-K}\}$ where $o'_{s'} - o'_{s'-1} = r_{n+1}$. Moreover, $o'_{s-K} - o^n_{s-T-1} = r_n - h_{n+1}r_{n+1} = r_{n+2}$.
Therefore, if $o^{n+1}_{t'}$ and $o^{n+1}_{t'-1} \in \{o^{n+1}_t \mid o^n_s \geq o^{n+1}_t \geq o^n_{s-T-1}\}$, $o^{n+1}_{t'} - o^{n+1}_{t'-1} = r_{n+1}$ or r_{n+2}. ■

Proof 6.2.9 *Now one proves theorem 7.*

Suppose that n is even.
Then, $\overline{k_{n-1}a} = r_{n-1}$, $\overline{2k_{n-1}a} = 2r_{n-1}$, ...$\overline{h_{n-1}k_{n-1}a} = h_{n-1}r_{n-1} = r_n - r_{n-2}$.
Now, $\overline{ak_{n-1}\ell} = \overline{\ell r_{n-1}} = \ell r_{n-1}$ for $\ell = 0, 1, 2, ..., h_{n-1}$.
Therefore,
$\big\{(k_{n-1}\ell, r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\} = \big\{(k_{n-1}\ell, \overline{ak_{n-1}\ell})\big|\ell = 0, 1, 2, ..., h_{n-1}\big\} \subset E_2$.
Moreover, $r_{n-2} = h_{n-1}r_{n-1} + r_n$. On the other hand, by lemma 6.2.8 , all the points of $E_2 = (t, \overline{at})$, $t \leq H_{n-1}$, have ordinates distant of r_n or r_{n-1}.
Therefore, if there is other points of $E_2 \cap \big\{[0, H_{n-1}] \otimes [0, r_{n-2}[\big\}$ that the points $\big\{(k_{n-1}\ell, r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$, there exists $\ell_0 \in \{1, 2, ..., h_{n-1}\}$ and $(x_1, y_1) \in E_2 \cap \big\{[0, H_{n-1}] \otimes [0, r_{n-2}]\big\}$ such that $r_{n-1}\ell_0 - y_1 = r_n$.

Because $H_{n-1} = k_n + k_{n-1} - 1 < k_{n+1} \leq k_{d+1}$, by lemma 6.2.5, there exists an only $t \in \{1, ..., H_{n-1}\}$, such that $\overline{at} = y_1 : t = x_1$. Because $y_1 \neq 0$, there exists an only $t \in \{0, 1, ..., H_{n-1}\}$, such that $\overline{at} = y_1$.
Now, $r_{n-1}\ell_0 - y_1 = \overline{a\ell_0 k_{n-1}} - \overline{at} = r_n = \overline{-ak_n}$. Then, $\overline{a\ell_0 k_{n-1}} - \overline{-ak_n} = \overline{at}$. Then, $\overline{a(\ell_0 k_{n-1} + k_n)} = \overline{at}$.
Because $r_{d-1} = h_d r_d$ with $r_{d-1} > r_d$, $h_d \geq 2$. Moreover, $d \geq n \geq 2$. Then, $d - 1 > 0$. Then, $k_{d-1} > 0$.
Then, by lemma 6.2.4, $0 < k_{n-1} + k_n \leq \ell_0 k_{n-1} + k_n \leq h_{n-1}k_{n-1} + k_n \leq k_n - k_{n-2} + k_n = 2k_n - k_{n-2} \leq 2k_d < 2k_d + k_{d-1} \leq h_d k_d + k_{d-1} = k_{d+1} \leq m$. Then, $0 < \ell_0 k_{n-1} + k_n < k_{d+1}$.
Now $0 < t \leq H_{n-1} = k_n + k_{n-1} - 1 < k_d + k_{d-1} \leq k_{d+1}$. Moreover, $0 < \ell_0 k_{n-1} + k_n < k_{d+1}$.
Then, because $\overline{a(\ell_0 k_{n-1} + k_n)} = \overline{at}$, by lemma 6.2.5, $t = \ell_0 k_{n-1} + k_n$.
Then, $t = \ell_0 k_{n-1} + k_n \geq k_{n-1} + k_n > H_{n-1}$. It is a contradiction.

Therefore, there is not other points of $E_2 \cap \big\{[0, H_{n-1}] \otimes [0, r_{n-2}[\big\}$ that $\big\{(k_{n-1}\ell, r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$.
Therefore, there is not other points of $E_2 \cap \big\{[0, k_n[\otimes [0, r_{n-2}[\big\}$ that the points $\big\{(k_{n-1}\ell, r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$.

Therefore,

$$E_2 \cap \left\{[0, k_n[\otimes[0, r_{n-2}[\right\} = \left\{(k_{n-1}\ell, r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\} .$$

According to what precedes,

$$\left\{(k_{n-1}\ell, \overline{ak_{n-1}\ell})\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\} = \left\{(k_{n-1}\ell, r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}$$

is located on the straight line $y = (r_{n-1}/k_{n-1})x$ if n is even.

Suppose that n is odd. Then, $\overline{k_{n-2}a} = r_{n-2}$, $\overline{k_{n-2}a + k_{n-1}a} = r_{n-2} - r_{n-1}$, $\overline{k_{n-2}a + 2k_{n-1}a} = r_{n-2} - 2r_{n-1}$, ..., $\overline{k_{n-2}a + h_{n-1}k_{n-1}a} = r_{n-2} - h_{n-1}r_{n-1}$.

Therefore,

$$\left\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}$$

$$= \left\{(k_{n-2} + k_{n-1}\ell\ , \overline{k_{n-2}a + \ell k_{n-1}a}\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}\ \subset E_2.$$

For $\ell = 0, 1, 2, ..., h_{n-1}$, $k_{n-2} + k_{n-1}\ell \leq k_{n-2} + h_{n-1}k_{n-1} = k_n$. Therefore,

$$\left\{(k_{n-2}+k_{n-1}\ell\ ,\ r_{n-2}-r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\} \subset E_2 \cap \left\{]0, k_n]\otimes]0, r_{n-2}]\right\} .$$

Moreover, $r_{n-2} - h_{n-1}r_{n-1} = r_n$. On the other hand, by lemma 6.2.8 , all the points of $E_2 = (t, \overline{at})$, $t \leq H_{n-1}$, have ordinates distant of r_n or r_{n-1}.

Therefore, if there is other points of $E_2 \cap \left\{[0, H_{n-1}]\otimes]0, r_{n-2}]\right\}$ that the points $\left\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\middle|\ell = 0, 1, 2, ..., h_{n-1}\right\}$, there exists $\ell_0 \in \{1, 2, ..., h_{n-1}\}$ and $(x_1, y_1) \in E_2 \cap \left\{[0, H_{n-1}]\otimes[0, r_{n-2}]\right\}$ such that $y_1 - (r_{n-2} - r_{n-1}\ell_0) = r_n$.

Because $H_{n-1} = k_n + k_{n-1} - 1 < k_{n+1} \leq k_{d+1}$, by lemma 6.2.5, there exists an only $t \in \{1, ..., H_{n-1}\}$, such that $\overline{at} = y_1$. Because $y_1 \neq 0$, there exists an only $t \in \{0, 1, ..., H_{n-1}\}$, such that $\overline{at} = y_1$.

Then, $y_1 - (r_{n-2} - r_{n-1}\ell_0) = \overline{at} - \overline{k_{n-2}a + \ell_0 k_{n-1}a} = r_n = \overline{ak_n}$. Then, $\overline{at} = \overline{k_{n-2}a + \ell_0 k_{n-1}a} + \overline{ak_n}$. Then, $\overline{at} = \overline{a(k_{n-2} + \ell_0 k_{n-1} + k_n)}$.

Now, because $r_{d-1} = h_d r_d$ with $r_{d-1} > r_d$, $h_d \geq 2$. Now, $n \geq 3$. Then, $d - 1 \geq n - 1 > 1$. Then, $k_{d-1} > 0$.

Then $0 < k_{n-2} + \ell_0 k_{n-1} + k_n \leq k_{n-2} + h_{n-1}k_{n-1} + k_n \leq 2k_n \leq 2k_d < 2k_d + k_{d-1} \leq h_d k_d + k_{d-1} = k_{d+1} \leq m$.

Now $0 < t \leq H_{n-1} = k_n + k_{n-1} - 1 < k_d + k_{d-1} \leq k_{d+1}$.

Then, because $\overline{a(k_{n-2} + \ell_0 k_{n-1} + k_n)} = \overline{at}$, by lemma 6.2.5, $t = k_{n-2} + \ell_0 k_{n-1} + k_n$.

Then, $t = k_{n-2} + \ell_0 k_{n-1} + k_n \geq k_{n-1} + k_n > H_{n-1}$. It is a contradiction.

Therefore, there is not other points of $E_2 \cap \big\{[0, H_{n-1}]\otimes]0, r_{n-2}]\big\}$ that the points $\big\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$.

Therefore, there is not other points of $E_2 \cap \big\{]0, k_n]\otimes]0, r_{n-2}]\big\}$ that the points $\big\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$: i.e.

$$E_2 \cap \big\{]0, k_n]\otimes[0, r_{n-2}[\big\} = \big\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}\ .$$

According to what precedes,

$$\big\{(k_{n-2} + k_{n-1}\ell\ ,\ r_{n-2} - r_{n-1}\ell)\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$$

$$= \big\{(k_{n-2} + k_{n-1}\ell\ , \overline{ak_{n-2} + ak_{n-1}\ell})\big|\ell = 0, 1, 2, ..., h_{n-1}\big\}$$

is located on a straight line. ■

Chapter 7

Congruences of Fibonacci

In this chapter, we study the transformation of a random variable $Y \in F(m)$ defined a probability space $(\Omega, \mathcal{A}, P)$: $X = \widehat{T}(Y)$ where T is Fibonacci congruence and et where $\widehat{T}$ is defined by the following way.

Notations 7.0.1 *We set* $\widehat{T}(k/m) = \overline{T}(k)/m$.

We suppose that the probability density function of Y with respect to μ_m is written in a form :

$$h(y)\big[1 + \frac{\eta(y)}{co}\big] \ ,$$

where $\eta(y)$ is a sample of a white noise independent of h and where

$$\int h(y)\mu_m(dy) \approx 1.$$

One will study $P\{X \in I\}$ where I is an interval $I = [c/m, c'/m[$. One will study especially the cases where h is the probability density function of the normal or uniform distribution.

7.1 Distribution of normal type

One thus will calculate $P\{X \in I\} = P\{Y \in \widehat{T}^{-1}(I)\} = P_Y\{\widehat{T}^{-1}(I)\} = P_X\{I\}$.

It will be supposed that $y \mapsto h(y)$ has a curve in the shape of bell, (e.g. a Gaussian curve). That will result in the following assumption.

Hypothesis 7.1.1 *Let I be an interval of F(m). Let N(I) be the number of points of F(m) belonging to I. We assume that, for all I such that* $N(I) >> 1$,

$$\sqrt{N(I)}\big[\sum_{k=c}^{c'} \frac{h\big(\widehat{T}^{-1}(k/m)\big)}{N(I)} - 1\big] \approx 0 \ .$$

This assumption means well that the curve of h has properties enough close to those of a normal curve. Indeed, by property 7.1.11 the following approximation holds

$$\frac{1}{N(I)}\sum_{k=c}^{c'} h\big(\widehat{T}^{-1}(k/m)\big) = 1 + \frac{O(1)}{N(I)} \ ,$$

when h is the Gaussian curve : $y \mapsto \frac{10.e^{-[10(y-0.5)]^2/(2\sigma^2)}}{2\pi\sigma^2}$.

Now, it is thus a question of calculating the approximation of L(I) by $P_X\{I\}$. We set $\epsilon_I^G = \big|(N(I))/m - P_Y\{\widehat{T}^{-1}(I)\}\big|$

For example, let us suppose that h is an approximation of the normal law $N(0,\sigma^2)$ corresponding to a summation of S terms $Z_n \in F(m)$, with S=10, $E\{Z_n\} = 1/2$, $E\{(Z_n - 1/2)^2\} = \sigma^2$: $0 \leq Z_1 + ... + Z_{10} < 10$, i.e. $Z_1 + ... + Z_{10} - 10/2$ has a distribution close to $N(0, 10.\sigma^2)$. One thus considers that $Y = (Z_1 + ... + Z_{10} - 5)/\sqrt{10}$ has a distribution close to $N(0,\sigma^2)$. Finally,

$$P_X\{I\} = \sum_{k=c}^{c'-1} P_Y(\widehat{T}^{-1}(k/m))$$

$$\approx \sum_{k=c}^{c'-1} \frac{10}{m} \ \frac{e^{-[10(\widehat{T}^{-1}(k/m)-0.5)\ ]^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \big[1 + \frac{\eta(\widehat{T}^{-1}(k/m))}{co}\big] \ .$$

7.1.1 Case $co \geq 10$

At first, let us consider the following example where $co \geq 10$: it represents an approximation of the law of $g(j) = \sum_{i=1}^{10} f(i,j)$ used in the construction of the random bits $b^1(n')$: cf section 11.1.2.

Example 7.1.1 *Suppose that $\eta(y) = u(y)$ is a sample which has the uniform distribution over [-1/2,1/2]. Suppose that $co \geq 10$. Assume that $h(y) \approx 10 \ \frac{e^{-[10(y-0.5)]^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$.*

Study Under these assumptions,

$$P_Y(Y = y) \approx \frac{10}{m} \ \frac{e^{-[10(y-0.5)]^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \Big[1 + \frac{\eta(y)}{co}\Big] \ .$$

Let us notice that to suppose that $\eta(y)$ is a sample of uniform distribution is a difficult assumption to ensure in the construction of the random bits $b^1(n')$

used in section 11.2.5). But the important thing, it is the variance associated with $\eta(y)$ (e.g. cf proposition 7.2.2). ■

In the general case, one can consider the model

$$P_X(I) = \sum_{k=c}^{c'-1} \frac{1}{m}\ h\big(\widehat{T}^{-1}(k/m)\big)\ \Big[1 + \frac{\eta(\widehat{T}^{-1}(k/m))}{co}\Big]\ .$$

In order to simplify, the following notations are adopted.

Notations 7.1.1 *We set*

$$P_X(I) = \frac{N(I)}{m} \sum_{k=c}^{c'-1} \frac{g(k)}{N(I)}\big[1 + \frac{U_k(\omega)}{co}\big]\ ,$$

where $g(k) = h\big(\widehat{T}^{-1}(k/m)\big)$ *and where* U_k *is an IID sequence of random variable with mean 0 satisfying* $U_k(\omega) = u_k = \eta(\widehat{T}^{-1}(k/m))$.

One can then study the approximation of L(I). At first, one will use the following notation.

Notations 7.1.2 *We set* $\epsilon_I^G = P_X(I) - N(I)/m$.

For this study, the following lemmas are used.

Lemma 7.1.2 *The following equalities hold.*

$$\epsilon_I^G = \frac{N(I)}{m}\Big(\sum_k \frac{g(k)\big[1 + \frac{U_k}{co}\big] - 1}{N(I)}\Big) = \frac{\sqrt{N(I)}}{m}\ \sum_k \frac{g(k)\big[1 + \frac{U_k}{co}\big] - 1}{\sqrt{N(I)}}\ .$$

Lemma 7.1.3 *Let* σ_g^2 *be the variance of* $\sum_k \frac{g(k)\left[1+\frac{U_k}{co}\right]-1}{N(I)}$. *Let* σ_U^2 *be the variance of* U_1.

Then, $\sigma_g^2 = \frac{\sigma_U^2 E(g^2)}{co^2 N(I)}$, *where* $E(g^2) = \sum_k \frac{g(k)^2}{N(I)}$

Proof At first, $E\Big\{\sum_k \frac{g(k)\left[1+\frac{U_k}{co}\right]}{N(I)}\Big\} = \sum_k \frac{g(k)}{N(I)}$.

Moreover,

$$E\Big\{\Big[\sum_k \frac{g(k)\big[1 + \frac{U_k}{co}\big]}{N(I)}\Big]^2\Big\}$$

$$= E\Big\{\sum_{k,k'} \frac{g(k)g(k')\big[1 + U_k/co + U_{k'}/co + U_kU_{k'}/co^2\big]}{N(I)^2}\Big\}$$

$$= \sum_{k,k'} \frac{g(k)g(k')}{N(I)^2} + \sum_{k,k'} \frac{g(k)g(k')E\{U_k U_{k'}/co^2\}}{N(I)^2}$$

$$= \Big[\sum_k \frac{g(k))}{N(I)}\Big]^2 + \sum_k \frac{g(k)^2 E\{U_k^2/co^2\}}{N(I)^2}$$

$$= \Big[\sum_k \frac{g(k))}{N(I)}\Big]^2 + \sum_k \frac{\sigma_U^2 g(k)^2}{co^2 N(I)^2}$$

$$= \Big[\sum_k \frac{g(k))}{N(I)}\Big]^2 + \frac{\sigma_U^2}{co^2 N(I)} \sum_k \frac{g(k)^2}{N(I)}$$

$$= \Big[\sum_k \frac{g(k))}{N(I)}\Big]^2 + \frac{\sigma_U^2 E(g^2)}{co^2 N(I)} \ .$$

Moreover,

$$\sigma_g^2 = E\Big\{\Big[\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]}{N(I)} - 1\Big]^2\Big\} - \Big[\sum_k \frac{g(k)}{N(I)} - 1\Big]^2$$

$$= E\Big\{\Big[\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]}{N(I)}\Big]^2\Big\} - 2E\Big\{\Big[\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]}{N(I)}\Big\} + 1 - \Big[\sum_k \frac{g(k)}{N(I)}\Big]^2 + 2\sum_k \frac{g(k)}{N(I)} - 1$$

$$= E\Big\{\Big[\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]}{N(I)}\Big]^2\Big\} - 2\sum_k \frac{g(k)}{N(I)} \ - \Big[\sum_k \frac{g(k)}{N(I)}\Big]^2 + 2\sum_k \frac{g(k)}{N(I)}$$

$$= E\Big\{\Big[\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]}{N(I)}\Big]^2\Big\} \ - \Big[\sum_k \frac{g(k)}{N(I)}\Big]^2$$

$$= \frac{\sigma_U^2 E(g^2)}{co^2 N(I)}. \ \blacksquare$$

This result means that σ_g^2 decreases if N(I) increase.

Example 7.1.4 *Assume that the assumptions of the example 7.1.1 hold with co=10. Assume that σ is not too big. Assume that the points $\widehat{T}^{-1}(k/m)$ are well distributed in F(m). Assume m and N(I) are enough big. Then,*

$$\sigma_g^2 \approx \frac{1}{120\sqrt{4\pi\sigma^2}\ N(I)} \ .$$

Proof We have

$$\sum_k \frac{g(k)^2 E\{U_k^2/co^2\}}{N(I)^2} = \sum_k \frac{g(k)^2}{12*100*N(I)^2}$$

$$= \frac{1}{1200N(I)} \sum_k \frac{g(k)^2}{N(I)}$$

$$= \frac{1}{120N(I)} \frac{1}{N(I)} \sum_k \frac{10e^{-[10\widehat{T}^{-1}(k/m)-0.5]^2/\sigma^2}}{2\pi\sigma^2}$$

$$= \frac{1}{120\ N(I)} \frac{1}{\sqrt{2*2\pi\sigma^2}} \frac{1}{N(I)} \sum_k \frac{10e^{-[10\widehat{T}^{-1}(k/m)-0.5]^2/(2\sigma^2/2)}}{\sqrt{2\pi[\sigma^2/2]}}$$

$$\approx \frac{1}{120\sqrt{4\pi\sigma^2}\ N(I)}. \blacksquare$$

Lemma 7.1.5 *Let* $Y_{No}^* = \sum_k \frac{g(k)\frac{U_k}{co}}{N(I)\sigma_g}$. *If N(I) is enough big, if* σ_U *and* $E(g^2)$ *are not too small,*

$$Y_{No}^* \approx \sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)\sigma_g}\ .$$

Proof By hypothesis 7.1.1, we know that $\sqrt{N(I)}\big[\sum_k \frac{g(k)}{N(I)} - 1\big] \approx 0$ if N(I) is enough big. Then,

$$Y_{No}^* = \sum_k \frac{g(k)\frac{U_k}{co}}{N(I)\sigma_g}$$

$$= \sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)\sigma_g} - \sum_k \frac{g(k)-1}{N(I)\sigma_g}$$

$$= \sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)\sigma_g} - \frac{\sum_k \frac{g(k)}{N(I)} - 1}{\sigma_g}$$

$$= \sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)\sigma_g} - co\sqrt{N(I)}\frac{\sum_k \frac{g(k)}{N(I)} - 1}{\sigma_U\sqrt{E(g^2)}}$$

$$\approx \sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)\sigma_g}\ . \blacksquare$$

Lemma 7.1.6 *Let* $Y_{No} = \sum_k \frac{g(k)\left[1+\frac{U_k}{co}\right]-1}{N(I)\sigma_g}$. *Then,*

$$P\Big\{\frac{m.co|\epsilon_I^G|}{\sigma_U\sqrt{E(g^2)}\sqrt{N(I)}} \geq b\Big\} = P\{|Y_{No}| \geq b\} \ .$$

Proof We have

$$\epsilon_I^G = \frac{N(I)}{m}\Big(\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)}\Big) = \frac{N(I)\sigma_g}{m}\Big(\sum_k \frac{g(k)\big[1+\frac{U_k}{co}\big]-1}{N(I)\sigma_g}\Big)$$

$$= \frac{N(I)\sigma_g}{m}\ Y_{No} = \frac{N(I)}{m}\ \frac{\sigma_U\sqrt{E(g^2)}}{co\sqrt{N(I)}}\ Y_{No} = \frac{\sqrt{N(I)}}{m}\ \frac{\sigma_U\sqrt{E(g^2)}}{co}\ Y_{No}\ . \blacksquare$$

Of course, $E\{Y_{No}^2\} = 1$ and Y_{No} has asymptotically the distribution N(0,1). Then one can use the following result.

Proposition 7.1.1 *If N(I) is enough big,*

$$P\Big\{\frac{m.co|\epsilon_I^G|}{\sigma_U\sqrt{E(g^2)}\sqrt{N(I)}} \geq b\Big\} \approx \Gamma(b) \ .$$

This result means that, if b is enough big, one will be able to admit that

$$|\epsilon_I^G| \leq \frac{\sigma_U\sqrt{E(g^2)}\sqrt{N(I)}.b}{m.co} \ . \tag{7.1}$$

By this way, one has an increase of $|\epsilon_I^G|$.

Normal distribution

Now assume that h(y) is a normal probability density function and apply the previous result. Under the condition of example 7.1.1

$$P\Big\{\frac{m\sqrt{\sigma}|\epsilon_I^G|}{0.0485\sqrt{N(I)}} \geq b\Big\} \approx \Gamma(b) \ .$$

This result thus gives us a probable increase of $|\epsilon_I^G|$.

Example 7.1.7 *Suppose that we do not have more than* 10^6 *possible intervals I. We know that* $\Gamma(6) \leq 10^{-9}$. *Then, if N(I) is not too small, one can assume*

$$|\epsilon_I^G| \leq \frac{0.291\sqrt{N(I)}}{m\sqrt{\sigma}} \ . \tag{7.2}$$

Therefore, normally, ϵ_I^G increases if σ decrease or if N(I) increases.

In fact, the checking of inequality 7.2 depends on the number of possible intervals I in F(m) (that is equal to m(m+1)/2). Indeed, at first, one can suppose that the intervals behave like a sample. In this case, one must choose b according to m(m+1)/2 as one will do it in section 7.2.

Thus, if there are much more possible intervals " I " than $2 * 10^9$, the inequality 7.2 can be not checked, and, this, with a non-negligible probability : $P\{|Y_{No}| \geq 6\} \approx 2 * 10^9$. For example, if $m \geq 10^9$, there is $10^9(10^9 + 1)/2$ possible intervals " I ".

However, generally, one can have a better increase. That is confirmed by various numerical results.

Example 7.1.8 *Suppose m= 267914296 , a= 165580141. We choose intervals I length* $L(I) = (1/5)10^{-j}$ *for j =1,...,6 .*

Choose standard deviations $\sigma = 1/2$*, 1/4, 1/8, 1/40.*

Study For each j, one calculated each ϵ_I^G for 50 intervals I_s, s=1,2,...,50 length $(1/5)10^{-j}$.
Then, one obtains the following table of the maxima $Max_s\{0.5*10^6|\epsilon_I^G| \mid I_s,\ \sigma\}$ on these 50 terms.

L(I) $\setminus\ \sigma$	1/2	1/4	1/8	1/40
$(1/5)10^{-1}$	0.0708	0.0837	-0.0321	0.5361
$(1/5)10^{-2}$	0.1096	-0.0114	-0.1507	-0.1077
$(1/5)10^{-3}$	0.0067	0.0328	0.0097	-0.1834
$(1/5)10^{-4}$	-0.0008	0.0004	0.0046	0.0083
$(1/5)10^{-5}$	-0.0008	-0.0014	0.0025	-0.0152
$(1/5)10^{-6}$	-0.0000	-0.0010	0.0010	0.0013
$(1/5)10^{-7}$	-0.0006	0.0002	-0.0032	-0.0044

As each term $Max_s\{0.5*10^6|\epsilon_I^G|\}$ of the table is to be multiplied by $2/10^{-6}$, it is understood that one has results much better than those given by the example 7.1.7. ■

One could thus admit that for our calculations of construction of the sequence $b^1(n')$ used in section 11.2, the inequality 7.2 is always checked.

Hypothesis 7.1.2 *In the construction of the sequence g(j) used in section 11.2.5 , one can suppose that*

$$|\epsilon_I^G| \leq \frac{0.291 * 12^{1/4} * 10\sqrt{N(I)}}{m} \leq \frac{5.42\sqrt{N(I)}}{m} \ .$$

Indeed, that amounts supposing the inequality 7.2 with $\sigma^2 \geq 8.334 * 10^{-6} \geq \frac{1}{12}\ \frac{1}{10000}$.

Now, $\sigma^2 = \frac{1}{12}\ \frac{1}{10000}$ means that the variance of the conditional probability of the G_n, $P\{G_n \in I \mid G_{n+j_2} = g_2, ..., G_{n+j_p} = g_p\}$ is 10000 times smaller than the variance of the uniform distribution. It is a completely reasonable increase with the data used in section 11.2.

Because $N(I) \leq m$, the following result holds.

Proposition 7.1.2 *Suppose that hypothese 7.1.2 holds. Then,*

$$|\epsilon_I^G| \leq \frac{5.42}{\sqrt{m}}\ .$$

Remark 7.1.9 *The proof of previous results have only numerical proofs about sequences g(j). We only give these results to carry out a complete study of congruences of Fibonacci, but not to build the sequences of bits IID $b^1(n')$. For this construction we use indeed the sequences h(j) and mathematically proved results .*

Another distributions

We gave the previous results when h is a Gauss density. They remain equivalent for other curves in the shape of bell, especially those which one obtains as densities of sums of S random variables.

7.1.2 Case $co = \infty$

Now, it is possible that the curve of the law of the sums of random variables associated with our data is smoother than that which we have just studied: cf figures of section 5.4. It is the case $co = \infty$. In this section, we study this case : then, we assume that the probability of Y is

$$Proba\{Y = y\} = \frac{10}{m}\ \frac{e^{-[10(y-0.5)]^2/(2\pi\sigma^2)}}{\sqrt{2\pi\sigma^2}}\ .$$

Remark 7.1.10 *It is an assumption which could be checked for the curves of the conditional probabilities of the sequence G(j) obtained in section 11.2. Indeed, we impose in section 11.2 that the lines are independent. Moreover we make uniform the probabilities by pseudo-random generators and functions of Fibonacci to avoid finding models like those of the section 5.4 : figures 5.7 and 5.8. Finally in the example of the figures 5.26 and 5.25 the laws are with values in $F^*(q+1) = \{0, 1, ..., 80\}$, whereas the laws which we use in section 11.2 are with values in $F^*(m)$ where $m > 32^{20}$. However, the numerical results of the section 5.4 show that more m is large, better is convergence to the normal distribution.*

General study

As previously,

$$P_Y(\widehat{T}^{-1}(I)) = \sum_k P_Y(\widehat{T}^{-1}(k/m)) = \frac{N(I)}{m} \sum_k \frac{g_N(k)}{N(I)} \ ,$$

where $g_N(k) = h_N\big(\widehat{T}^{-1}(k/m)\big)$ and where $h_N(y) = c_0 \frac{10.e^{-[10(y-0.5)]^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$ with $\int h_N(y).\mu_m(dy) = 1$: c_0 is an adequate constant.

Then, the following property holds.

Property 7.1.11 *Let K_N be the Lipschitz's coefficient of the curve*

$$y \mapsto \frac{10.e^{-[10(y-0.5)]^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \ .$$

Then, the following approximation holds

$$\frac{1}{N(I)} \sum_{r=1}^{N(I)} g_N(k) = 1 + O(1)\frac{K_N}{N(I)} \ .$$

Proofs We point out that a more complete proof of this property in section 7.1.3.

The $\widehat{T}(k)$'s are almost uniformly distributed : that derives from the properties of congruences of Fibonacci (one can also understand this result by numerical simulations).

One deduce from these numerical studies that

$$\frac{1}{N(I)} \sum_r h_N(r/N(I)) = 1 + \frac{2Ob(1)K_N}{N(I)} \ ,$$

if N(I) and m are big (cf also lemma 7.1.16).

As a matter of fact,

$$\frac{1}{N(I)} \sum_r h_N(r/N(I)) \approx \int h_N(y)dy = 1$$

if N(I) and m are big.

Now compute these sums. For that, one can use the method of Riemann : it is noted that

$$\frac{1}{N(I)} \sum_{r=1}^{N(I)} h_N(r/N(I)) = 1 + \frac{O(1)K_0}{N(I)} \ .$$

The numerical studies confirm that one obtains the same results for the function g_N :

$$\frac{1}{N(I)} \sum_{k=1}^{N(I)} g_N(k) = 1 + \frac{O(1)K_0}{N(I)} \ .$$

We recall that a more complete proof of this property in section 7.1.3. ■

Then, we deduce from these results that $P_X(I) = N(I)/m + \frac{O(1)C_N}{N(I)}$ where C_N is a constant as soon as N(I) is not too small.

The previous results are confirmed by numerical studies.

Example 7.1.12 *One takes the same parameters as in the example 7.1.8. One calculates the maximum of* $0.5 * 10^7 \epsilon_I^G$.

We have the following table.

L(I) $\setminus$ σ	1/2	1/4	1/8	1/40
$(1/5)10^{-1}$	0.0033	-0.0160	0.0197	-0.0613
$(1/5)10^{-2}$	-0.0003	-0.0168	0.0097	0.0643
$(1/5)10^{-3}$	-0.0080	-0.0014	0.0048	0.1157
$(1/5)10^{-4}$	-0.0032	-0.0059	0.0176	-0.1179
$(1/5)10^{-5}$	-0.0002	-0.0118	-0.0153	0.0058
$(1/5)10^{-6}$	-0.0001	0.0079	-0.0314	-0.0266
$(1/5)10^{-7}$	0.0020	-0.0247	0.0051	-0.0186

It is noticed that the ϵ_I^G's seem almost independent of L(I). Then, one can admit the following increases.

Property 7.1.13 *If N(I) and m are enough big, the following inequalities hold*

$$|\epsilon_I^G| \leq \frac{Ob(1)C_N}{N(I)} \ ,$$

$$|\epsilon_I^G| \leq \frac{Ob(1)C_N}{m} \ .$$

Remark 7.1.14 *Of course, these results confirm our study of the section 7.1.1: one can admit the hypothesis 7.1.2 :* $\epsilon_I^G \leq \frac{5.42\sqrt{N(I)}}{m}$.

Numerical study

We suppose that $X = \overline{T}(Y)/m \in F(m)$. We suppose again that the probability of Y is

$$P\{Y = y\} \approx \frac{10}{m} \frac{10.e^{-[10(y-0.5)]^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \ .$$

We suppose that T is a congruence $T(x) \equiv ax$ mod m. One will carry out the calculation of $\epsilon = \left| N(I)/m - P_X(I) \right|$ for various intervals I, a, m and σ^2.

Some results are consigned in the following paragraphs. We carried out many others of them. All of them give equivalent results.

Variation of I In this paragraph, we suppose a = 1346269, m= 2178309 , $\sigma^2 = 1/4$, and one varies intervals I.

The second column indicates the coefficient by which it is necessary to multiply the result.

I	x	ϵ	$\frac{\epsilon}{N(I)}$	$\frac{\epsilon}{N(I)/m}$
$[0, 10000[$	10^{-3}	0.00171	0.0000001	0.373341
$[0, 100000[$	10^{-4}	0.010658	0.00000010	0.232165650
$[0, 1000000[$	10^{-5}	0.111392	0.0000001	0.242646
$[51236, 1000000[$	10^{-6}	0.24874828	0.00000026	0.57111213
$[151236, 1000000[$	10^{-7}	-0.3573127	0.00000042	0.9170246
$[851236, 1000000[$	10^{-4}	0.0169700	0.000000114	0.24848780
$[851236, 974502[$	10^{-4}	0.02839505	0.00000023	0.501786457
$[951236, 974502[$	10^{-3}	0.00261713	0.0000001	0.245032
$[951236, 952502[$	10^{0}	0.00000069	0.00000000055	0.0011915
$[952236, 952502[$	10^{0}	0.0000005	0.000000002	0.004321

Variation of a In this paragraph, we suppose m= 2178309 , $\sigma^2 = 1/4$, I = [1250312,1948077[.

One varies a. One does not suppose more that a is in the Fibonacci sequence : only a= 1346269 belongs to the Fibonacci sequence.

a	x	ϵ	$\frac{\epsilon}{N(I)}$	$\frac{\epsilon}{N(I)/m}$
1346269	10^{-5}	0.0708762	0.0000001	0.2212642
1000000	10^{-5}	0.1209353	0.00000017	0.3775404
2034476	10^{-5}	0.237045	0.00000033	0.7400157
1985474	10^{-5}	0.264947	0.00000038	0.8271217
1407735	10^{-4}	0.032717	0.000000047	0.102137
1250149	10^{-4}	0.06093	0.000000087	0.1902183
425987	10^{-4}	0.0723728	0.0000001	0.225936
54812	10^{-4}	0.0715460	0.0000001	0.223355
5412	10^{-3}	0.07495	0.0000001	0.233986

Variation of m : 1 In this paragraph, we suppose m= 2178309 , $\sigma^2 = 1/4$, I = [980235,1730228[.

One varies m. One does not suppose more that m is in the Fibonacci sequence : only m=2178309 belongs to the Fibonacci sequence.

m	x	ϵ	$\frac{\epsilon}{N(I)}$	$\frac{\epsilon}{N(I)/m}$
2178309	10^{-5}	0.2535020117322	0.0000003380058	0.7362811568567
2000000	10^{-5}	0.1304767415277	0.0000001739706	0.3479412248588
3520470	10^{-5}	0.0290115355239	0.0000000386824	0.1361802583036
1988242	10^{-5}	0.1736900782922	0.0000002315889	0.4604548424371

Variation of m : 2 In this paragraph, we suppose $\sigma^2 = 1/4$. Moreover, we suppose I = [541231,1905574[.

One suppose that m and a are in the Fibonacci sequence : T is a Fibonacci congruence. One varies m.

It is noted that ϵ is about divided by 2 with each step: it is pointed out that m and a are also about multiplied by 2 ($fi_{n+1} = 2.O(fi_n)$).

m	a	x	ϵ	$\frac{\epsilon}{N(I)}$	$\frac{\epsilon}{N(I)/m}$
2178309	1346269	10^{-5}	0.423066	0.000000	0.675468
3524578	2178309	10^{-5}	0.261470	0.0000002	0.675469
5702887	3524578	10^{-5}	0.161597	0.0000001	0.675468
267914296	165580141	10^{-5}	0.003439	$2 * 10^{-9}$	0.675468
3 295128	2036 501	10^{-5}	0.00002	$2 * 10^{-11}$	0.675468

Variation of σ^2 In this paragraph, we suppose m= 2178309. Moreover, we suppose I = [541231,1905574[.

One varies σ^2 : ϵ is about divided by 2 with each step.

σ^2	x	ϵ	$\frac{\epsilon}{N(I)}$	$\frac{\epsilon}{N(I)/m}$
1/4	10^{-5}	0.4230665591742	0.00000031008812	0.6754677478084
1/2	10^{-5}	0.2848750343620	0.00000020880016	0.4548312640048
1	10^{-5}	0.1654087269709	0.00000012123691	0.2640914481470
1/16	10^{-4}	0.2253644022687	0.00000016518163	0.3598166338974
1/128	10^{-3}	0.15111108927857	0.0000001107574	0.2412638506411
1/1024	10^{-3}	0.3016983971605	0.00000022113090	0.4816914323014

Distribution different from the Normal distribution

The previous results hold when h_N is a Gauss density. Many numerical studies of the same type show that they remain equivalent for other curves in the shape of bell which one can obtain as limit densities of sums of random variables.

7.1.3 Theoretical study

Ones can almost prove mathematically a result similar at that one of property 7.1.11 when T is a Fibonacci congruence :

$$P\{\widehat{T}(Y) \in I\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] \ .$$

As a matter of fact, it will be necessary to complete our mathematical study of Fibonacci congruences in order to completely prove this result. But we can already study this property with numerical simulations : all those give the same result. Remark that theses simulations are used to know if some mathematical results hold.

First, we use the following notations.

Notations 7.1.3 *Let h_N be the probability density function of Y with respect to μ_m : $\int_0^1 h_N(u)\mu_m(du) = 1$.*

Let $h'_N = (1/c_0)h_N$ be the probability density function such that $\int_0^1 h'_N(u)du = 1$. Let $K_0 \in \mathbb{R}_+$ such that $|h_N(r) - h_N(r')| \leq K_0|r' - r|$ and $|h'_N(r) - h'_N(r')| \leq K_0|r' - r|$ when $r, r' \in [0, 1]$.

Then, we need the following lemmas.

Lemma 7.1.15 *The following equality holds :*

$$c_0 = \frac{1}{1 + \frac{Ob(1)K_0}{m}} = 1 + \frac{O(1)K_0}{m} \ .$$

Proof The following equalities hold :

$$1 = \int_0^1 h'_N(u)du = \sum_t \int_{t/m}^{(t+1)/m} h'_N(u)du$$

$$= \sum_t \int_{t/m}^{(t+1)/m} \big[h'_N(t/m) + Ob(1)K_0/m\big]du$$

$$= \frac{1}{m}\sum_t h'_N(t/m) + \frac{Ob(1)K_0}{m}$$

$$= \int_0^1 h'_N(u)\mu_m(du) + \frac{Ob(1)K_0}{m} \ .$$

Then, $\int_0^1 h'_N(u)\mu_m(du) = 1 + \frac{Ob(1)K_0}{m}$.

Therefore,

$$1 = \int_0^1 h_N(u)\mu_m(du) = c_0 \int_0^1 h'_N(u)\mu_m(du) = c_0\Big[1 + \frac{Ob(1)K_0}{m}\Big] \ .$$

Then,

$$c_0 = \frac{1}{1 + \frac{Ob(1)K_0}{m}} = 1 + \frac{O(1)K_0}{m} \ . \ \blacksquare$$

Lemma 7.1.16 *We keep the notations of lemma 7.1.15. Assume that N(I) is large enough. Then*

$$\frac{1}{N(I)}\sum_r h_N(r/N(I)) = 1 + \frac{2Ob(1)K_0}{N(I)} \ .$$

Proof The following equalities hold :

$$1 = \int_0^1 h'_N(u)du = \sum_r \int_{r/N(I)}^{(r+1)/N(I)} h'_N(u)du$$

$$= \sum_r \int_{r/N(I)}^{(r+1)/N(I)} \big[h'_N(r/N(I)) + Ob(1)K_0/N(I)\big]du$$

$$= \frac{1}{N(I)} \sum_r h'_N(r/N(I)) + \frac{Ob(1)K_0}{N(I)} \ .$$

Therefore $1 = \frac{1}{N(I)} \sum_r h'_N(r/N(I)) + \frac{Ob(1)K_0}{N(I)}$. Therefore $c_0 = \frac{1}{N(I)} \sum_r h_N(r/N(I)) + \frac{Ob(1)c_0K_0}{N(I)}$.

Therefore, if N(I) is large enough, by lemma 7.1.15,

$$c_0 = 1 + \frac{O(1)K_0}{m} = \frac{1}{N(I)} \sum_r h_N(r/N(I)) + \frac{Ob(1)[1 + \frac{O(1)K_0}{m}]K_0}{N(I)} \ .$$

Therefore, if N(I) is large enough,

$$\frac{1}{N(I)} \sum_r h_N(r/N(I)) = 1 + \frac{2Ob(1)K_0}{N(I)} \ . \blacksquare$$

Remark 7.1.17 *If $N(I) << m$,*

$$\frac{1}{N(I)} \sum_r h_N(r/N(I)) \approx 1 + \frac{Ob(1)K_0}{N(I)} \ .$$

Then, the following property holds.

Property 7.1.18 *Let $g_N(k) = h_N\big(\widehat{T}^{-1}(k/m)\big)$. Assume again that T is a Fibonacci congruence. The following approximation holds*

$$\frac{1}{N(I)} \sum_{k=1}^{N(I)} g_N(k) = 1 + \frac{6Ob(1)K_0}{N(I)} \ .$$

Proof Let k^n, n=1,2,...,c'-c, be a permutation of $T^{-1}(I) = \{a_1, ..., a_{c'-c}\}$ such that $\overline{T}^{-1}(k^1) < \overline{T}^{-1}(k^2) < \overline{T}^{-1}(k^3) < ..... < \overline{T}^{-1}(k^{c'-c})$. Then, for all numerical simulations which we executed, one has always obtained that

$$|T^{-1}(k^r/m) - r/N(I)| \leq 4/N(I) \ .$$

We deduce that $|g_N(k^r) - h_N(r/N(I))| \leq 4K_0/N(I)$.

Therefore,

$$\frac{1}{N(I)}\sum_k g_N(k) = \frac{1}{N(I)}\sum_r g_N(k^r)$$

$$= \frac{1}{N(I)}\sum_r h_N(r/N(I)) + \frac{1}{N(I)}\sum_r \left[g_N(k^r) - h_N(r/N(I))\right]$$

$$= \frac{1}{N(I)}\sum_r h_N(r/N(I)) + \frac{1}{N(I)}\sum_r 4Ob(1)K_0/N(I)$$

$$= \frac{1}{N(I)}\sum_r h_N(r/N(I)) + \frac{4Ob(1)K_0}{N(I)}$$

$$= 1 + \frac{2Ob(1)K_0}{N(I)} + \frac{4Ob(1)K_0}{N(I)}$$

$$= 1 + \frac{6Ob(1)K_0}{N(I)} \ . \blacksquare$$

Remark 7.1.19 *The only result which is not proven mathematically is*

$$|T^{-1}(k^r) - r/N(I)| \leq 4/N(I) \ .$$

It is enough to prove this result in order that property 7.1.20 is mathematically proven. We point out that, by our numerical study, this result seems sure.

From the previous result, we deduce the following property.

Property 7.1.20 *The following equality holds :*

$$P\{\widehat{T}(Y) \in I\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] \ .$$

Proof By the previous equalities,

$$P\{\widehat{T}(Y) \in I\} = \frac{1}{m}\sum_k g_N(k) = \frac{N(I)}{m}\frac{1}{N(I)}\sum_k g_N(k)$$

$$= \frac{N(I)}{m}\Big[1 + \frac{6Ob(1)K_0}{N(I)}\Big] = L(I)\Big[1 + \frac{1}{m}\Big]\Big[1 + \frac{6Ob(1)K_0}{N(I)}\Big]$$

$$= L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] \ . \blacksquare$$

Remark 7.1.21 *One can easily generalize the proof of proposition 7.1.20 to the two-dimensional case:*

Indeed, $P\{X \in I\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big]$. *Now, if p=2,* $N(I_1) = N(I_2)$ *and with a Lipchitz coefficient* K_0',

$$P\big\{(X_1, X_2) \in I_1 \otimes I_2\big\} = L(I_1)L(I_2)\Big[1 + \frac{O(1)K_0'}{N(I_1)}\Big] \ .$$

The previous results can be proved in another manner. In this case, there is a less fine approximation. But the assumptions that we deduce from the digital simulations seem easier to prove. We give this result to show that one is almost sure approximation of $P\{\widehat{T}(Y) \in I\}$.

Then, the following result holds.

Property 7.1.22 *Let* $Y \in F(m)$ *be a random variable which has a probability density function f with respect to* μ_m *such that* $|f(k/m) - f(k'/m)| \leq K_0|k - k'|/m$ *where* K_0 *is not too large.*

Let $I = [c/m, c'/m[\subset [0,1]$. *Assume that* $\widehat{T}^{-1}(I) = \{a_1, a_2, ..., a_{c'-c}\}$. *Then, there exists a constant* c_2 *which is not too large, such that*

$$P\{\widehat{T}(Y) \in I\} = L(I)\Big[1 + \frac{c_2}{\sqrt{c' - c}}\Big] \ .$$

Proof There exists a sequence η_j, j=0,1,2,...,c'-c such that $\eta_{j-1} < a_j < \eta_j$, $\eta_0 = 0$, $\eta_{c'-c} = 1$.
By numerical computations one understand that one can assume $\eta_{j+1} - \eta_j \leq 3/(c' - c)$.
We set $\eta_j - \eta_{j-1} = \frac{1}{c'-c} + e_j$. Then,

$$1 = \int_0^1 f(u)\mu_m(du) = \sum_{j=1}^{c'-c} \int_{\eta_{j-1}}^{\eta_j} f(u)\mu_m(du)$$

$$= \sum_{j=1}^{c'-c} \int_{\eta_{j-1}}^{\eta_j} \big[f(a_j) + Ob(1)K_0 Max(a_j - \eta_{j-1}, \eta_j - a_j)\big]\mu_m(du)$$

$$= \sum_{j=1}^{c'-c} \big[(\eta_j - \eta_{j-1})f(a_j) + (\eta_j - \eta_{j-1})Ob(1)K_0 Max(a_j - \eta_{j-1}, \eta_j - a_j)\big]$$

$$= \sum_{j=1}^{c'-c} \Big(\frac{1}{c' - c} + e_j\Big)f(a_j) + Ob(1)K_0 Max_j\big\{Max(a_j - \eta_{j-1}, \eta_j - a_j)\big\}$$

$$= \frac{1}{c'-c}\sum_{j=1}^{c'-c} f(a_j) + \sum_{j=1}^{c'-c} e_j f(a_j) + Ob(1)K_0\frac{3}{c'-c} \ .$$

In order to have an idea of the value of $\sum_{j=1}^{c'-c} e_j f(a_j)$ one can use the fact that one has a sum and generalize the corollary 8.1.4. Then, one can admit that the e_j's are a realization of a sequence of random variables E_j :

$$\sum_{j=1}^{c'-c} e_j f(a_j) = \sum_{j=1}^{c'-c} E_j(\omega) f(a_j),$$

where $\mathbb{E}(E_j) = 0$. As a matter of fact, because one uses sums, one can assume that E_j is IID (cf corollary 8.1.4).

Then, by theorem page 103 of [19],

$$\frac{\sum_{j=1}^{c'-c} E_j(\omega) f_(a_j)}{\sqrt{\sum_{j=1}^{c'-c} \mathbb{E}(E_j^2) f(a_j)^2}}$$

has asymptotically the N(0,1) distribution.

Because $0 < \eta_j - \eta_{j-1} = \frac{1}{c'-c} + e_j \leq \frac{3}{c'-c}$, $\mathbb{E}(E_j^2) \leq \left[\frac{2}{c'-c}\right]^2$. Moreover, $\sum_{j=1}^{c'-c} f(a_j)^2 \leq (c'-c).K_2$. For example, $\sum_{j=1}^{c'-c} f(a_j)^2 \leq (c'-c)[K_0+1]^2$. Then, one can admit

$$\frac{\sum_{j=1}^{c'-c} E_j(\omega) f(a_j)}{\sqrt{\sum_{j=1}^{c'-c} \mathbb{E}(E_j^2) f(a_j)^2}} = \frac{\sum_{j=1}^{c'-c} E_j(\omega) f(a_j)}{\sigma_E \sqrt{\sum_{j=1}^{c'-c} f(a_j)^2}}$$

$$= \frac{\sum_{j=1}^{c'-c} E_j(\omega) f(a_j)}{\frac{2Ob(1)}{c'-c}\sqrt{Ob(1)K_2(c'-c)}} \ .$$

Then, if b is large enough

$$1 \approx 1 - \Gamma(b) \approx P\Big\{ \frac{\sum_{j=1}^{c'-c} E_j(\omega) f(a_j)}{\frac{2Ob(1)}{c'-c}\sqrt{Ob(1)K_2(c'-c)}} < b \Big\}$$

$$= P\Big\{ \sum_{j=1}^{c'-c} E_j(\omega) f(a_j) < \frac{2b.Ob(1)}{c'-c}\sqrt{K_2(c'-c)} \Big\}$$

$$\leq P\Big\{ \sum_{j=1}^{c'-c} E_j(\omega) f(a_j) < \frac{2b\sqrt{K_2}}{\sqrt{c'-c}} \Big\} \ .$$

Then, normally,

$$1 - \Gamma(b) \le P\Big\{ \sum_{j=1}^{c'-c} E_j(\omega) f(a_j) < \frac{2b\sqrt{K_2}}{\sqrt{c'-c}} \Big\} \le 1 \ .$$

Then,

$$\sum_{j=1}^{c'-c} e_j f(a_j) = \frac{2b\sqrt{K_2} Ob(1)}{\sqrt{c'-c}} = \frac{c_1}{\sqrt{c'-c}} \ ,$$

where c_1 is a constant not too large.

Finally,

$$\begin{aligned} 1 &= \frac{1}{c'-c} \sum_{j=1}^{c'-c} f(a_j) + \sum_{j=1}^{c'-c} e_j f(a_j) + \frac{2O(1)K_0}{c'-c} \\ &= \frac{1}{c'-c} \sum_{j=1}^{c'-c} f(a_j) + \frac{c_1}{\sqrt{c'-c}} + \frac{2O(1)K_0}{c'-c} \\ &= \frac{1}{c'-c} \sum_{j=1}^{c'-c} f(a_j) + \frac{c_1' Ob(1)}{\sqrt{c'-c}} \ . \end{aligned}$$

Then,

$$\sum_{j=1}^{c'-c} f(a_j) = (c'-c) - \frac{c_1'(c'-c)}{\sqrt{c'-c}} \ .$$

Then,

$$\begin{aligned} P\{\widehat{T}(Y) \in I \ \} &= \sum_{j=1}^{c'-c} (1/m) f(a_j) \\ &= \frac{c'-c}{m} \Big[1 - \frac{c_1'}{\sqrt{c'-c}} \Big] \\ &= L(I) \big[1 + Ob(1)/m) \big] \Big[1 - \frac{c_1'}{\sqrt{c'-c}} \Big] \\ &= L(I) \Big[1 + \frac{c_2}{\sqrt{c'-c}} \Big] \ . \ \blacksquare \end{aligned}$$

This result has not the same form as property 7.1.20. It does not correspond either to the preceding numerical studies: too strong increases are obtained.

It is normal because the same techniques are not employed. Here, the e_j are increased: it is a result easier to check than the distribution of the a_j compared to the $r/N(I)$ used in the study of property 7.1.20. Let us notice that the smaller K_0 will be, the more this result is probable. Therefore, in a certain way, this result seems surer.

In fact, it seems well that in all the cases one has at minimum

$$P\{\widehat{T}(Y) \in I\ \} = L(I)\Big[1 + \frac{c_2}{\sqrt{c' - c}}\Big]\ .$$

But, more precisely, as we already said, it seems that

$$P\{\widehat{T}(Y) \in I\ \} = L(I)\Big[1 + \frac{O(1)K_0}{c' - c}\Big]\ .$$

It remain to prove mathematically these results which are yet only almost sure conjectures.

7.1.4 Connexion with the Lipschitz coefficient

We now will detail the function of the Lipschitz coefficient K_N in the method used in section 11.

Indeed, in the section 11, we do not use in the same way the Lipschitz coefficient K_N as in the fundamental theorem 1.

First, we remind the following result.

Lemma 7.1.23 *Let $Pe(y, \sigma^2)$ be the slope in y of the Gaussian curve $h(y) = \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$. Let $K_N = Sup_y\{|Pe(y, \sigma^2)\}|$. Then,*

$$K_N = \frac{1}{\sigma^2}\ \frac{e^{-1/2}}{\sqrt{2\pi}}\ .$$

Proof First, $h'(y) = -(y/\sigma^2)\ \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$.

Then,

$$h"(y) = -(1/\sigma^2)\ \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} + (-y/\sigma^2)(-y/\sigma^2)\frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$$

$$= -(1/\sigma^2)\ \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} + (y^2/\sigma^4)\ \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$$

$$= (1/\sigma^2)\big[y^2/\sigma^2 - 1\big]\frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}} \ ,$$

which is cancelled in $y = \pm\sigma$.

Therefore,

$$K_N = sup\Big\{(|y|/\sigma^2)\ \frac{e^{-y^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}\Big\} = (\sigma/\sigma^2)\ \frac{e^{-\sigma^2/(2\sigma^2)}}{\sqrt{2\pi\sigma^2}}$$

$$= (1/\sigma)\frac{e^{-1/2}}{\sqrt{2\pi\sigma^2}} = \frac{1}{\sigma^2}\ \frac{e^{-1/2}}{\sqrt{2\pi}} \ . \blacksquare$$

Therefore $\frac{1}{\sqrt{\sigma}} = \frac{(2\pi)^{1/8}K_N^{1/4}}{e^{-1/8}}$.

Let K_G be the Lipschitz coefficient associated to G(j) used to build the sequences of random bits $b^1(n')$ (cf section 11.2.5). One can write $K_N = c_G K_G$. In particular, one can choose $c_G = 1$.

Therefore, $\frac{1}{\sqrt{\sigma}} = \frac{(2\pi)^{1/8}c_G^{1/4}K_G^{1/4}}{e^{-1/8}}$.

If the inequality 7.2 is admitted, $|\epsilon_I^G| \leq \frac{0.291\sqrt{N(I)}}{m\sqrt{\sigma}}$. Then, because of the CLT applied to G(n) and to conditional probabilities $P\{G(n) = g \mid G(n+j_s) = g(s)\}$ (cf section 5.7),

$$|\epsilon_I^G| \leq \frac{0.291\sqrt{N(I)}}{m}\ \frac{(2\pi)^{1/8}c_G^{1/4}(K_G)^{1/4}}{e^{-1/8}} \ .$$

Then, one has the following property.

Property 7.1.24 *Let* $C_0 = \frac{0.291(2\pi)^{1/8}c_{Gj}^{1/4}}{e^{-1/8}}$. *Assume that the inequality* 7.2 *holds. Then,*

$$|\epsilon_I^G| \leq \frac{C_0.(K_G)^{1/4}\sqrt{N(I)}}{m} \ .$$

Remark 7.1.25 *The fact that one obtains an equality of the type*

$$|\epsilon_I^G| \leq \frac{C_0 K_G^{1/4}\sqrt{N(I)}}{m}$$

instead of

$$|\epsilon_I^G| \leq \frac{C_1.K_G}{\sqrt{m}}$$

as one obtained for the theorem 1 (or for the proposition 4.1.1 which is equivalent) should not surprise.

In the proof of the proposition 4.1.1, one used increases of bad quality. Here one refined them by using numerical calculations.

But contrary to the proposition 4.1.1 the result is not proved mathematically. However, one can consider that it is true: we chose sufficiently strong increases to be sure of the result. That will be even surer if one takes b larger in the equation 7.1 .

Nonnormal distribution

Logically these results should relate to only case

$$P_Y\{Y=y\} = \frac{10}{m}\ \frac{e^{-[10(y-0.5)]^2/(2\pi\sigma^2)}}{\sqrt{2\pi\sigma^2}}\Big[1+\frac{u(y)}{co}\Big]\ .$$

But increases which we chose are sufficiently strong so that they remain true for the other cases

$$P_Y\{Y=y\} = \frac{h(y)}{m}\big[1+\frac{\eta(y)}{co}\big]\ .$$

Moreover, we carried out various numerical simulations. They show that this result is always checked if one takes the numerical data used for the construction of the sequences of random bits $b^1(n')$ (cf section 11.2.5).

In fact the following assumption is verified per many files that one generally finds on computers : e.g. texts, programmings, musics, etc.

Hypothesis 7.1.3 *We suppose that the following inequality hold.*

$$|\epsilon_I^G| \leq \frac{0.291(2\pi)^{1/8}}{e^{-1/8}}\ \frac{K_G^{1/4}\sqrt{N(I)}}{m}\ .$$

We studied these results only to specify the connection between the fundamental theorem using congruences modulo $d^{p+q}-1$ and congruences of Fibonacci: in the construction of the sequences of random bits $b^1(n')$, we do not use this assumption, but the assumption 7.1.2 which is much weaker.

Numerical increase of K_0

We now calculate the increase of K_0 with the numerical data which we used in the construction of $b^1(n')$ in section 11.2. It is the K_0 associated with the conditional probabilities of D(j) and $E^3(j)$ defined in section 11.2.3.

We studied many numerical examples : for example, for the conditional probability $P\{D(j) = d \mid D(j+j_2) = d_2, ..., D(j+j_p) = d_p\}$ we have the following increases

K_0	3.4	1.2	3.8	15.7	2.2	0.6	4.1	1.5	3.3

For $P\{E^3(j) = e \mid E^3(j+j_2) = e_2, ..., E^3(j+j_p) = e_p\}$, we have the following increases

K_0	0.15	0.9	1.1	0.3	0.2	0.05	0.6	0.2

The K_0 are thus enough small, especially those associated with the $e^3(j)$. It is not surprising because, first we made uniform our data by the applications defined in chapter 8. But this transformation makes them also independent : cf section 8.4.

7.2 Uniform distribution

In the construction of the sequence of random bits $b^1(n')$, we apply the functions of Fibonacci after having applied the XORLT (cf section 11.2.5). One thus has the uniform distribution as limit distribution. Then, in this section we study the model

$$P_Y\{Y = k/m\} = \frac{1}{m}\Big[1 + u_k\Big] \ ,$$

where u_k is a sample of an IID sequence U_k.

We use the following property.

Lemma 7.2.1 *Let I be an interval. We set $\epsilon_I = P_X(I) - N(I)/m$. Let $Y_{Nu} = \sum_{k/m\in I} \frac{U_k}{\sigma_U\sqrt{N(I)}}$. Then,*

$$P\Big\{\frac{m|\epsilon_I|}{\sigma_U\sqrt{N(I)}} \geq b\Big\} = P\{|Y_{Nu}| \geq b\} \ .$$

Proof We have

$$P_X(I) = \sum_{\widehat{T}(k/m)\in I} \frac{1}{m}\ \Big[1 + u_k\Big] \ .$$

Then, $\epsilon_I = P_X(I) - N(I)/m = \frac{N(I)}{m}\sum_k \frac{1}{N(I)}\big[1+u_k\big] - N(I)/m$
$= \frac{N(I)}{m}\Big(\sum_k \frac{u_k}{N(I)}\Big) = \frac{\sqrt{N(I)}\sigma_U}{m}\Big(\sum_k \frac{u_k}{\sigma_U\sqrt{N(I)}}\Big) = \frac{\sqrt{N(I)}\sigma_U}{m}\ Y_{Nu}(\omega)$.

Then, $\frac{m.\epsilon_I}{\sigma_U\sqrt{N(I)}} = \ Y_{Nu}(\omega)$.

Then,

$$P\Big\{\frac{m|\epsilon_I|}{\sigma_U\sqrt{N(I)}} \geq b\Big\} = P\{|Y_{Nu}| \geq b\}. \ \blacksquare$$

Now, Y_{Nu} has asymptotically the normal distribution. Then, the following assumption can be admitted when N(I) is enough big.

Hypothesis 7.2.1 *In this section, we suppose*

$$P\Big\{\frac{m|\epsilon_I|}{\sigma_U\sqrt{N(I)}} \geq b\Big\} \approx \Gamma(b) \ .$$

For example choose b = 10. Because $\Gamma(10) = 1.524 * 10^{-23}$, generally

$$|\epsilon_I| \leq \frac{10\sigma_U\sqrt{N(I)}}{m} \ .$$

As previously, ϵ_I can increase if N(I) increases.

In a more general way, let us consider the set $F(2^q) = \{0/2^q, ..., (2^q-1)/2^q\}$. There are 2^q intervals $I_k = [k/2^q, (k+1)/2^q[$. Then, choose b_q such that $\Gamma(b_q) = 4^{-q}$.

Let $N_{Iel} = Sup_k\Big|card\Big[F(m) \cap [k/2^q, (k+1)/2^q[\Big]\Big|$.

Then, $N_{Iel} = \lfloor m/2^q \rfloor + 1$.

Then, for all interval I_k, generally,

$$P\Big\{\frac{m|\epsilon_{I_k}|}{\sigma_U\sqrt{N_{Iel}}} \geq b_q\Big\} \leq 4^{-q} \ .$$

Because there are only 2^q intervals $I_k = [k/2^q, (k+1)/2^q[$, the following property holds

Property 7.2.2 *Assume that the hypothesis 7.2.1 holds. Then, one can admit*

$$|\epsilon_{I_k}| \leq \frac{b_q\sigma_U\sqrt{N_{Iel}}}{m} \ .$$

Now one wants numerical results usable in the construction of the random sequence $b^1(n')$. However, what is important in this model, it is the variance σ_U^2 of U_k. In fact with our data, one can admit the increase $\sigma_U^2 \leq 1$. It is even maybe too weak.

For example, let $m \geq 1.4 * 10^{31}$. If $\sigma_U \leq 1$ and if q=84, $2^{84} \approx 1.9343 * 10^{25}$, one choose $b_q = 15$, $N_{Iel} \approx 7520$. Then, one can suppose

$$|\epsilon_{I_k}| \leq \frac{b_{84}\ \sqrt{N_{Iel}}}{m} \leq \frac{15\sqrt{7520}}{1.4 * 10^{31}} \leq \frac{9.3}{10^{29}} \ .$$

Chapter 8

To make uniform by the functions of Fibonacci

8.1 Study of the problem

In this section, one will understand how one can make uniform the marginal distributions of a sequence of random variable thanks to the congruence of Fibonacci. It will be understood that this technique can also to make these variables independent

8.1.1 Function T_q of Fibonacci

One reminds the following definition (cf Definition 1.3.5).

Definition 8.1.1 *Let $q, d \in \mathbb{N}^*$. Let T be the congruence of Fibonacci modulo m where m belongs to the Fibonacci sequence.*

We define the Fibonacci functions $T_q^d : F(m) \to F(d^q)$ by $T_q^d = Pr_q^d \circ \widehat{T}$, where $Pr_q^d(z) = \overline{\overline{0, d_1 d_2 .... d_q}}$ when $z = \overline{\overline{0, d_1 d_2 ...}}$ is the writing of z base d.

If d=2, one simplifies T_q^d in T_q and Pr_q^d in Pr_q.

Some notations

In this chapter, the following notations are used.

Notations 8.1.2 *In this chapter 8, $q, d \in \mathbb{N}^*$. Moreover, m is an element of the Fibonacci sequence : $m = fi_{n_0}$. Moreover, $Y_n \in F(m)$ is a sequence of random variables defined on a probability space $(\Omega, \mathcal{A}, P)$ and $X_n = T_q^d(Y_n)$.*

Notations 8.1.3 *We denote by k a element of $F^*(d^q)$: $k \in \{0, 1, ..., d^q - 1\}$.*

Then, $P\{X_n = k/d^q\} = P\{\widehat{T}(Y_n) \in [k/d^q, k'/d^q[\}$.

Notations 8.1.4 *Let $I_k = [k/d^q, (k+1)/d^q[$. We define the interval $[c_k/m, c'_k/m[$ with $c_k, c'_k \in F^*(m)$ by $[c_k/m, c'_k/m[\cap F(m) = [k/d^q, (k+1)/d^q[\cap F(m)$.*

More generally, we denote by I the intervals $I = [k/d^q, k'/d^q[$. Then, we define $[c/m, c'/m[$ with $c, c' \in F^*(m)$ by $[c/m, c'/m[\cap F(m) = [k/d^q, k'/d^q[\cap F(m)$.

Sometimes, by misusing of our notations, we set also $I_k = I = [k/d^q, k'/d^q[= [c/m, c'/m[$.

Then, $P\{X_n = k/d^q\} = P\{\widehat{T}(Y_n) \in [k/d^q, k'/d^q[\} = P\{\widehat{T}(Y_n) \in [c_k/m, c'_k/m[\}$.

Notations 8.1.5 *We set $m = d^Q$ where $Q \in \mathbb{R}_+$.*

Now, the following lemma holds.

Lemma 8.1.1 *With the previous notations, $(c_k - 1)/m < k/d^q \leq c_k/m$ and $(c'_k - 1)/m < (k+1)/d^q \leq c'_k/m$.*

Lemma 8.1.2 *Let $N(I_k)$ be the number of $t/m \in F(m)$ such that $k/d^q \leq t/m < (k+1)/d^q$. Then, $N(I_k) = c'_k - c_k$.*

Let $1/d^q = h_0/m + r$ where $0 \leq r < 1/m$ and $h_0 \in \mathbb{N}$. Then, $N(I_k) = h_0$ or $N(I_k) = h_0 + 1$.

Lemma 8.1.3 *We keep the notation of lemma 8.1.2. Then, $m/d^q = h_0 + e$ where $0 \leq e < 1$.*

Notations 8.1.6 *Let $x_s \in F(m)$. We set $p_{x_s} = P\{\overline{T}(mY_n) = mx_s\} = P\{\widehat{T}(Y_n) = x_s\}$.*

Of course, we can write $P\{T_q^d(Y_n) = k/d^q\} = P\{\widehat{T}(Y_n) \in [k/d^q, k'/d^q[\} = P\{\widehat{T}(Y_n) \in [c_k/m, c'_k/m[\} = \sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}$.

8.1.2 Sequence of real numbers regarded as IID

We show now that about any sequence of real numbers can be regarded as the permutation of an IID sequence. It is thus also the case if $z_n = Z_n(\omega)$ where Z_n is a sequence of unspecified random variables, even if Z_n is deterministic.

Proposition 8.1.1 *Let z_n , $n = 1, 2, ..., n_0$ be a sequence of integers $z_n \in F^*(m)$ such that all the z_n's are different.*

Then, there exists a permutation ϕ such that $z'_n = z_{\phi(n)}$, $n = 1, 2, ..., n_0$ can be regarded as an IID sample having a distribution M_Z .

Proof Let $x_n = x(n)$ be an IID sample with uniform distribution. For any function f, $f(x_n)$ is a priori an IID sample. But it is necessary to be careful: it is better than f is not too complicated. For example $f(x_n)$ can be increasing : it is a classical problem of the samples.

To avoid it, we denote by r_x and r_z the number of order of $x(n)$ and $z(n) = z_n$, respectively : $r_x(n)$ and $r_z(n)$ are the permutations of $\{1, 2, ..., n_0\}$ such that $x_{r_x^{-1}(1)} < x_{r_x^{-1}(2)} < .... < x_{r_x^{-1}(n_0)}$ and $z_{r_z^{-1}(1)} < z_{r_z^{-1}(2)} < .... < z_{r_z^{-1}(n_0)}$.

Then, there exists a continuous function f such that $f(x_{r_x^{-1}(n)}) = z_{r_z^{-1}(n)}$ for $n = 1, 2, ..., n_0$. One can force this function to be smoothest possible

For this function f , $f(x_n)$ can be regarded as an IID sample which has the same law as $f(X_1)$.

Now, $\{f(x_n) \mid n = 1, 2, ..., n_0\} = \{z_n \mid n = 1, 2, ..., n_0\}$. Then, there exists a permutation ϕ such that $z_{\phi(n)} = f(x_n)$ for $n = 1, 2, ..., n_0$. Then, $z_{\phi(n)}$ can be regarded as an IID sample. ■.

We deduce the following propositions.

Corollary 8.1.4 *Let $\sum_{n\in F} z_n$ where $F \subset \{1, 2, ..., n_0\}$.*

Then $\sum_{n\in F} z_n = \sum_{n\in F'} z'_n$, where $z'_n = z_{\psi(n)}$ is an IID sample which has the distribution M_Z and where $F' = \psi(F)$ and $\psi = \phi^{-1}$.

Proof Let $n = \phi(n')$. Then, $n = \phi(n') \in F$ is equivalent to $n' \in \phi^{-1}(F)$. Then, $\sum_{n\in F} z_n = \sum_{n\in F} z_{\phi(n')} = \sum_{n'\in\phi^{-1}(F)} z_{\phi(n')} = \sum_{n'\in\phi^{-1}(F)} z'_{n'}$. ■

Corollary 8.1.5 *Let $\sum_{n\in F'} z'_n$ be a sum where z'_n is an IID sample which has the distribution M_Z.*

Then, for all sets F" $= \psi'(F')$, except a negligible minority, $\sum_{n\in F"} z'_n$ behaves as the sum of an IID sample which has this same distribution M_Z.

Proof If the sequence of random variables $Z'_n = Z_{\psi(n)}$ is IID, any sequence $Z'_{\psi'(n)}$ is IID (where ψ' is a permutation). The sequence $z'_{\psi'(n)}$ will not behave like a sample IID only with one negligible probability given by the traditional laws about IID sample. ■

Now, let us suppose that F is chosen randomly. Then, one can admit that F' is also chosen randomly.

Thus each time one has a sum over a set chosen randomly, one carries out a sum of a sample of an IID sequence of random variable Z'_n. One thus finds the usual limit distributions, for example, $\frac{1}{card(F)} \sum_{n\in F} Z'_n \to L$ where $L = E\{Z'_1\}$, and where Z'_1 has a fixed distribution M_Z.

We deduce the following result.

Corollary 8.1.6 *For almost all the sets F, $\frac{1}{card(F)} \sum_{n\in F} z_n \approx L$ where L does not depend on F.*

Remark 8.1.7 *A sum can thus always be comparable to a sum of an IID sample. But, it is necessary to pay attention to which sum it is in question exactly. For example, let us consider the sequence of the* $x'_n \in F^*(10^8)$ *:* $x'_1, x'_2, ... = 110000, 250000, 780000, 1020000, 111000, ...$ *It will not be transformed into a IID sequence belonging to* $F^*(m)$*, but to a subset of* $F^*(m)$*. If one does not pay attention to that, one can have certain problems during the use of the CLT : cf section 5.4.4.*

8.1.3 To make uniform the marginal distributions

We will try to understand why the functions T_q make uniform the marginal probabilities. For that, we suppose for example that our data are provided by text.

Then, we have $P\{X_n = k/d^q\} = \sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}$.

Now, there is no logical connection between text (translated in numbers) and the distribution of the points of $\{a_1, a_2, ...\} = \widehat{T}^{-1}(I)$ where I is an interval. These two events are logically independent. Indeed, the sequence $\{a_1, a_2, ...\} = \widehat{T}^{-1}(I)$ is built by a specific and relatively simple mathematical application whereas the data y_n are the realization of a succession of random variables Y_n and thus unpredictable in an exact way. Moreover the sequence $\{a_1, a_2, ...\}$ is well distributed in F(m). It is reasonable to think that this set is independent of sequences obtained starting from text. One can thus regard this set as randomly chosen.

That means that $\sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}$ can be regarded as a sum $\sum_{s \in F} p_{x_s}$ where the set F is a Borel set chosen randomly. According to the corollary 8.1.6, that means that, for all k, $(d^q/m) \sum_{x_s \in [k/d^q, (k+1)/d^q[} p_{x_s}$ converges to the same limit L.

One is all the more sure of this result that only a negligible minority of the possible sets F will not check this property: because there is only d^q "k" possible, it is thus enough to choose m enough large compared to d^q.

At last, $\sum_k \sum_{x_s \in [k/d^q, (k+1)/d^q[} p_{x_s} = 1$. Therefore, $\sum_{x_s \in [k/d^q, (k+1)/d^q[} p_{x_s} \approx 1/d^q$.

Now, in order to understand this approximation more, it is necessary to use the CLT: it is what we will do in sections 8.2 and 8.3

Example

Let us suppose that the curve of the probabilities of Y_n have the shape of a normal curve. The $\{a'_1, a'_2, ...\} = \overline{T}^{-1}\big([c, c'[\cap F^*(m)\big)$ are about uniformly distributed.

For example, in figure 8.1, one supposes $card\big([c,c'[\cap F^*(m)\big) = 10$. One understands that $\{a'_1, a'_2, ..., a'_{10}\}$ is about uniformly distributed in [-4,4].

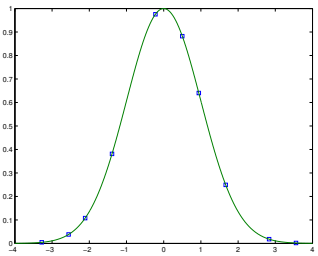

Figure 8.1: Example : Normal curve

Let D be the length of interval [-4,4] on which one studied the normal curve f_N^1 and let N=10 be the number of points. Then, the sum $(D/N)\sum_s f_N^1(y_s)$ over 10 points x_s distributed about uniformly is then close to 1. Therefore, $P\{X_n = k/d^q\} \approx 1/d^q$.

8.1.4 Empirical Probability

Let us be interested with a sample $x_n^* = T_q(y_n)$, $n = 1, 2, ..., n_0$, where all the y_n are distinct. For the empirical associated probabilities P_e, one has

$$P_e\big\{T_q(Y_n) \in [c/m, c'/m[\big\} = \sum_{x_s \in [c/m, c'/m[} pe_{x_s} \ ,$$

where $pe_{x_s} = P_e\big\{\widehat{T}(Y_n) = x_s\big\}$.

Let us set $\widehat{T}^{-1}([c/m, c'/m[) = \{a_1, ..., a_{c'-c}\}$. Then,

$$P_e\big\{T_q(Y_n) \in [c/m, c'/m[\big\} = \frac{1}{n_0}\ Nomb\big\{y_n \in \{a_1, ..., a_{c'-c}\}\big\} \ ,$$

where $Nomb\big(y_n \in \{a_1, ..., a_{c'-c}\}\big)$ is the number of points of the sample y_n such that $y_n \in \{a_1, ..., a_{c'-c}\}$.

One can write $Nomb\big(y_n \in \{a_1, ..., a_{c'-c}\}\big)$ in two ways:

$$Nomb\big\{y_n \in \{a_1, ..., a_{c'-c}\}\big\} = \sum_{n=1}^{n_0} 1_{\{a_1, ..., a_{c'-c}\}}(y_n) = \sum_i 1_{\{y_1, ..., y_{n_0}\}}(a_i) \ .$$

Let us suppose that the y_n's are data resulting from texts: a priori, there is no logical connection between the y_n and $\{a_1, a_2, ...\}$.

For example, choose the text beginning by << Newton's theory of gravitation was soon accepted without question, and it remained unquestioned until the beginning of this century. Then Albert Einstein shook the foundations of physics with the introduction of his Special Theory of Relativity in 1905, and his General Theory of Relativity in 1915 (Here is an example of a thought experiment in special relativity). .....>>. There is no logical connection between the set $\mathcal{J}$ =

[Newton's theory]
[of gravitation was]
[soon accepted wit]
[hout question, and]
[it remained unques]
[tioned until the begin]
[ning of this century.]
[Then Albert Einstein]
[shook the foundation]
[s of physics with the i]
[ntroduction of his Spe]
.....

and the set $\mathcal{H}$ =

[whgkudf ly cuqhjg]
[aamxgusdggbxckmp]
[x;cbkutcc ze xycyc x]
[qtdxucdzlcxy yx vyxy]
[uezuxcuazvxaoaqzq]
[,hqdsgcize cqy bxq]
[a picykhgkkl hfqfqqq]
[ory of Relativity in 190]
[xwtex pez! i yi qy yqhfg]
.....

Thus a priori, the probability that $[Newton'stheory] \in \mathcal{H}$ is approximatively of $card(\mathcal{H})/32^{18}$ (if it is considered that the 26 letters, capital letters and small letters and punctuation belong to a set with 32 elements and if each y_n contains 18 letters or signs or space).

Finally, one can admit that $Nomb\big(y_n \in \{a_1, ..., a_{c'-c}\}\big)$ is the number of points y_n belonging to a set selected randomly of size c'-c.

Therefore, one can admit that the a_i are taken randomly compared to $\{y_1, ..., y_{n_0}\}$. Therefore, one can apply the limit laws. Let $N_A = card\{a_1, a_2, ...\}$ $= c' - c$ and $n_0 = card\{y_n\}$. Then

$$\frac{1}{N_A} \sum_{i=1}^{c'-c} 1_{\{y_1,.....,y_{n_0}\}}(a_i) \rightarrow \frac{n_0}{m} \quad as\ c' - c \rightarrow \infty\ .$$

Now, it was known already that, under simple assumptions,

$$P_e\big\{Y_n \in \{a_1, ..., a_{c'-c}\}\big\} = \frac{1}{n_0} \sum_{n=1}^{n_0} 1_{\{a_1,.....,a_{c'-c}\}}(y_n) \rightarrow \frac{N_A}{m} \quad as\ n_0 \rightarrow \infty\ .$$

This result means that, if the set $\{a_1, ..., a_{c'-c}\}$ is well chosen randomly, in all the cases, $\frac{1}{n_0} \sum_{n=1}^{n_0} 1_{\{a_1,.....,a_{c'-c}\}}(y_n) \rightarrow \frac{N_A}{m}$.

Therefore, that does not depend on the sequence y_n. This result is due to the independence of both events between the sequence y_n and the sequence a_i: since it is admitted that $\{a_1, ..., a_{c'-c}\}$ is chosen randomly, one can apply corollary 8.1.6

Finally, $P_e\big\{X_n \in [k/d^q, (k+1)/d^q[\big\} = P_e\big\{Y_n \in \{a_1, ..., a_{c'-c}\}\big\}$ converges to $1/d^q$ for any k, i.e. one has well the uniformity for empirical measures.

Let us notice that it is possible that the $\{a_1, a_2, ...\}$ have a connection with the empirical probability. It is always possible. But that is likely to occur with a negligible probability as it is the case when one tests if an IID sample is well IID.

Thus it is always possible that, despite all our constructions, the sequence b_i which we finally obtain in section 11 is not IID, but it would be with a negligible probability.

Tests

One checked these results by testing them with the sample provided for each line i by the sequence f(i,j), j=1,2,... : cf section 11.2.4. One has tested the uniformity of the law of the $f(i, j)$, j=1,2,.... (which have been made uniform by the transformation $e^3(j) = m.T_1^m(e^2(j)/m^1)$). All the tests conclude to the uniformity

In figure 8.2 , we have the histogram for the first line f(1,j), j=1,2,.... One can compare this figure with the figure 8.3 which represents the histogram for a pseudo-random sequence of uniform distribution.

8.2 Theoretical probabilities: first method

To simplify the presentation, we consider the case where data y_n, $n = 1, 2, ..., n_0$ are provided by text. There are several possible models representing this text

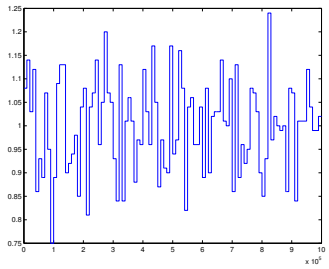

Figure 8.2: Histogram of f(1,j)

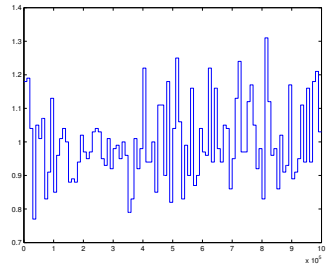

Figure 8.3: Histogram of uniform data

(there is even an infinity of it). I.e. there is several possible random sequences Y_n such as $y_n = Y_n(\omega)$. That thus means that there are several systems of possible probabilities $p_{y^1,y^2,.....,y^{n_0}} = P\{Y_1 = y^1, Y_2 = y^2, ..., Y_{n_0} = y^{n_0}\}$.

In this section, we want to prove that the marginal distributions of the $T_q(Y_n)$ are close to the uniform distribution. We are thus interested in the marginal probability $p_{y^s} = P\{Y_n = y^s\}$ for the $y^s \in \{0/m, 1/m, ..., (m-1)/m\}$.

It is equivalent to study the p_{x_s} : $p_{x_s} = P\{\widehat{T}(Y_n) = x_s\} = P\{Y_n = y^s\} = p_{y^s}$.

Because there are several possible correct models meaning this text, we will study the various possible probabilities p_{x_s}.

One will provide the set of the p_{x_s} of a law of probability. I.e. the set of the possible probabilities p_{x_s} is itself the realization of a probability space.

In this section, one admits the following assumptions.

Definition 8.2.1 *One supposes that the set of possible probabilities of $\widehat{T}(Y_n)$ over F(m) is symbolised by*

$$\left\{ \left(\frac{p'_{1/m}}{\sum_{i=1}^m p'_{i/m}}, \frac{p'_{2/m}}{\sum_{i=1}^m p'_{i/m}}, ..., \frac{p'_{m/m}}{\sum_{i=1}^m p'_{i/m}} \right) \;\middle|\; p'_{i/m} \in \mathbb{R}_+ \right\} .$$

One supposes that $p'_{i/m} = P'_{i/m}(\omega_1)$, $\omega_1 \in \Omega_1$, *where the* $P'_{i/m} \geq 0$ *are IID random variables with a law M, defined on a probability space* $(\Omega_1, \mathcal{A}_1, Proba_1)$.

Then, $p_{x_s} = \frac{p'_{x_s}}{\sum_{i=1}^m p'_{i/m}}$.

Now, we shall need the following notations.

Notations 8.2.2 *One supposes that the* P'_{x_t} *have a fixed distribution M. For all* $b > 0$, *we denote by* $\Gamma_0(b)$ *the function* $\Gamma_0(b) = 2.Max_{n_1 \geq n_2} Proba_1\big\{|S_{n_1}| \geq b\big\}$, *where* $n_2 = min_{k \in F^*(d^q)}(card\{r/m | k/d^q \leq r/m < (k+1)/d^q)\}$, *and where*

$$S_{n_1} = \frac{P'_{1/m} + P'_{2/m} + ... + P'_{n_1/m}}{\sigma_M \sqrt{n_1}} \ .$$

Of course $\Gamma_0(b) \approx 2\Gamma(b)$ as soon as m/d^q is big.

For example, if one wants to choose the probability of X_n randomly, one can admit that M has the uniform distribution on [0,1].

This choice is reasonable. Indeed, according to the proposition 8.1.1, there exists a permutation ψ_0 such as the sequence $p_{\psi_0(y_s)} = p^f_{y_s}$ can be regarded as an IID sample of random variables which have the distribution M_Z.

Therefore, for almost all the permutations ψ, the sequence $p^f_{\psi(y_s)}$ can be regarded as an IID sample of random variables which have the distribution M_Z.

Because T can be regarded as independent of the text (therefore independent of probabilities p_{y_s}) the permutation $\psi_1(y_s) = \psi_0^{-1}(\overline{T}(y_s))$ - associated to $p^f_{y_s} \mapsto p^f_{\psi_1(y_s)} = p_{\psi_0(\psi_1(y_s))} = p_{\psi_0(\psi_0^{-1}(\overline{T}(y_s))} = p_{\overline{T}(y_s)} = p_{x_s}$ - can be regarded as chosen randomly.

Therefore, the sequence $p_{x_s} = p^f_{\psi_1(y_s)}$ can be regarded as an IID sample of random variables which have the distribution M_Z.

Then, we use this model. Now, we study it.

First, the following proposition holds.

Proposition 8.2.1 *We assume* m/d^q *is great enough. Then, the following approximation holds.*

$$P\big\{X_n = k/d^q\big\} \approx 1/d^q \ .$$

Proof Let $c = c_k$ and $c' = c'_k$. We have

$$P\big\{X_n = k/d^q\big\} = P\big\{\widehat{T}(Y_n) \in [c/m, c'/m[\big\} = \sum_{x_s \in [c/m, c'/m[} p_{x_s} \ .$$

Because $p_{x_s} = \frac{p'_{x_s}}{\sum_{i=1}^m p'_{i/m}}$ and because the P'_{x_s} have the distribution M,

$$P\big\{Y_n \in \widehat{T}^{-1}([c/m, c'/m[)\big\} = \frac{c'-c}{m} \ \frac{(1/(c'-c)) \sum_{x_s \in [c_k/m, c'_k/m[} p'_{x_s}}{(1/m) \sum_{i=1}^m p'_{i/m}}$$

which converges in probability to $\frac{c'-c}{m}$ considering that $p'_s s/m$ is an IID sample : $\frac{1}{c'-c}\ \sum_{x_s\in[c/m,c'/m[} p'_{x_s} = \frac{\sum_s P'_{x_s}(\omega_1)}{N(I)}$ where $I=[c/m,c'/m[$, $N(I)=c'-c$. ■

Then,

$$P\big\{X_n = k/d^q\big\} \approx 1/d^q\ .$$

For example, if b=20,

$$Proba_1\Big\{\frac{|\sum_s (P'_{x_s}-E_M)|}{\sqrt{N(I)}\sigma_M} \geq b\Big\} \leq \Gamma_0(b)/2 \approx \Gamma(b) \approx \frac{5.6}{10^{89}}\ ,$$

where $E_M = \int x.M(dx)$ and where σ_M^2 is the variance of M .

Therefore, if b is big enough, one can admit, with a very strong probability

$$\frac{|\sum_s (P'_{x_s}-E_M)|}{\sqrt{N(I)}\sigma_M} \leq b\ .$$

For example, if M has the uniform distribution on [0,1], $|\sum_s (P'_{x_s}-1/2)| \leq \frac{20}{\sqrt{12.N(I)}}$.

More precisely the following lemma holds.

Lemma 8.2.1 *Suppose that m is big enough with respect to N(I) and suppose N(I) big enough. Let us suppose that b and σ_M are not too large. Let us suppose that E_M is not too small. Then, with a probability larger than $1-\Gamma_0(b)$,*

$$P\big\{Y_n \in \widehat{T}^{-1}([c/m,c'/m[)\big\} \approx \frac{c'-c}{m}\ \Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{N(I)}}\Big]\ .$$

Proof With a probability larger than $1-\Gamma_0(b)/2$,

$$\frac{1}{N(I)} \sum_{x_s\in[c/m,c'/m[} p'_{x_s} = E_M + \frac{Ob(1).b\sigma_M}{\sqrt{N(I)}}\ .$$

Then, with a probability larger than $1-\Gamma_0(b)$,

$$P\big\{Y_n \in \widehat{T}^{-1}([c/m,c'/m[)\big\} = \frac{c'-c}{m}\ \frac{[1/(c'-c)]\sum_{x_s\in[c/m,c'/m[} p'_{x_s}}{(1/m)\sum_{i=1}^m p'_{i/m}}$$

$$= \frac{c'-c}{m}\ \frac{E_M + \frac{Ob(1).b\sigma_M}{\sqrt{N(I)}}}{E_M + \frac{Ob(1).b\sigma_M}{\sqrt{m}}} = \frac{c'-c}{m}\ \frac{1+(Ob(1).b\sigma_M)/[E_M\sqrt{N(I)}]}{1+(Ob(1).b\sigma_M)/[E_M\sqrt{m}]}$$

$$= \frac{c'-c}{m}\ \Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{N(I)}}+\frac{O(1)}{\sqrt{m.N(I)}}+\frac{Ob(1).b\sigma_M}{E_M\sqrt{m}}\Big]$$
$$\approx \frac{c'-c}{m}\ \Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{N(I)}}\Big]\ .\blacksquare$$

Then, generally,

$$P\big\{X_n = k/d^q\big\} \approx \frac{1}{d^q}\ \Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{N(I_k)}}\Big]\ .$$

More precisely the following proposition holds.

Proposition 8.2.2 *Let us suppose that m is enough large compared to d^q and $N(I_k)$. Let us suppose that b and σ_M are not too large. Let us suppose that E_M is not too small. Then, with a probability larger than $1-\Gamma_0(b)$,*

$$P\big\{X_n = k/d^q\big\} = \frac{1}{d^q}\ \Big[1+\frac{Ob(1).2b\sigma_M}{E_M\sqrt{N(I_k)}}\Big]\ .$$

Proof In this case, one can choose $[c/m, c'/m[= I_k$. By the proof of lemma 8.2.1,

$$P\big\{Y_n \in \widehat{T}^{-1}([c/m, c'/m[)\big\} = \frac{c'-c}{m}\ \Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{N(I_k)}}+\frac{O(1)}{\sqrt{m}}\Big]\ .$$

By lemma 8.1.3 and 8.1.2, $h_0/m \leq 1/d^q \leq (h_0+1)/m$. By lemma 8.1.2, $N(I_k) = c'-c = h_0$ or $N(I_k) = h_0+1$. Then,

$$\frac{c'-c}{m} = \frac{1}{d^q}+\frac{Ob(1)}{m}\ .$$

Then,

$$P\big\{Y_n \in \widehat{T}^{-1}([c/m, c'/m[)\big\} = \frac{1}{d^q}\ \Big[1+\frac{Ob(1).2b\sigma_M}{E_M\sqrt{N(I_k)}}\Big]\ .\blacksquare$$

For example, if d=10, b=20, if $m \geq 10^{Q'}$, Q'=100, and q=50 : $N(I_k) \geq 10^{50}$. Then, if M is the uniform distribution on [0,1], one can admit with a probability very close to 1, that,

$$P\{X_n = k/d^q\} \approx \frac{1}{10^{50}}\ \Big[1+\frac{80.Ob(1)}{\sqrt{12}.10^{25}}\Big] = \frac{1}{10^{50}}\ \Big[1+\frac{23.2}{10^{25}}\Big]\ .$$

This result is true for a n and a k fixed. But, one can have a result for all the k and all the X_n, n=1,2,., N. Indeed one has the following lemma.

Lemma 8.2.2 *One supposes $m = d^Q$ big. Let us suppose m enough large compared to d^q and suppose d^q large enough. Let us suppose that $b << d^{Q-q}$ and that M is the uniform law. One supposes that the size of sequence X_n is N. Let $P_{X_n}(k) = P\{X_n = k/d^q\}$. Then,*

$$Proba_1\left\{\bigcap_{n,k}\left\{\left|P_{X_n}(k) - \frac{1}{d^q}\right| \leq \frac{4b}{\sqrt{12.d^{Q+q}}}\right\}\right\} \geq 1 - Nd^q\Gamma_0(b) \ .$$

This lemma is proved by the same way as lemma 8.2.3.

One keep the same example as previously and one takes a sample x_n^* of size 10^8. Then, because, $\Gamma(20) = \frac{5.6}{10^{89}}$ approximately,

$$Proba_1\left\{\bigcap_{n,k}\left\{\left|P_n^X(k) - \frac{1}{d^q}\right| \leq \frac{2b}{\sqrt{12.d^{Q+q}}}\right\}\right\} = 1 - \frac{5.6 * 10^{50} * 10^8}{10^{89}} \geq 1 - \frac{5.6}{10^{31}} \ .$$

Moreover, $P\{X_n = k/d^q\} = P_{X_n}(k/d^q) \approx \frac{1}{10^{50}}\left[1 + \frac{23.2.Ob(1)}{10^{25}}\right]$. Therefore, for a sample of x_n^* of size 10^8, the X_n have a distribution extremely close to the uniform distribution for a vast majority of the possible models.

However, if for example, the probability $\Gamma(20) = 5.6/10^{89}$ were still too large, it would be enough to choose b even larger. Thus, for $b = 40$,

$$Proba_1\left\{\frac{|\sum_s (P_{x_s} - E_M)|}{\sqrt{N(I)}\sigma_M} \geq 40\right\} \approx \frac{1}{10^{340}} \ .$$

In this case, for q=50, $P\{X_n = k/d^q\} \approx \frac{1}{10^{50}}\left[1 + \frac{46.4}{10^{25}}\right]$.

For q=48 : $P\{X_n = k/d^q\} \approx \frac{1}{10^{48}}\left[1 + \frac{46.4}{10^{26}}\right]$.

Then, one notices that if the marginal probabilities are taken randomly, these marginal probabilities are close to the uniform law for almost all the models possible of probability.

8.2.1 Case of Borel sets

One has just understood that the probabilities of the $X_n = k/d^q$ are very close to the uniform probability. One studies now the probability $X_n \in Bo$ where Bo is Borel set of $\{0/d^q, 0/d^q, ..., (d^q - 1)/d^q\}$: we set $Bo = \cup_{k\in\Theta}\{k/d^q\}$, where $\Theta \subset \{0, 1, ..., d^q - 1\}$. Let $K_\Theta = card(\Theta)$.

One can also consider that $Bo = \cup_{k\in\Theta}\{[k/d^q, (k+1)/d^q[\}$. Then, $L(Bo) = \frac{K_\Theta}{d^q}$.

Here we write $I_k = [c_k/m, c'_k/m[\, : \, P\{X_n = k/d^q\} = \sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}$. Donc, $P\{X_n \in Bo\} = \sum_{x_s \in \cup_{k\in\Theta}[c_k/m, c'_k/m[} p_{x_s}$. Let $N(Bo)$ be the number of $t/m \in \cup_{k\in\Theta}[c_k/m, c'_k/m[$. Then, by the CLT,

$$Proba_1\Big\{ \frac{|\sum_{x_s \in \cup_{k\in\Theta}[c_k/m, c'_k/m[}(P'_{x_s} - E_M)|}{\sqrt{N(Bo)}\sigma_M} \geq b \Big\} \approx \Gamma(b) \; .$$

Then, one has the following property.

Proposition 8.2.3 *Let us suppose m enough large compared to d^q and $N(I_k)$. Let us suppose that b and σ_M are not too large. Let us suppose that E_M is not too small. Then, with a probability larger than $1 - \Gamma_0(b)$,*

$$P\big\{X_n \in Bo\big\} = L(Bo) \; \Big[1 + \frac{Ob(1).2b\sigma_M}{E_M d^{(Q-q)/2}}\Big] \; .$$

Proof Clearly $K_\Theta \leq d^q$. By the proof of proposition 8.2.2,

$$P\big\{X_n \in Bo\big\} = \sum_{k\in\Theta} P\big\{Y_n \in \widehat{T}^{-1}([c_k/m, c'_k/m[)\big\}$$

$$= \Big[\frac{\sum_{k\in\Theta}(c'_k - c_k)}{m}\Big] \; \Big[1 + \frac{Ob(1).b\sigma_M}{\sqrt{N(I_k)}} + \frac{O(1)}{\sqrt{m}}\Big]$$

$$= \Big[\sum_{k\in\Theta}\Big(\frac{1}{d^q} + \frac{Ob(1)}{m}\Big)\Big] \; \Big[1 + \frac{Ob(1).b\sigma_M}{E_M\sqrt{N(I_k)}} + \frac{O(1)}{\sqrt{m}}\Big]$$

$$= \Big[\frac{K_\Theta}{d^q} + \frac{Ob(1)K_\Theta}{m}\Big] \; \Big[1 + \frac{Ob(1).b\sigma_M}{E_M\sqrt{Inf(N(I_k))}} + \frac{O(1)}{\sqrt{m}}\Big]$$

$$= \Big[\frac{K_\Theta}{d^q} + \frac{Ob(1)d^q}{d^Q}\Big] \; \Big[1 + \frac{O(1).b\sigma_M}{E_M d^{(Q-q)/2}} + \frac{O(1)}{\sqrt{m}}\Big]$$

$$= \frac{K_\Theta}{d^q} \; \Big[1 + \frac{Ob(1).2b\sigma_M}{E_M d^{(Q-q)/2}}\Big]$$

$$= L(Bo) \; \Big[1 + \frac{Ob(1).2b\sigma_M}{E_M d^{(Q-q)/2}}\Big] \; .\blacksquare$$

For example, for $b = 40$, $d = 10$, $m \geq 10^{100}$, $L(Bo) = 1/10$, $q = 50$, if M is the uniform distribution, then, with a probability larger than $1 - \frac{1}{10^{340}}$ in the set of the probabilities,

$$P\{X_n \in Bo\} \approx \frac{1}{10} \; \Big[1 + \frac{4.64}{10^{49}}\Big] \; .$$

It is thus understood that one still has completely satisfactory results for all the Borel sets.

A priori, this result is normal: there is always no logical connection between the set of the type $\mathcal{J} =$

[Newton's theory]
[of gravitation was]
[soon accepted wit]
[hout question, and]
[it remained unques]
.....

and the union of the set of the type $\mathcal{H}_s =$

[whgkudf ly cuqhjg]
[aamxgusdggbxckmp]
[a picykhgkkl hfqfqqq]
[ory of Relativity in 190]
[xwtex pez! i yi qy yqhfg]
.....

However, it is not exact. Indeed, there exist some connections : if one chooses Borel sets built from the $\mathcal{H}_k = T_q^{-1}(k/d^q)$ containing more parts of text than the others ones, one will obtain Borel sets which have measure enough different from uniform measure.

It is just the same for the empirical probabilities when the sample size n_0 is much smaller than d^q: in general, the $T_q^{-1}(k/d^q)$ contains 0 or 1 points.

But in the case of empirical probability, one knows that one can always find Borel sets of nonuniform measure however associated with an IID sample.

It is not astonishing when one considers the number of possible Borel sets : it is equal to the number of subsets of $F(d^q)$. A priori, there is
1 subset of $F(d^q)$ which is empty
d^q subsets of $F(d^q)$ containing 1 element
$C_{d^q}^2$ subsets of $F(d^q)$ containing 2 element
.....
Altogether, there is $\sum_t C_{d^q}^t = 2^{d^q}$ possible subsets of $F(d^q)$.

Therefore, one can prove the following lemma.

Lemma 8.2.3 *Let us suppose m enough large compared to d^q and $N(I_k)$. Let us suppose that $b << d^{Q-q}$ and that M is the uniform law. Let $1 \leq n \leq N$. Let*

$P_{X_n}(Bo) = P\big\{X_n \in Bo\big\}$. *Then,*

$$Proba_1\Big\{ \bigcap_{n,Bo} \Big\{ \big|P_{X_n}(Bo) - L(Bo)\big| \leq \ L(Bo)\frac{4b}{\sqrt{12.d^{Q-q}}} \Big\} \Big\} \geq 1 - N2^{d^q}\Gamma_0(b) \ . \tag{8.1}$$

Proof By proposition 8.2.3, for all Borel set Bo, we have thus

$$Proba_1\Big\{ \big|P_{X_n}(Bo) - L(Bo)\big| > \ L(Bo) \ \frac{4b.}{\sqrt{12.d^{Q-q}}} \Big\} \leq \Gamma_0(b) \ .$$

For $h \in \{0, 1, ..., d^q\}$, there are $C_{d^q}^h$ Borel sets Bo_h such that $card(Bo_h) = h$. Moreover, there is at the most N possible "n". Then

$$Proba_1\Bigg\{ \bigcap_{n,Bo} \Big\{ \big|P_{X_n}(Bo) - L(Bo)\big| \leq L(Bo) \ \frac{4b}{\sqrt{12.d^{Q-q}}} \Big\} \Bigg\}$$

$$= Proba_1\Bigg\{ \bigcap_{n,h,Bo_h} \Big\{ \big|P_{X_n}(Bo_h) - L(Bo_h)\big| \leq L(Bo_h) \ \frac{4b}{\sqrt{12.d^{Q-q}}} \Big\} \Bigg\}$$

$$= 1 - Proba_1\Bigg\{ \complement \bigcap_{n,h,Bo_h} \Big\{ \big|P_{X_n}(Bo_h) - L(Bo_h)\big| \leq L(Bo_h) \ \frac{4b}{\sqrt{12.d^{Q-q}}} \Big\} \Bigg\}$$

$$= 1 - Proba_1\Bigg\{ \bigcup_{n,h,Bo_h} \Big\{ \big|P_{X_n}(Bo_h) - L(Bo_h)\big| > L(Bo_h) \ \frac{4b}{\sqrt{12.d^{Q-q}}} \Big\} \Bigg\}$$

$$\geq 1 - \sum_{n,h,Bo_h} Proba_1\Bigg\{ \Big\{ \big|P_{X_n}(Bo) - L(Bo_h)\big| > L(Bo_h) \ \frac{4b}{\sqrt{12.d^{Q-q}}} \Big\} \Bigg\}$$

$$= 1 - \sum_h \sum_{n,Bo_h} \Gamma_0(b)$$

$$= 1 - \sum_h NC_{d^q}^h \Gamma_0(b)$$

$$= 1 - N2^{d^q}\Gamma_0(b) \ . \ \blacksquare$$

In order that the probability can be uniform for all Borel sets, certain conditions will thus have to be imposed.

Property 8.2.4 *In order that the inequality 8.1 is useful,* $b = K_{10}d^{q/2}$ *should be imposed, where the* K_t *'s are constant and* $K_{10} > 1$.

Proof Indeed, $(1/2)\Gamma_0(b) \approx \Gamma(b) \approx \frac{\sqrt{2}}{\sqrt{\pi}b}\ e^{-b^2/2}$ when b is big (cf (28)' page 56 [44]). Then,

$$Proba_1\left\{ \bigcap_{n,Bo} \left\{ \left| P_{X_n}(Bo) - L(Bo) \right| \leq L(Bo)\ \frac{4b}{\sqrt{12.d^{Q-q}}} \right\} \right\}$$

$$\geq 1 - N2^{d^q}\Gamma_0(b) \approx 1 - N2^{d^q}\frac{\sqrt{2}}{\sqrt{\pi}b}\ e^{-b^2/2}\ .$$

In order that this inequality are useful, It is necessary that

$$N2^{d^q}\frac{\sqrt{2}}{\sqrt{\pi}b}\ e^{-b^2/2} \leq 1\ .$$

One will thus impose that

$$e^{Log(2)d^q}\ \leq K_7 e^{b^2/2}\ .$$

One will thus impose that

$$2Log(2)d^q\ \leq K_8 b^2\ .$$

One will thus impose that

$$\sqrt{2Log(2)}d^{q/2}\ \leq K_9 b\ .$$

One will thus impose that

$$b = K_{10}d^{q/2}\ . \blacksquare$$

Thus,

$$N2^{d^q}\frac{\sqrt{2}}{\sqrt{\pi}b}\ e^{-b^2/2} = Ne^{log(2)d^q}\frac{\sqrt{2}}{\sqrt{\pi K_{10}}d^{q/4}}\ e^{-(K_{10}^2/2)d^q}$$

$$= N\frac{\sqrt{2}}{\sqrt{\pi K_{10}}d^{q/4}}\ e^{-(K_{10}^2/2-log(2))d^q}\ .$$

For example, if $d^q \geq 10^6$, $K_{10} = 2$, $N\frac{\sqrt{2}}{\sqrt{\pi K_{10}}d^{q/4}}\ e^{-(K_{10}^2/2-log(2))d^q} = \epsilon_8$ where $\epsilon_8 \approx 0$.

Now, b has to be not too large.

Property 8.2.5 *One supposes b large. In order that the inequality 8.1 is useful, one can impose* $2q < Q$.

Proof Choose $b = K_{10} d^{q/2}$. There are the following relations

$$Proba_1 \left\{ \bigcap_{n,Bo} \left\{ \left| P_{X_n}(Bo) - L(Bo) \right| \leq L(Bo) \ \frac{4b}{\sqrt{12}.d^{(Q-q)}} \right\} \right\}$$

$$= Proba_1 \left\{ \bigcap_{nBo} \left\{ \left| P_{X_n}(Bo) - L(Bo) \right| \leq L(Bo) \ \frac{4K_{10} d^{q/2}}{\sqrt{12}.d^{(Q-q)/2}} \right\} \right\}$$

$$= Proba_1 \left\{ \bigcap_{nBo} \left\{ \left| P_{X_n}(Bo) - L(Bo) \right| \leq L(Bo) \ \frac{4K_{10}}{\sqrt{12}.d^{(Q-2q)/2}} \right\} \right\}$$

$$\geq 1 - N 2^{d^q} \Gamma_0(b) \ .$$

In order that this inequality is useful, one can impose

$$\frac{4K_{10}}{\sqrt{12}.d^{(Q-2q)/2}} \leq 1 \ .$$

One will thus impose that

$$2q < Q \ . \ \blacksquare$$

One must thus choose $2q < Q$ and $b = K_{10} d^{q/2}$ so that $\frac{4K_{10}}{\sqrt{12}.d^{(Q-2q)/2}} << 1$ and $2N 2^{d^q} \Gamma(b) \approx 0$ pour $b = O(d^{q/2})$.

Are we obliged to impose this condition to have x_n IID? The answer seems not. Indeed, let us suppose that the sample size n_0 checks $n_0 << d^q$: in this case, the sets $\{y_n \mid T_q(y_n) = k/d^q\}$ contain 0 points y_n in general. Some ones contain 1 point. There are obligatorily important breakdowns of empirical independence for some Borel sets. Then, it is not annoying if that results also in breakdowns of theoretical independence.

8.2.2 Well distributed measure

The previous results were obtained by considering that one chose randomly a measure in the set of the possible probabilities.

But, so that one can associated a probability chosen randomly with a sample, one needs that the probabilities of the X_n are not concentrated nearly a small number of points. If not, this choice is not correct. Indeed, in this case, the majority of the p_{x_s} will be equal to 0 and could not thus be regarded as chosen randomly, at least not according to a uniform law for example.

That means that it is necessary that there is not some $P\{X_n = k/m\}$ too large.

Now let us take into account the empirical probability.

For example, if one chooses data resulting from texts and if m is not too large, (for example $m = 32^7$), certain words of 7 letters can appear several times in the text, for example "theorem" if mathematical text is used. In this case, one will have to consider that there is a too strong probability in certain points.

The previous model of the section 8.2 is not then appropriate.

That thus should be avoided. For that, first, it is necessary that one has a sample of y_n which all are different.

A simple method for that is to choose m large enough, compared to the size N of the sample, for example $m \geq 10^{100}$, $N = 10^7$. Indeed, in this case, it is not much possible that there is two y_n and $y_{n'}$ which are equal. If not, that would mean that two sequences of approximately 70 letters could be identical in a set of N terms. There is a quasi negligible number of chances that such an event occurs.

Now, it is necessary that *a priori* all the possible values of F(m) can exist in a sample.

It is reasonably the case when one adds modulo m a pseudo-random sequence g_n of period m : $my'_n = \overline{g_n + my_n}$. Normally any value k/m has a chance reasonable to be realized a priori. There is no reason that can not occur. Moreover, a priori all k/m has about as much chance to be an image than any other k'/m.

With this method, there is very little chance that there is $y'_n = y'_{n'}$ when $n \neq n'$. In particular, an empirical sample will not be associated a priori with a probability concentrated in a small number of points.

Moreover, by using this way, there is a first standardization of the marginal laws.

8.2.3 Counterexample

One has just understood that, for the set of the probabilities chosen randomly with the uniform law, one can consider that

$$P\{X_n \in Bo\} = L(Bo)\Big[1 + \frac{4b.Ob(1)}{\sqrt{12.N(Bo)}}\Big] .$$

We made this study in the section 8.2 without supposing that T is the congruence of Fibonacci. Then, this result is true for any function $Pr_q \circ Perm$ where Perm is a permutation of F(m).

Let us suppose that $Perm$ is the identity and that the curve of the probabilities of Y_n have the shape of a normal curve. Then,

$$\{X_n = k/d^q\} = \{Y_n \in T_q^{-1}(I)\} = \left\{Y_n \in \{c_k/m, (c_k+1)/m, ..., (c'_k-1)/m\}\ \right\},$$

where $I = [c_k/m, c'_k/m[$.

For example, suppose $d^q = 10^{50}$. Then, $P\{Y_n \in T_q^{-1}(I)\}$ depends on c_k and varies considerably according to c_k. Therefore there is no a uniform probability.

Then, the form of the probability intervenes. In the case of a curve of normal law, to study the probabilities as if they were randomly selected as above is not appropriate.

Let us notice that, contrary to the congruence of Fibonacci, there is a dependence between T and text. If y_n means an extract of texts, $T(y_n)$ means the same extract of text. Therefore, $T_q^{-1}(k/m)$ is a set of extract of texts. This counterexample is thus not valid in the case of a congruence of Fibonacci.

One would obtain the same type of result (that with $Perm = Id$), if one used congruences $T(x) \equiv d^p x$ modulo $d^{2p} - 1$ defined in proposition 4.1.1, considering in this case, T inverts the first decimals with the last ones : cf section 4.1.2. Then, $T_q^{-1}(k/m)$ means also a set of extracts of text.

Also let us notice that, a priori, this counterexample does not use probabilities associated to numbers built with text according to the method described in section 11.1.2. Indeed, in this case, one can consider that the probabilities associated can be regarded as chosen randomly.

8.2.4 Validity of the previous system

Let us suppose again that one is in the case where T is the congruence of Fibonnacci and that our data are provided by text y_n to which one adds a pseudo-random generator g_n : $my'_n = \overline{g_n + my_n}$.

We will understand that to use the probability space $(\Omega_1, \mathcal{P}_1, Proba_1)$ defined in the definition 8.2.1 is a reasonable assumption.

First reason of validity

Let us suppose that the probability $p'_{x_s} = P'_{x_s}(\omega_1)$ corresponding to these y'_n is not extracted from an IID sequence P'_{x_s}, but of a sequence of random variables which have a certain law.

One has always

$$P\{X_n = k/d^q\} = \sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}\ .$$

One knows that there is no logical connection between text and the points of the $\{a_1, a_2, ...\}$. Therefore, $\sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}$ can be regarded as a sum $\sum_{s \in F} p_{x_s}$ where the set F is a Borel set chosen randomly. According to the results of the section 8.1.1, the sum $\sum_{x_s \in [c_k/m, c'_k/m[} p_{x_s}$ is also the sum of an IID sample which has a distribution Q.

Therefore, $\sum_{x_s \in [c_k/m, d_k/m[} P_{x_s} \ \rightarrow \ E_Q$.

One thus has indeed sums of an IID sequence P'_{x_s} with a law Q.

Therefore, one can indeed apply the results of the section 8.2 : that validates this system.

Second reason of validity

The results of section 8.2 were obtained in considering that one randomly chose a measure in the set of the possible probabilities.

Now, if one wants to have a complete idea of the marginal probability for a sample, it is necessary that the size of this one is sufficiently large. Without that, one does not know enough this probability.

Therefore, one can imagine that one complements the sample y'_n by a virtual sample of big size which constitutes a natural continuation of the y'_n. For this virtual sample, there is nothing a priori which prevents from having all the points of $\{0, 1, ..., m-1\}$ even several times.

For example, for a = 1346269 m=2178309 , d= 10, q = 4, one has used samples of size 30m of data y_n to which one added a pseudo-random sequence $g_n \in F(m)$ of period larger than 30m: $my'_n = \overline{g_n + my_n}$. The number of times n_{x_s} where each $x_s \in F(m)$ was equal to a y'_n/m : $y'_n/m = x_s$ checked $2 \leq n_{x_s} \leq 10$.

Now it is known that, if one makes uniform with a pseudo-random generator, one can logically admit that this probability can be enough close to uniform distribution.

Then, a correct model is that where the probabilities p_{x_s} are not too different from each other: it is generally the case when the probabilities of Y_n are taken randomly.

Thus, one can admit that one is in this case here.

Conclusion

Thus, the model chosen for the probabilities of the section 8.2 seems valid for the two previous reasons.

Thus one can consider that the distribution of X_n is very close to the uniform distribution : that takes place with a probability $Proba_1$ extremely near to"1.

8.3 Theoretical probability: second method

8.3.1 Introduction

In this section, one will not suppose that the probabilities are chosen randomly with a certain distribution. This one will remain fixed. They are the a_i that one will choose randomly where $I = I_k$ and where $\widehat{T}^{-1}(I) = \{a_1, a_2, ..., a_{c'-c}\}$. Therefore, one supposes the p_{x_s} fixed.

Moreover, let us suppose that the y_n's mean text and that g_n is a pseudo-random generator.

At last, the following notations are always used.

Notations 8.3.1 *We suppose that $I = [c/m, c'/m[$ with $c = c_k$ and $c' = c'_k$. We set $\widehat{T}^{-1}([c/m, c'/m[) = \{a_1, a_2, ..., a_{c'-c}\}$. We denote by p_{x_s} the probabilities of the x_i's for the model associated to $my'_n = \overline{g_n + my_n}$.*

Moreover, one suppose that $c' - c$ is small compared to m.

Then, one can write p_{a_i} by the following way.

Lemma 8.3.1 *We have the equality $p_{a_i} = \sum_{s=0}^{m-1} p_{x_s} 1_{x_s}(a_i)$.*

Distribution of the sets chosen randomly

Now, if one chooses a set F randomly, that means that the set $F = \{f_1, ..., f_p\}$ has as much chance to be selected than the set $F' = \{f'_1, ..., f'_p\}$:

$$Proba\big\{F = \{f_1, ..., f_p\}\big\} = Proba\big\{F' = \{f'_1, ..., f'_p\}\big\}$$

for all p, for all F and F'.

In particular, for p=1, $Proba\big\{F = \{f_1\}\big\} = Proba\big\{F' = \{f'_1\}\big\} = C_7$ where C_7 is a constant. At first, one thus chooses f_1 randomly, therefore with a uniform probability. Then, one chooses f_2 randomly, i.e. independent and so on.

In fact, it would be necessary to choose f_2 being able to be equal to f_1. But if p is small compared to m and if m is large, one can regard such a possibility as negligible. It is what one will do in this section.

Therefore, the f_i's forms an independent sequence. Finally, one can consider that the f_i are an IID sample with uniform distribution.

One can thus admit that to choose a set $\{a_1, a_2, ..., a_{c'-c}\} \subset F(m)$ of size $p = c' - c$ randomly, it is to choose $a_1, ..., a_p$ with a uniform distribution and independent for each $a_i \in F(m)$. Thus one can admit the following assumption.

Hypothesis 8.3.1 *For all i, one supposes $a_i = A_i(\omega_2)$ where $\omega_2 \in \Omega_2$ and where A_i is an IID sequence of random variables with the uniform distribution on F(m), defined on a probability space $(\Omega_2, \mathcal{A}_2, Proba_2)$.*

Expectation and variance

With the previous notations, $P\{\widehat{T}^{-1}(I)\} = \sum_{i=1}^{c'-c} p_{a_i}$. Therefore, the following lemma holds.

Lemma 8.3.2 *Let $N_A = c' - c$. The following equalities hold:*
$\mathbb{E}\{p_{A_i}\} = 1/m$,
$\mathbb{E}\{p_{A_i}^2\} = (1/m)\sum_{s=1}^{m} p_{x_s}^2$,
$Var = \mathbb{E}\Big\{\Big(\sum_{i=1}^{c'-c}[p_{A_i} - 1/m]\Big)^2\Big\} = (N_A/m)\sum_{s=1}^{m}[p_{x_s}(p_{x_s} - 1/m)]$.

Proof We have

$$\mathbb{E}\{p_{A_i}\} = \mathbb{E}\Big\{\sum_{s=1}^{m} p_{x_s} 1_{x_s}(A_i)\Big\} = \sum_{s=1}^{m} p_{x_s}\mathbb{E}\Big\{1_{x_s}(A_i)\Big\}$$

$$= \sum_{s=1}^{m} p_{x_s}\int 1_{x_s}(u)\mu_m(du) = \sum_{s=1}^{m} p_{x_s}(1/m) = 1/m \ .$$

Moreover,

$$\mathbb{E}\{p_{A_i}^2\} = \mathbb{E}\Big\{\Big[\sum_{s=1}^{m} p_{x_s} 1_{x_s}(A_i)\Big]^2\Big\} = \mathbb{E}\Big\{\sum_{s,s'} p_{x_s} p_{x_{s'}} 1_{x_s}(A_i) 1_{x_{s'}}(A_i)\Big\}$$

$$= \sum_{s=1}^{m}\mathbb{E}\Big\{p_{x_s}^2 1_{x_s}(A_i)\Big\} = (1/m)\sum_{s=1}^{m} p_{x_s}^2 \ .$$

Now the A_i's are independent. Then,

$$\mathbb{E}\Big\{\Big(\sum_{i=1}^{c'-c}[p_{A_i} - 1/m]\Big)^2\Big\} = \sum_{i=1}^{c'-c}\mathbb{E}\Big\{\Big(p_{A_i} - 1/m\Big)^2\Big\} = \sum_{i=1}^{c'-c}\Big(\mathbb{E}\{p_{A_i}^2\} - 1/m^2\Big)$$

$$= (N_A/m)\sum_{s=1}^{m} p_{x_s}^2 - N_A/m^2$$

$$= (N_A/m)\Big[\sum_{s=1}^{m} p_{x_s}^2 - 1/m\Big] = (N_A/m)\sum_{s=1}^{m}[p_{x_s}(p_{x_s} - 1/m)] \ . \ \blacksquare$$

8.3.2 Study of the Central Limit Theorem

CLT for double independent sequences

Having an IID sequence a_i, it is easier to obtain results with the Central Limit Theorem. But, to have infinite limits with sequences of a_i smaller than m, one will have to employ the results on the double sequences. Indeed, one cannot choose samples a_i increasingly large when m is fixed. One thus will make increase m and, thus the number of a_i.

The aim is to understand by using the CLT that $P\{\widehat{T}^{-1}(I)\} - 1/d^q$ is small

One thus will suppose that one has IID sequences $a_i^n \in F(m_n)$, $i = 1, ..., m_n$ and that one makes to tend $m = m_n$ to infinity. As a matter of fact, one studies the samples of $a_i^n \in F(m_n)$, all distinct, of size N_{A^n}. Then, one uses sequences of probabilities $p_{x_s}^n = p_{x_s}$, $s \in \{0, 1, ..., m_n\}$.

Moreover what interests us, it is to apply the functions T_q when q is fixed: i.e. one admits that $L_{A^n} = N_{A^n}/m_n = 1/d^q + Ob(1)/m_n$ remains about constant.

On the other hand, one will admit also the following assumption.

Hypothesis 8.3.2 *In this section, one supposes that $m_n p_{x_s} \leq m_n^\beta$ where $\beta < 1/4$. We set $n_{x_s} = m_n p_{x_s}$.*

This assumption means that the associated probability is not concentrated nearly a small number of points.

Is what this assumption is correct?

A priori, that could be not always the case. Thus, for an unspecified sample, it would be always possible that a sample of elements all distinct with size 10^{10} corresponds for example to a uniform probability concentrated in 10^{30} points. Such a model would be a priori a correct model.

But, other assumptions intervene: it would not be a logical model for our model $my'_n = \overline{g_n + my_n}$. Indeed, any point $k \in \{0, 1, ..., m-1\}$ has a reasonable chance to be a my'_n. One must thus reject this possibility.

Moreover, to suppose that $n_{x_s} \geq m_n^{1/4}$, that means that as soon as m_n is large, there is a point which is likely much more to appear in a sequence my'_n. It is not the case.

It is even known that, normally the probability of y'_n is close to a uniform probability (cf above). However, it is always possible that an estimate of the density can be slightly different from the uniform distribution. That means that there remain unknown properties about this law. We use the function of Fibonacci to eliminate these unknown properties.

Therefore, the assumption $n_{x_s} \leq m_n^{1/4}$ is logical with the sample which we have. In fact, it is even too strong. For the same reasons that described above, one can even admit the following assumption.

Hypothesis 8.3.3 *In this section, in some cases, one will suppose* $\beta < 1/5$.

Some lemmas

To avoid complicating the notations, one keeps the notations a_i and A_i instead of a_i^n and A_i^n. Then,

$$\frac{p_{a_i}^2}{Var} = \frac{p_{a_i}^2}{(N_{A^n}/m_n)\Big[\sum_s [p_{x_s}(p_{x_s} - 1/m_n)]\Big]}$$

$$= \frac{n_{a_i}^2/m_n^2}{L_{A^n}\Big[\sum_s (n_{x_s}/m_n)(n_{x_s}/m_n - 1/m_n)\Big]} = \frac{n_{a_i}^2}{L_{A^n}\Big[\sum_s n_{x_s}(n_{x_s} - 1)\Big]} \ .$$

Let $d > 0$. By the theorem of page 103 of [19], we know that $\sum_{i=1}^{c'-c} \frac{p_{A_i} - 1/m_n}{\sqrt{Var}} \xrightarrow{D} N(0,1)$ if and only if, for all $d > 0$,

$$\sum_{i=1}^{c'-c} E\Big\{1_{]d,\infty[}\Big(\Big|\frac{p_{A_i} - 1/m_n}{\sqrt{Var}}\Big|\Big)\frac{(p_{A_i} - 1/m_n)^2}{Var}\Big\}$$

$$= \sum_{i=1}^{c'-c} E\Big\{1_{]d,\infty[}\Big(\Big|\frac{n_{A_i} - 1}{\sqrt{m_n^2.Var}}\Big|\Big)\frac{(n_{A_i} - 1)^2}{m_n^2.Var}\Big\} \to 0.$$

In order to study this limit, we shall need the following lemma.

Lemma 8.3.3 *Let* $d > 0$. *Let* $\alpha(n) = d\sqrt{L_{A^n}\big(\sum_s n_{x_s}(n_{x_s} - 1)\big)}$. *We denote by* $F_{n,j}$ *the distribution of* $\frac{p_{A_j} - 1/m_n}{\sqrt{Var}}$. *Then,*

$$\sum_j \int 1_{]d,\infty[}(|x|)x^2.dF_{n,j}(x) = \sum_{s, n_{x_s} - 1 > \alpha(n)} \frac{(n_{x_s} - 1)^2}{\big(\sum_s n_{x_s}(n_{x_s} - 1)\big)} \ .$$

Proof We have

$$\int 1_{]d,\infty[}(|x|)x^2.dF_{n,j}(x)$$

$$= E\Big\{1_{]d,\infty[}\Big(\frac{|p_{A_j} - 1/m_n|}{\sqrt{Var}}\Big)\frac{(p_{A_j} - 1/m_n)^2}{Var}\Big\}$$

$$= E\Big\{\sum_s 1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{|p_{A_j} - 1/m_n|}{\sqrt{Var}}\Big)\frac{(p_{A_j} - 1/m_n)^2}{Var}\Big\}$$

$$= \sum_s E\Big\{1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{|n_{A_j} - 1|}{m_n\sqrt{Var}}\Big)\frac{(p_{A_j} - 1/m_n)^2}{Var}\Big\}$$

$$= \sum_s E\Big\{1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{|n_{A_j}-1|}{m_n\sqrt{L_{A^n}\big(\sum_s p_{x_s}(p_{x_s}-1/m_n)\big)}}\Big)\frac{(p_{A_j}-1/m_n)^2}{Var}\Big\}$$

$$= \sum_s E\Big\{1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{|n_{A_j}-1|}{\sqrt{L_{A^n}\big(\sum_s n_{x_s}(n_{x_s}-1)\big)}}\Big)\frac{(p_{A_j}-1/m_n)^2}{Var}\Big\}$$

$$= \sum_s E\Big\{1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{|n_{A_j}-1|}{\alpha(n)/d}\Big)\frac{(p_{A_j}-1/m_n)^2}{Var}\Big\}$$

$$= \sum_{s,n_{x_s}-1>\alpha(n)} E\Big\{1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{d|n_{A_j}-1|}{\alpha(n)}\Big)\frac{(p_{A_j}-1/m_n)^2}{Var}\Big\}$$

$$= \sum_{s,n_{x_s}-1>\alpha(n)} E\Big\{1_{x_s}(A_j)\frac{(p_{A_j}-1/m_n)^2}{Var}\Big\}$$

$$= \sum_{s,n_{x_s}-1>\alpha(n)} E\Big\{1_{x_s}(A_j)\frac{(n_{A_j}-1)^2}{L_{A^n}\big(\sum_s n_{x_s}(n_{x_s}-1)\big)}\Big\}$$

$$= \sum_{s,n_{x_s}-1>\alpha(n)} E\Big\{1_{x_s}(A_j)\frac{(n_{x_s}-1)^2}{L_{A^n}\big(\sum_s n_{x_s}(n_{x_s}-1)\big)}\Big\}$$

$$= \frac{1}{m_n}\sum_{s,n_{x_s}-1>\alpha(n)} \frac{(n_{x_s}-1)^2}{L_{A^n}\big(\sum_s n_{x_s}(n_{x_s}-1)\big)} \ .$$

Therefore,

$$\sum_j \int 1_{]d,\infty[}(|x|)x^2.dF_{n,j}(x)$$

$$= \frac{N(A^n)}{m_n}\sum_{s,n_{x_s}-1>\alpha(n)} \frac{(n_{x_s}-1)^2}{L_{A^n}\big(\sum_s n_{x_s}(n_{x_s}-1)\big)}$$

$$= \sum_{s,n_{x_s}-1>\alpha(n)} \frac{(n_{x_s}-1)^2}{\big(\sum_s n_{x_s}(n_{x_s}-1)\big)} \ . \ \blacksquare$$

Study of the assumption "$\sum_s n_{x_s}(n_{x_s}-1) \to \infty$".

In the set of the probabilities p_{x_s}, the assumption $\sum_s n_{x_s}(n_{x_s}-1) \to \infty$ is an often satisfied assumption. Now, one will understand that, or $\sum_{s,n_{x_s}-1>\alpha(n)} \frac{(n_{x_s}-1)^2}{L_{A^n}\left(\sum_s n_{x_s}(n_{x_s}-1)\right)} \to \ 0$, or the uniformity holds.

Let us try to understand why.

First, let us suppose that $\sum_s n_{x_s}(n_{x_s}-1) \to \infty$ and consider the $x_s = x^n_{s_n}$ such as $\frac{(n_{x_s}-1)^2}{\sum_s n_{x_s}(n_{x_s}-1)}$ does not converge to 0. Let us consider two cases.

1) Either $\sum_s n_{x_s}(n_{x_s}-1)$ tends very quickly to ∞, i.e. is normally much larger than any $(n_{x_s}-1)^2$. In this case, it is possible that there exists such x_s. But, for all d, for n large enough, $d^2(n_{x_s}-1)^2$ is much smaller than $\sum_s n_{x_s}(n_{x_s}-1)$, and therefore, $E\Big\{1_{x_s}(A_j)1_{]d,\infty[}\Big(\frac{|n_{A_j}-1|}{m_n\sqrt{Var}}\Big)\frac{(p_{A_j}-1/m_n)^2}{Var}\Big\} = 0$.

Therefore $\sum_{s,n_{x_s}-1>\alpha(n)} E\Big\{1_{x_s}(A_j)\frac{(n_{x_s}-1)^2}{L_{A^n}\Big(\sum_s n_{x_s}(n_{x_s}-1)\Big)}\Big\} = 0$.

2) Or, $\sum_s n_{x_s}(n_{x_s}-1)$ tends slowly to ∞.

In this case, $(1/m_n)\sum_s n_{x_s}(n_{x_s}-1) \approx 0$ means that $n_{x_s} \approx 1$. Therefore $p_{x_s} \approx 1/m_n$. Therefore the uniformity holds.

These results are proved by the many numerical simulations that we made. It are also confirmed for the following examples.

Example 8.3.4 *Let us suppose that the probability is concentrated in $m_n^{3/4}$ distinct points x_s such as $n_{x_s} = m_n^{1/4}$.*

Study Under these assumptions, $\sum_s n_{x_s}(n_{x_s}-1) = m_n^{3/4}(m_n^{1/4})(m_n^{1/4}-1) \approx m_n^{5/4}$. Moreover, $n_{a_i}^2 \leq m_n^{1/2}$ Therefore,

$$\left\{\left|\frac{n_{a_i}^2}{L_{A^n}\Big[\sum_s n_{x_s}(n_{x_s}-1)\Big]}\right| \geq d\right\} = \Big\{n_{a_i}^2 \geq d.L_{A^n}m_n^{3/4}m_n^{1/4}(m_n^{1/4}-1)\Big\} \to \emptyset\ .\blacksquare$$

Example 8.3.5 *Suppose $\sum_s n_{x_s}(n_{x_s}-1) \to \infty$ as $n \to \infty$ and $n_{a_i} < K$.*

Study In a obvious way, close to infinity,

$$\left\{\left|\frac{n_{a_i}^2}{L_{A^n}\Big[\sum_s n_{x_s}(n_{x_s}-1)\Big]}\right| > d\right\} = \emptyset\ .\blacksquare$$

In order to prove asymptotic normality, one could try to completely make the mathematical proofs associated with the previous reasoning. But, there is simpler: it is enough to use the following example.

Example 8.3.6 *Suppose that, for all s,* $n_{x_s} = 1 + u_s^n$*, where for all n,* $u_s^n = U_s^n(\omega_2)$ *, and where* U_s^n *,* $s = 1, 2, ...m_n$ *is a sequence of IID random variables, with mean 0, defined on a probability space* $(\Omega_2, \mathcal{A}_2, P_2)$ *such that* $E\{(U_1^n)^t\} = e_t^n$ *for t=2,3,4.*

Study First, remark that one is interested only in the sum $\sum_s n_{x_s}^2 - \sum_s n_{x_s}$. Then, because $\sum_s n_{x_s} = m_n$, according to proposition 8.1.1, one can always suppose $n_{x_s} = 1 + U_s^n$.

Under these assumptions

$$\frac{\sum_s n_{x_s}(n_{x_s} - 1) - m_n e_2^n}{\sqrt{m}_n} = \frac{\sum_s n_{x_s}^2 - \sum_s n_{x_s} - m_n e_2^n}{\sqrt{m}_n}$$

$$= \frac{\sum_s [1 + (U_s^n)^2 + 2U_s^n - 1 - U_s^n] - m_n e_2^n}{\sqrt{m}_n} = \frac{\sum_s [(U_s^n)^2 - e_2^n] + U_s^n}{\sqrt{m}_n}$$

which has asymptotically a Gaussian distribution.

The associated variance is

$$\sigma_{V_n}^2 = E\big\{\big[(U_s^n)^2 - e_2^n + U_s^n\big]^2\big\}$$

$$= E\big\{(U_s^n)^4 + (e_2^n)^2 + (U_s^n)^2 - 2e_2^n(U_s^n)^2 + 2(U_s^n)^3 - 2e_2^n U_s^n\big\}$$

$$= E\big\{(U_s^n)^4\big\} + (e_2^n)^2 + e_2^n - 2(e_2^n)^2 + 2E\big\{(U_s^n)^3\big\}$$

$$= E\big\{(U_s^n)^4\big\} + 2E\big\{(U_s^n)^3\big\} - (e_2^n)^2 + e_2^n$$

$$= (e_4^n) + 2(e_3^n) - (e_2^n)^2 + e_2^n \ .$$

Therefore, if b is large enough, with a probability very close to 1,

$$\sum_s n_{x_s}^2 - \sum_s n_{x_s} = m_n e_2^n + Ob(1)b\sigma_{V_n}\sqrt{m}_n \to \infty \ . \tag{8.2}$$

Indeed, $n_{x_s} \leq m_n^\beta$, $\beta < 1/4$. Therefore, $e_4^n \leq m_n^{4\beta}$, $e_3^n \leq m_n^{3/4}$. Finally, one can admit $\sigma_{V_n}^2 \leq 3m_n^{4\beta}$.

Therefore, $b\sigma_{V_n}\sqrt{m}_n/m_n \to 0$. Therefore, if $e_2^n \geq C_3 > 0$,

$$m_n e_2^n + Ob(1)b\sigma_{V_n}\sqrt{m}_n \to \infty \ .$$

Now, because $n_{x_s}^2 \leq m_n^{1/2}$, for all $d > 0$, there exists N_0 such that, for all $d' \leq dL_{A^n}[e_2^n + Ob(1)b\sigma_{V_n}/\sqrt{m}_n]$,

$$\left\{ \left| \frac{n_{a_i}^2}{L_{A^n}\left[\sum_s n_{x_s}(n_{x_s}-1)\right]} \right| > d \right\} \subset \left\{ n_{a_i}^2 > d' m_n \right\} = \emptyset$$

This result holds again if $\sum_s n_{x_s}^2 - \sum_s n_{x_s} \geq \gamma(n) m_n^{1/2}$ where $\gamma(n) > c_6 > 0$ where c_6 is a constant. ■

Therefore, finally, if $\sum_s n_{x_s}(n_{x_s}-1) \geq \gamma(n) m_n^{1/2}$, one can always admit that we have or the CLT whose we shall deduce uniformity or the uniformity itself.

8.3.3 Study of the other assumptions

Case $m_n^{-1/2} \sum_s n_{x_s}(n_{x_s}-1) < K_1$.

It one of the cases not studied in the example 8.3.6. It occurs if $e_2^n \to 0$.

However, this case is simpler: it implies obligatorily the uniformity.

Indeed if $\sum_s n_{x_s}(n_{x_s}-1) < K_1 m_n^{1/2}$,

$$\sum_s (n_{x_s}-1)^2 = \sum_s n_{x_s}(n_{x_s}-1) - \sum_s (n_{x_s}-1)$$

$$= \sum_s n_{x_s}(n_{x_s}-1) - (m_n - m_n) < K_1 m_n^{1/2} \ .$$

Then, we can admit $O(e_2^n) = (1/m_n)\sum_s (n_{x_s}-1)^2 < K_1/m_n^{1/2}$.

Now, let $Var_U(N_{A_i})$ be the variance of $N_{x_s} = 1 + U_s^n$. Then

$$Var_U(N_{A_i}) = e_2^n \to 0 \ .$$

If $Var_U(N_{A_i})$ is close to 0, that means that n_{a_i} is close to 1, i.e. p_{a_i} is close to $1/m_n$. i.e. that one has a uniform probability.

There is thus always the uniformity in this case. Therefore, finally that proves that the quasi-uniformity of the marginal probabilities is satisfied in all the cases.

Case $e_2^n \to 0$.

It is the last case not studied of the example 8.3.6. However, it is the same case as previously. Indeed,

$$Var_U(N_{A_i}) = e_2^n \to 0 \ .$$

There is thus always the uniformity in this case. Therefore, finally that proves that the quasi-uniformity of the marginal probabilities is satisfied in all the cases.

Case $m_n^{-1/2} \sum_s n_{x_s}(n_{x_s}-1)$ is not bounded and does not converge to ∞

In fact, these cases do not interest us. One is in a concrete study. One chose a "m", i.e. "n". The $n_{x_s} = n_{x_s}^n$ are fixed. Thus $m_n^{-1/2} \sum_s n_{x_s}(n_{x_s}-1)$ is fixed. It is large or small, not both. The probabilities that we will really obtain will check, either $\sum_s n_{x_s}(n_{x_s}-1)$ enough large, or $\sum_s n_{x_s}(n_{x_s}-1)$ enough small. One will be able to thus consider that one is in the case $\sum_s n_{x_s}(n_{x_s}-1) \to \infty$ quickly or in the case $\sum_s n_{x_s}(n_{x_s}-1) < K_1 m^{1/2}$.

We want only that, numerically, $\left|P(\widehat{T}^{-1}(I) - 1/d^q\right|$ is small. Now, in all the cases, one obtains this result.

8.3.4 Numerical study of the Central Limit Theorem

Thus the uniformity is checked in all the limit theoretical cases. Now, the $\frac{n_{A_i}-1}{\sqrt{L_{A^n} \sum_s n_{xs}(n_{xs}-1)}}$ are IID. Now, convergence for an IID sequence is very fast.

Moreover, which interests us, it is that $\left|\sum_i (p_{a_i} - 1/m_n)\right|$ is small enough. It is a result less strong than the CLT.

Thus, because convergence is very fast, one could almost do without the theoretical study of the double sequences. Numerically, it will be almost always found that $\left|\sum_i p_{a_i} - 1/d^q\right|$ is enough small. It is what is arisen from the numerical simulations that we made (but it was good to confirm this result by a theoretical study).

Thus, for a = 165580141 , m= 267914296 , d= 10, q = 4, one tested more than 250 different probabilities in 100 possible points k/m. The maxima for $\left|\sum_i (p_{a_i} - 1/10^4)\right|$ are

$4.715 * 10^{-9}$	$11.7052 * 10^{-9}$	$6.0254 * 10^{-9}$	$0.8687 * 10^{-9}$	$2.1077 * 10^{-9}$

8.3.5 Consequences of the Central Limit Theorem

Meaning of the assumption CLT

Therefore, one supposes that

$$\frac{\sum_i (n_{A_i} - 1)}{\sqrt{L_{A^n} \sum_s n_{x_s}(n_{x_s} - 1)}}$$

has asymptotically a standard Gaussian distribution. I.e. , for the probability of A_i,

$$Proba_2\Big\{ \frac{|\sum_i (p_{A_i} - 1/m_n)|}{(1/m_n)\sqrt{L_{A^n} \sum_s n_{x_s}(n_{x_s} - 1)}} \geq b \Big\} \approx \Gamma(b) \ .$$

Therefore,

$$Proba_2\Big\{ \big| \sum_i (p_{A_i} - 1/m_n) \big| \geq \frac{\sqrt{L_{A^n}} \sqrt{\sum_s n_{x_s}(n_{x_s} - 1)}\ b}{m_n} \Big\} \approx \Gamma(b) \ .$$

Now we have the assumption $n_{x_s} \leq m_n^{1/4}$: cf hypothesis 8.3.2. Therefore, $\sum_s n_{x_s}(n_{x_s} - 1) \leq m_n(m_n^{1/4})(m_n^{1/4} - 1) \leq m_n^{3/2}$. Therefore, generally,

$$Prob\Big\{ \big| \sum_i (p_{A_i} - 1/m_n) \big| \geq \frac{\sqrt{L_{A^n} m_n^{3/2}}.b}{m_n} \Big\} \leq \Gamma(b) \ .$$

It is thus that which means that the CLT holds.

Consequences under assumptions $n_{x_s} \leq m_n^{1/5}$

Now, the previous results are not specific enough: if they thus are applied one could lose data during construction of the random sequence $b^1(n')$. Also we will study now what occurs under hypothesis 8.3.3.

First, let us suppose that one has a probability concentrated in a number of points x_s near of $m_n^{4/5}$ where $n_{x_s} \approx m_n^{1/5}$ for these points. In this case, one has about $\sum_s n_{x_s}(n_{x_s} - 1) \leq m_n^{4/5}(m_n^{1/5})(m_n^{1/5} - 1) \leq m_n^{6/5}$.

One can understand that the result is true by numerical studies. In fact, the maximum seems reached for a probability concentrated in a number of points x_s near to $m_n^{4/5}$ with $n_{x_s} \approx m_n^{1/5}$.

In this case, one has about $\sum_s n_{x_s}(n_{x_s} - 1) \leq m_n^{6/5}$.

Remark 8.3.7 *This result remain true in all the cases if one admits the assumption $m_n^{1/5} \geq n_{x_s}$. That seems rather difficult to show completely by mathematical reasoning, but intuitively, it appears true. In any case, one can always circumvent the problem during the construction of $b^1(n')$ by choosing better parameters. But that would probably lead to an useless loss of data.*

Let us suppose again that one has a probability close to $m_n^{4/5}$ concentrated in a number of points x_s such that $n_{x_s} \approx m_n^{1/5}$. Then, generally,

$$Proba_2\Big\{ \big| \sum_i (p_{A_i} - 1/m_n) \big| \geq \frac{\sqrt{L_{A^n} m_n^{6/5}}.b}{m_n} \Big\} \leq \Gamma(b) \ .$$

Therefore,

$$Proba_2\Big\{ \big| \sum_i (p_{A_i} - 1/m_n) \big| \geq \frac{\sqrt{L_{A^n}}.b}{m_n^{4/10}} \Big\} \leq \Gamma(b) \ .$$

Let us deal the case where one uses the application $T_1^m : F(m^1) \to F(m)$, i.e. q=1, d=m, where m and m^1 belong to the Fibonacci sequence : cf section 11.1.2. As a matter of fact, with our current notations, it would be better to write $(m^1)_n$ and $(m)_n$ instead of m^1 and m. Suppose that $(m)_n \approx (m^1)_n^{3/5}$. In this case, $L_{A^n} \approx (m^1)_n^{-3/5}$. Therefore, we have

$$Proba_2\Big\{ \big| \sum_i (p_{A_i} - 1/(m^1)_n) \big| \geq \frac{b}{(m^1)_n^{2/5} (m^1)_n^{3/10}} \Big\} \leq \Gamma(b) \ .$$

Therefore, with a probability approximately larger than $1 - \Gamma(b)$,

$$\big| P_{X_n}\{X_n = k/(m)_n\} - 1/(m)_n \big| \leq \frac{b}{(m^1)_n^{7/10}} \ .$$

Therefore,

$$P_{X_n}\{X_n = k/(m)_n\} = \frac{1}{(m)_n} \Big[1 + \frac{O(1).b}{(m^1)_n^{1/10}} \Big] \ .$$

For example, if $(m^1)_n \geq 10^{30}$, $(m)_n \approx 10^{20}$, b=20, with a probability larger than $1 - \Gamma(b) = 1 - 10^{-89}$,

$$P_{X_n}\{X_n = k/(m)_n\} = \frac{1}{(m)_n} \Big[1 + \frac{O(1).20}{10^{30/10}} \Big] = \frac{1}{(m)_n} \Big[1 + \frac{O(1)}{50} \Big] \ .$$

These results are sufficient to guarantee that the marginal laws are sufficiently close to the uniform law. They are the assumptions that one will choose in section 11.1.2.

Case where the assumption CLT is not satisfied

Therefore the assumption CLT means that

$$Proba_2\Big\{ \frac{|\sum_i (p_{A_i} - 1/m_n)|}{(1/m_n)\sqrt{L_{A^n} \sum_s n_{x_s}(n_{x_s}-1)}} \geq b \Big\} \approx \Gamma(b) \ .$$

Now with a very weak probability on the A_j, it is possible that $\big|\sum_i (p_{a_i} - 1/m_n)\big|$ is not very small and thus that $\big|P(\widehat{T}^{-1}(I) - 1/d^q\big|$ is not small enough.

What is what that means?

We keep the notations of example 8.3.6 : $n_{x_s} = 1 + u_s^n$ where $u_s^n = U_s^n(\omega)$ and where U_s^n is a IID sequence of random variables.

The fact that $\big|\sum_i (p_{a_i} - 1/m_n)\big|$ is not too small can happen, for example, if a large number of the n_{a_i} checks $n_{a_i} \geq 3/2$ for i=1,2,...,c'-c.

But, in this case, that would mean that there is a connection between the set of $\{a_1, a_2, ...\}$ and the probability induced by the y_n. Then, it was understood that it is not the case.

In reality, this possibility always exists. That means that the sample of the data y_n does not behave as it should. It is known that it is always possible, but with a very weak probability: in this case, the $\{a_1, a_2, ...\}$ can indeed have a connection with the empirical probability associated with the text y_n.

But that does not correspond to our assumption a priori. One thus rejects it with regard to the theoretical probabilities (but not with regard to empirical probabilities).

Thus the random sequences of bits $b^1(n')$ obtained in section 11.2 will check the theoretical assumption IID. But it is always possible that, despite all our constructions, the sequence b_i which we obtain finally does not check all the tests of randomness, as it is the case for any sample really IID, i.e. with a negligible probability.

8.3.6 Conclusion

Finally under our assumptions, one finds no theoretical case where the p_{x_s} are not close to $1/m$. As a matter of fact, one can prove that it is also true for many Borel sets Bo.

It is an important result. Indeed, for a sequence of data y_n, there exists very large number of possible models Y_n. But the reasoning above shows that $\big|Proba_2\{\widehat{T}^{-1}(I)\} - L(I)\big|$ is small for **all** the reasonable and logical models.

It is concluded that, for all the correct models X_n of the sequence of data $x_n^* = T_q(y_n)$ the marginal probability of the X_n is very close to the uniform probability. That is thus true for sequences $e_S^3(j) = m_S T_1^{m_S}(e_S^2(j)/m_S^1)$ (cf

section 11.2).

This result is important : **in the section 11.2, we will be sure that our model for the sequence $E^3(j)$ is good** : the probability marginal of the $E^3(j)$ is very close to the uniform probability.

8.4 Multidimensional case

One will understand now that the previous results seems to be again valid in the multidimensional case. We did not make a detailed study of this problem because the techniques which we employ in the chapters 11 and 12 do not require it.

8.4.1 Empirical probability

Let $x_n^* = T_q(y_n)$, $n = 1, 2, ..., n_0$, be a sample of the random variables $X_n = T_q(Y_n)$, where all the y_n's are distincts. For t=1,2,...,p, we set $I^t = \left[\frac{c_t}{m}, \frac{c'_t}{m}\right[$ where c_t, $c'_t \in F^*(m)$. In order to simplify notations, in some cases we identify I^t, $I^t \cap F(m)$ and $I^t \cap F(d^q)$. We denote by P_e the associated empirical probabilities. Therefore, we have

$$P_e\Big\{\big\{T_q(Y_{n+j_1}) \in I^1\big\} \cap .... \cap \big\{T_q(Y_{n+j_p}) \in I^p\big\}\Big\} = \sum_{x^1_{s_1} \in I^1} .... \sum_{x^p_{s_p} \in I^p} pe_{x^1_{s_1},...,x^p_{s_p}} \ ,$$

where $pe_{x^1_{s_1},...,x^p_{s_p}} = P_e\Big\{\big\{\widehat{T}(Y_{n+j_1}) = x^1_{s_1}\big\} \cap .... \cap \big\{\widehat{T}(Y_{n+j_p}) = x^p_{s_p}\big\}\Big\}$.

We set $A^t = \widehat{T}^{-1}(I^t) = \{a^t_1, ..., a^t_{c'_t - c_t}\}$ for t=1,2,...,p.

Let us suppose that the y_n is obtained starting from text. The y_n thus represent parts of texts. There is then no logical connection between this text and the sets A^t.

Relationship between texts and the sets A^t : 1

As in section 8.1.4 there is no logical connection between the set $\mathcal{J}$ =

[Newton's theory]
[of gravitation was]
[soon accepted wit]
.....

and the sets of type $\mathcal{H}_1 =$

[whgkudf ly cuqhjg]
[aamxgusdggbxckmp]
[x;cbkutcc ze xycyc x]
[oeHlm mlk Rdfg yu]
.....

In multidimensional case, the probability that {[Newton's theory], [of gravitation was], [soon accepted wit]} $\subset \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ is approximatively equal to $\prod_{t=1}^{3} \left[card(\mathcal{H}_s t)/32^{18}\right]$. These results can also be understood by numerical studies.

Relationship between texts and the sets A^t : 2

Because there is always no connection between parts of texts $(y_{n+j_1}, ..., y_{n+j_p})$ and the sets $A^1 \otimes .... \otimes A^p$, it is thus logical that sums on the various possible sets $A^1 \otimes .... \otimes A^p$ (where $p \leq Log(n_0)/log(2)$), behave as sums over sets chosen randomly, i.e.

$$\sum_{x_{s_1}^1 \in I^1} ..... \sum_{x_{s_p}^p \in I^p} pe_{x_{s_1}^1, ..., x_{s_p}^p} \approx C^1$$

for almost all the sets which have a same size : cf corollary 8.1.6 (C^1 is a constant).

As it is true for all the sets $A^1 \otimes .... \otimes A^p$, C^1 must check $C^1 \approx \prod_{t=1}^{p} L(I^t)$. For example, one can apply this result to several hypercubes of the same size $I^1 \otimes .... \otimes I^p$ forming a partition of $F(m)^p$.

Example : continuous functions

Consider the vector $(X_{n+j_1}, ..., X_{n+j_p})$. We know that

$$\sum_{x_{s_1}^1 \in \left[\frac{c_1}{m}, \frac{c_1'}{m}\right[} ..... \sum_{x_{s_p}^p \in \left[\frac{c_p}{m}, \frac{c_p'}{m}\right[} p_{x_{s_1}^1, ..., x_{s_p}^p}$$

$$= \sum_{(y_{n+j_1}, ..., y_{n+j_p}) \in A^1 \otimes .... \otimes A^p} \mathbb{E}\left\{1_{A^1 \otimes .... \otimes A^p}(y_{n+j_1}, ..., y_{n+j_p})\right\} .$$

Let us suppose that $(Y_{n+j_1}, ..., Y_{n+j_p})$ is approximately a function with continuous density and that the $c_t' - c_t$ are large enough. By applying the traditional

methods of integration, it is then clear that

$$\sum_{x^1_{s_1}\in I_1} ..... \sum_{x^p_{s_p}\in I_p} p_{x^1_{s_1},..,x^p_{s_p}} \approx \prod_{t=1}^{p} L(I_t) \ .$$

Under the assumption of this continuous model, one thus obtains easily IID sequences. Let us recall that, when $n_0 << d^p$, one often accepts like model a model with continuous density. That shows that the functions T_q are a good tool to obtain IID sequences.

8.4.2 Theoretical probability

Let us choose a random vector $(X_n, X_{n+1}, ..., X_{n+p})$. One generalizes the notations of section 8.2 . We set

$$p_{x^1_{s_1},..,x^p_{s_p}} = P\big\{(X_{n+1} = x^1_{s_1}) \cap ... \cap (X_{n+p} = x^p_{s_p})\big\}.$$

Probabilities chosen randomly

One generalizes the reasoning of the unidimensional case. At first a natural order is defined in $F(m)^p$. According to proposition 8.1.1, there exists a permutation ψ_0 such that the sequence $p_{\psi_0(y^1_{s_1},..,y^p_{s_p})} = p^f_{y^1_{s_1},..,y^p_{s_p}}$ can be regarded as an IID sample of random variables (according to the order defined on $F(m)^p$) which have the distribution M_Z.

Because T can be regarded as independent of the text, therefore independent of probabilities $p^f_{y^1_{s_1},..,y^p_{s_p}}$, the permutation $\psi_1(y^1_{s_1},..,y^p_{s_p}) = \psi_0^{-1}(\overline{T}(y^1_{s_1}),..,\overline{T}(y^p_{s_p}))$ corresponding to

$$p^f_{y^1_{s_1},..,y^p_{s_p}} \mapsto p^f_{\psi_1(y^1_{s_1},..,y^p_{s_p})} = p_{\psi_0(\psi_0^{-1}(\overline{T}(y^1_{s_1}),..,\overline{T}(y^p_{s_p})))} = p_{\overline{T}(y^1_{s_1}),..,\overline{T}(y^p_{s_p})} = p_{x^1_{s_1},..,x^p_{s_p}}$$

can be regarded as chosen randomly.

Therefore, the sequence $p_{x^1_{s_1},..,x^p_{s_p}} = p^f_{\psi_1(y^1_{s_1},..,y^p_{s_p})}$ can be regarded as an IID sample of random variables which have the distribution M_Z.

Model

Therefore, one can consider that the associated theoretical probabilities behave as if they were chosen randomly.

Hypothesis 8.4.1 *Suppose that*

$$p_{x^1_{s_1},..,x^p_{s_p}} = \frac{p'_{x^1_{s_1},..,x^p_{s_p}}}{\sum_{i_1=1}^{m} .... \sum_{i_p=1}^{m} p'_{i_1/m,...,i_p/m}} \ .$$

We assume that the $p'_{x^1_{s_1},..,x^p_{s_p}}$'s are a sample of a sequence of IID random variables $P'_{x^1_{s_1},..,x^p_{s_p}}$ defined on the probability space $(\Omega_1, \mathcal{A}_1, Proba_1)$: $p'_{x^1_{s_1},..,x^p_{s_p}} = P'_{x^1_{s_1},..,x^p_{s_p}}(\omega_1)$. One supposes that $P'_{x^1_{s_1},..,x^p_{s_p}}$ has the distribution M. Let E_M and σ^2_M be the associated expectation and the associated variance, respectively .

We shall need the following notations.

Notations 8.4.1 *Let $p_m = \lfloor Log(n_0)/log(2) \rfloor$. For all $b > 0$, we denote by $\Gamma_1(b)$ the fonction $\Gamma_1(b) = 2.Max_{p \leq p_m, n_s \geq n^2} Proba_1\big\{|S_{(n),p}| \geq b\big\}$ where $n^2 = min_{k \in F^*(d^q)}(card\{r/m \mid k/d^q \leq r/m < (k+1)/d^q)\}$, $(n) = (n_1, ..., n_p)$, and where*

$$S_{(n),p} = \frac{\sum_{i_1=1}^{n_1} .... \sum_{i_1=1}^{n_p} P'_{i_1/m,...,i_p/m}}{\sigma_M \sqrt{n_1 n_2 .... n_p}} \ .$$

Of course $\Gamma_1(b) \approx 2\Gamma(b)$ as soon as m/d^q is large enough.

Independence in each point

Now, one will understand that one has a good approximation of the independence in each point. First, one has the following proposition.

Proposition 8.4.1 *The following approximation holds.*

$$P\Big\{(X_n, ..., X_{n+p}) = (k_1, ..., k_p)/d^q\Big\} \approx 1/d^{pq} \ .$$

Proof We assume that $I_t = \big[k_t/d^q, (k_t+1)/d^q\big[$. Then,

$$P\Big\{(X_n, ..., X_{n+p}) = (k_1, ..., k_p)/d^q\Big\} = \sum_{x^1_s \in I_1} ..... \sum_{x^p_s \in I_p} p_{x^1_{s_1},...,x^p_{s_p}} \ .$$

Because $p_{x^1_{s_1},..,x^p_{s_p}} = \frac{p'_{x^1_{s_1},...,x^p_{s_p}}}{\sum_{i_1=1}^m .... \sum_{i_p=1}^m p'_{i_1/m,...,i_p/m}}$ and because the $P'_{x^1_{s_1},..,x^p_{s_p}}$ have the distribution M,

$$P\Big\{\big\{\widehat{T}(Y_1) \in I_1\big\} \cap ..... \cap \big\{\widehat{T}(Y_p) \in I_p\big\}\Big\}$$

$$= \frac{\prod_s (c'_s - c_s)}{m^p} \ \frac{\frac{1}{\prod_s (c'_s - c_s)} \sum_{x^1_s \in I_1} ..... \sum_{x^p_s \in I_p} p'_{x^1_{s_1},....,x^p_{s_p}}}{\frac{1}{m^p} \sum_{i_1=1}^m .... \sum_{i_p=1}^m p'_{i_1/m,...,i_p/m}}$$

which converges in probability to $\frac{\prod_s (c'_s - c_s)}{m^p}$ because the $p'_{x^1_{i_1},....,x^p_{i_p}}$ are an IID sample. ■

More precisely the following lemma hold.

Lemma 8.4.1 *We assume that* $I_t = \big[k_t/d^q, (k_t+1)/d^q\big[$. *Let* $N(I_t)$ *be the number of* $r/m \in F(m)$ *such that* $k_t/d^q \leq r/m < (k_t+1)/d^q$. *We suppose* $I_t \cap F(m) = [c_t/m, c'_t/m[\cap F(m)$.

Let us suppose m *enough large compared to* d^q *and to the* $N(I_t)$. *Let us suppose that* b *and* σ_M *are not too large. Let us suppose that* E_M *is not too small.*

Then, with a probability larger than $1-\Gamma_1(b)$,

$$P\Big\{\big\{Y_1 \in \widehat{T}^{-1}(I_1)\big\}\cap.....\cap\big\{Y_p \in \widehat{T}^{-1}(I_p)\big\}\Big\} \approx \frac{\prod_s (c'_s - c_s)}{m^p}\ \Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}}\Big]\ .$$

Proof We use the CLT. Then, with a probability larger than $1-\Gamma_1(b)/2$,

$$\frac{1}{\prod_s N(I_s)} \sum_{x_s^1 \in I_1} ..... \sum_{x_s^p \in I_p} p'_{x^1_{s_1},....,x^p_{s_p}} = E_M + \frac{Ob(1).b\sigma_M}{\sqrt{\prod_s N(I_s)}}\ .$$

Then, with a probability larger than $1-\Gamma_1(b)$,

$$P\Big\{\big\{Y_1 \in \widehat{T}^{-1}(I_1)\big\}\cap.....\cap\big\{Y_p \in \widehat{T}^{-1}(I_p)\big\}\Big\}$$

$$= \frac{\prod_s (c'_s - c_s)}{m^p}\ \frac{\frac{1}{\prod_s (c'_s - c_s)} \sum_{x_s^1 \in I_1} ..... \sum_{x_s^p \in I_p} p'_{x^1_{s_1},....,x^p_{s_p}}}{(1/m^p)\sum_{i_1=1}^m .... \sum_{i_p=1}^m p'_{i_1/m,....,i_p/m}}$$

$$= \frac{\prod_s (c'_s - c_s)}{m^p}\ \frac{E_M + \frac{Ob(1).b\sigma_M}{\sqrt{\prod_s N(I_s)}}}{E_M + \frac{Ob(1).b\sigma_M}{\sqrt{m^p}}}$$

$$= \frac{\prod_s (c'_s - c_s)}{m^p}\ \frac{1 + (Ob(1).b\sigma_M)/[E_M\sqrt{\prod_s N(I_s)}]}{1 + (Ob(1).b\sigma_M)/[E_M\sqrt{m^p}]}$$

$$\approx \frac{\prod_s (c'_s - c_s)}{m^p}\ \Big[1 + \frac{Ob(1).b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}} + \frac{Ob(1).b\sigma_M}{E_M\sqrt{m^p}}\Big]$$

$$\approx \frac{\prod_s (c'_s - c_s)}{m^p}\ \Big[1 + \frac{Ob(1).b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}}\Big]\ . \blacksquare$$

We deduce the following proposition.

Proposition 8.4.2 *Let us suppose* m *enough large compared to* d^q *and to the* $N(I_k)$*'s. Let us suppose that* b *and* σ_M *are not too large. Let us suppose that* E_M *is not too small. Then, with a probability larger than* $1-\Gamma_1(b)$,

$$P'\big\{X_1 = k_1/d^q, ..., X_p = k_p/d^q\big\}$$

$$= \frac{1}{d^{pq}}\Big[1 + \frac{Ob(1)2p.d^q}{m} + \frac{Ob(1).2b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}}\Big]\ .$$

Proof In this case, $[c_k/m, c'_k/m[= I_k$. By the proof of lemma 8.4.1,

$$P\Big\{\big\{Y_1 \in \widehat{T}^{-1}(I_1)\big\} \cap ..... \cap \big\{Y_p \in \widehat{T}^{-1}(I_p)\big\}\Big\}$$

$$= \frac{\prod_s (c'_s - c_s)}{m^p} \Big[1 + \frac{Ob(1).b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}} + \frac{O(1).b\sigma_M}{E_M\sqrt{m^p}}\Big] .$$

By our definition $h_0/m \leq 1/d^q \leq (h_0+1)/m$. By the lemma 8.1.2, $N(I_k) = c'_k - c_k = h_0$ or $N(I_k) = h_0 + 1$. Then,

$$\frac{c'_s - c_s}{m} = \frac{1}{d^q} + \frac{Ob(1)}{m} .$$

Then,

$$\prod \frac{c'_s - c_s}{m} = \Big[\frac{1}{d^q} + \frac{Ob(1)}{m}\Big]....\Big[\frac{1}{d^q} + \frac{Ob(1)}{m}\Big]$$

$$= \frac{1}{d^{pq}} + \frac{pOb(1)}{m.d^{(p-1)q}} + \frac{Ob(1)p(p-1)/2}{m^2.d^{(p-2)q}} + ...$$

$$= \frac{1}{d^{pq}}\Big[1 + \frac{p.d^qOb(1)}{m} + \frac{p(p-1)d^{2q}Ob(1)}{2m^2} + ...\Big] .$$

Then,

$$P\Big\{\big\{Y_1 \in \widehat{T}^{-1}(I_1)\big\} \cap ..... \cap \big\{Y_p \in \widehat{T}^{-1}(I_p)\big\}\Big\}$$

$$= \frac{1}{d^{pq}}\Big[1+\frac{p.d^qOb(1)}{m}+\frac{p(p-1)d^{2q}Ob(1)}{2m^2}+....\Big]\Big[1+\frac{Ob(1).b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}}+\frac{O(1).b\sigma_M}{E_M\sqrt{m^p}}\Big]$$

$$= \frac{1}{d^{pq}}\Big[1 + \frac{Ob(1)(3/2)p.d^q}{m}\Big]\Big[1 + \frac{Ob(1).(3/2)b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}}\Big]$$

$$= \frac{1}{d^{pq}}\Big[1 + \frac{Ob(1)2p.d^q}{m} + \frac{Ob(1).2b\sigma_M}{E_M\sqrt{\prod_s N(I_s)}}\Big] . \blacksquare$$

8.4.3 Case of Borel sets

One supposes "p" fixed with $p \leq p_m = \lfloor Log(n_0)/log(2)\rfloor$.

Now, we study the probability that $(X_{n+j_1}, ..., X_{n+j_p}) \in Bo$ where Bo is a Borel set of $F(d^q)^p$: we set $Bo = \cup_{(k_1,...,k_p)\in\Theta}\{(k_1/d^q, ..., k_p/d^q)\}$, where $\Theta \subset \{0, 1, ..., d^q - 1\}^p$. Let $K_\Theta = card(\Theta)$.

One can also write that

$$Bo = \cup_{(k_1,...,k_p)\in\Theta}\big\{[k_1/d^q, (k_1+1)/d^q[, ..., [k_p/d^q, (k_p+1)/d^q[\big\} .$$

Then, $L(Bo) = \frac{K_\Theta}{d^{pq}}$.

Let $I_{k_t} = [k_t/d^q, (k_t+1)/d^q[$. Then,

$$Proba_1\Big\{\frac{\big|\sum_{(x^1_{s_1},...,x^p_{s_p})\in\cup_{(k_1,...,k_p)\in\Theta} I_{k_1}\otimes....\otimes I_{k_p}}(P'_{x^1_{s_1},...,x^p_{s_p}} - E_M)\big|}{\sqrt{N(Bo)}\sigma_M} \geq b\Big\} \leq \Gamma_1(b)/2\,.$$

Moreover, we have the following proposition.

Proposition 8.4.3 *Let us suppose m enough large compared to d^q and to the $N(I_{k_s})$. Let us suppose that b and σ_M are not too large. Let us suppose that E_M is not too small. Then, with a probability larger than $1-\Gamma_1(b)$,*

$$P\big\{(X_{n+j_1},....,X_{n+j_p})\in Bo\big\} = L(Bo)\Big[1+\frac{Ob(1)2p.d^q}{m}+\frac{Ob(1).2b\sigma_M}{E_M.Inf\big\{\sqrt{\prod_s N(I_{k_s})}\big\}}\Big]\,.$$

Proof By proposition 8.4.2 ,

$$\begin{aligned}
&P\big\{(X_{n+j_1},....,X_{n+j_p})\in Bo\big\}\\
&= \sum_{(k_1,...,k_p)\in\Theta}\frac{1}{d^{pq}}\Big[1+\frac{Ob(1)2p.d^q}{m}+\frac{Ob(1).2b\sigma_M}{E_M.Inf\big\{\sqrt{\prod_s N(I_{k_s})}\big\}}\Big]\\
&= \frac{K_\Theta}{d^{pq}}\Big[1+\frac{Ob(1)2p.d^q}{m}+\frac{Ob(1).2b\sigma_M}{E_M.Inf\big\{\sqrt{\prod_s N(I_{k_s})}\big\}}\Big]\\
&= L(Bo)\Big[1+\frac{Ob(1)2p.d^q}{m}+\frac{Ob(1).2b\sigma_M}{E_M.Inf\big\{\sqrt{\prod_s N(I_{k_s})}\big\}}\Big]\,.\ \blacksquare
\end{aligned}$$

Now, one can prove the following lemma.

Lemma 8.4.2 *We assume that the assumptions of proposition 8.4.3 holds. Let $1\leq n\leq n_0$. Let $P_{X_n}(Bo) = P\big\{(X_{n+j_1},....,X_{n+j_p})\in Bo\big\}$. One supposes that M is the uniform distribution. Then,*

$$\begin{aligned}
Proba_1\Big\{\bigcap_{n+j_t,Bo}\Big\{\big|P_{X_n}(Bo)-L(Bo)\big|\leq L(Bo)\Big[\frac{2p.d^q}{m}+\frac{\sqrt{3}b.}{\sqrt{d^{p(Q-q)}}}\Big]\Big\}\Big\}&\\
\geq 1-n_0^p 2^{d^{pq}}\Gamma_1(b)\,.&\qquad(8.3)
\end{aligned}$$

Proof For all Borel set Bo, by proposition 8.4.3, we have thus with a probability larger than $1-\Gamma_1(b)$,

$$P\big\{(X_{n+j_1},....,X_{n+j_p})\in Bo\big\} = L(Bo)\Big[1+\frac{Ob(1)2p.d^q}{m}+\frac{Ob(1).2b\sigma_M}{E_M.Inf\big\{\sqrt{\prod_s N(I_{k_s})}\big\}}\Big]\,.$$

Then, because $\sigma_M^2 = 1/12$, $E_M = 1/2$,

$$Proba_1\Big\{\big|P_{X_n}(Bo) - L(Bo)\big| > \ L(Bo)\Big[\frac{2p.d^q}{m} + \frac{3b}{\sqrt{3.d^{p(Q-q)}}}\Big]\Big\} \leq \Gamma_1(b) \ .$$

For $h \in \{0, 1, ..., d^{pq}\}$, there are $C_{d^{pq}}^h$ Borel sets Bo_h such that $card(Bo_h) = h$. Moreover, there is at the maximum $(n_0)^p$ "$n + j_t$" possible. Then,

$$Proba_1\Bigg\{ \bigcap_{n+j_t, Bo} \Big\{\big|P_{X_n}(Bo) - L(Bo)\big| \leq \ L(Bo)\Big[\frac{2p.d^q}{m} + \frac{3b.}{\sqrt{3.d^{p(Q-q)}}}\Big]\Big\}\Bigg\}$$

$$= Proba_1\Bigg\{ \bigcap_{n+j_t, h, Bo_h} \Big\{\big|P_{X_n}(Bo_h) - L(Bo_h)\big| \leq \ L(Bo_h)\Big[\frac{2p.d^q}{m} + \frac{3b.}{\sqrt{3.d^{p(Q-q)}}}\Big]\Big\}\Bigg\}$$

$$= 1 - Proba_1\Bigg\{ \complement \bigcap_{n+j_t, h, Bo_h} \Big\{\big|P_{X_n}(Bo_h) - L(Bo_h)\big| \leq \ L(Bo_h)\Big[\frac{2p.d^q}{m} + \frac{3b.}{\sqrt{3.d^{p(Q-q)}}}\Big]\Big\}\Bigg\}$$

$$= 1 - Proba_1\Bigg\{ \bigcup_{n+j_t, h, Bo_h} \Big\{\big|P_{X_n}(Bo_h) - L(Bo_h)\big| > \ L(Bo_h)\Big[\frac{2p.d^q}{m} + \frac{3b.}{\sqrt{3.d^{p(Q-q)}}}\Big]\Big\}\Bigg\}$$

$$\geq 1 - \sum_{n+j_t, h, Bo_h} Proba_1\Big\{\big|P_{X_n}(Bo_h) - L(Bo_h)\big| > \ L(Bo_h)\Big[\frac{2p.d^q}{m} + \frac{3b.}{\sqrt{3.d^{p(Q-q)}}}\Big]\Big\}$$

$$= 1 - \sum_h \sum_{n+j_t, Bo_h} \Gamma_1(b)$$

$$= 1 - \sum_h n_0^p C_{d^{pq}}^h \Gamma_1(b)$$

$$= 1 - n_0^p 2^{d^{pq}} \Gamma_1(b) \ . \ \blacksquare$$

We deduce the following properties.

Property 8.4.3 *One supposes b big. In order that the inequality 8.3 is useful, we have to impose $b = K_{11} d^{qp/2}$ where $K_{11} > 1$.*

Proof One use the study of property 8.2.4. One reminds that $(1/2)\Gamma_1(b) \approx \Gamma(b) \approx \frac{\sqrt{2}}{\sqrt{\pi b}} \ e^{-b^2/2}$ if b is large.

Then, if one choose $b = K_{11} d^{qp/2}$ with K_{11} suitably chosen, $n_0^p 2^{d^{pq}} \Gamma_1(b) << 1$. Then, $1 - n_0^p 2^{d^{pq}} \Gamma_1(b) \approx 1$. $\blacksquare$.

Now, b has to be not too large.

Property 8.4.4 *One supposes b large. In order that the inequality 8.3 is useful, we can impose $2q < Q$.*

Proof One use the study of property 8.2.5. One reminds that if one chooses $b = K_{11} d^{qp/2}$,

$$\frac{2p.d^q}{m} + \frac{\sqrt{3}b}{\sqrt{d^{p(Q-q)}}} = \frac{2p.d^q}{m} + \frac{\sqrt{3}K_{11}}{\sqrt{d^{p(Q-2q)}}} \ .$$

In order that this equality is usefull, on can impose $\frac{\sqrt{3}K_{11}}{d^{p(Q-2q)/2}} = \epsilon_3 << 1$, and therefore, $2q < Q$. Now, $\frac{2p.d^q}{m} = \epsilon_4 \approx 0$. Then, $\epsilon_5 = \epsilon_4 + \epsilon_3 \approx 0$. We deduce that

$$Proba_1\Big\{ \bigcap_{n+j_t, Bo} \Big\{ \big| P_{X_n}(Bo) - L(Bo) \big| \leq \ L(Bo) \Big[\frac{2p.d^q}{m} + \frac{\sqrt{3}b}{\sqrt{d^{p(Q-q)}}} \Big] \Big\} \Big\}$$

$$= Proba_1\Big\{ \bigcap_{n+j_t, Bo} \Big\{ \big| P_{X_n}(Bo) - L(Bo) \big| \leq \ L(Bo)\epsilon_5 \ \Big\} \Big\}$$

$$\geq 1 - n_0^p 2^{d^{pq}} \Gamma_1(b) \approx 1 \ . \ \blacksquare$$

As in the case p=1, one notices that it seems that one is not obliged to impose this condition to have x_n IID when the size of the sample n_0 checks $n_0 << d^q$.

Remark 8.4.5 *Conditions $2q < Q$ of property 8.4.4 is much too strong. Indeed in lemma 8.4.2 one can decrease the number of conditions about Borel Set : cf [18]*

The problem of marginals laws

In the choice of spaces $(\Omega_1, \mathcal{A}_1, Proba_1)$ there is a problem: one does not take account of the marginal probabilities.

In fact with the measure defined into hypothesis 8.4.1, p is fixed. There will not be the same results if p is changed: in this case one changes space of measure. Thus let us note $(\Omega_1^p, \mathcal{A}_1^p, Proba_1^p)$ the probability spaces associated with the $(X_{n+j-1}, ., X_{n+j_p})$ for each p fixed.

Thus, in $(\Omega_1^2, \mathcal{A}_1^2, Proba_1^2)$ with a probability very close to 1 one has

$$P\Big\{ \big\{ X_1 \in I_1 \big\} \cap \big\{ X_2 \in I_2 \big\} \Big\} \approx \frac{(c_1' - c_1)(c_2' - c_2)}{m^2} \ \Big[1 + \frac{Ob(1).b}{\sqrt{3N(I_1)N(I_2)}} \Big] \ .$$

Therefore,

$$P\big\{ X_1 \in I_1 \big\} = P\Big\{ \big\{ X_1 \in I_1 \big\} \cap \big\{ X_2 \in F(m) \big\} \Big\} \approx \frac{c_1' - c_1}{m} \ \Big[1 + \frac{Ob(1)b}{\sqrt{3N(I_1)m}} \Big] \ .$$

Now, in space $(\Omega_1^1, \mathcal{A}_1^1, Proba_1^1)$, one has

$$P\big\{X_1 \in I_1\big\} \approx \frac{c_1' - c_1}{m} \ \Big[1 + \frac{Ob(1)b}{\sqrt{3N(I_1)}}\Big] \ .$$

This difference comes from the marginal probabilities. In space $(\Omega_1^2, \mathcal{A}_1^2, Proba_1^2)$ those will be already a sum of probabilities taken randomly. That means that with this measure, the marginal probabilities $p_{x_{s_1}^1} = \sum_{x_{s_2}^2} p_{x_{s_1}^1, x_{s_2}^2}$ in their vast majority will have a priori uniform distribution.

One thus does not take in account that the $p_{x_{s_1}^1}$ are probabilities in two dimensions with marginal laws, i.e. with constraints. It is thus a result which seems not to correspond to reality.

However it is not completely true. Let us place in empirical probabilities, which is always the case because $n_0 << m$. Then, the y_n seems quite independent of the sets $T_q^{-1}(I_k) \otimes T_q^{-1}(I_{k'})$. It is not true any more for the theoretical probabilities. It is normal : it is impossible to estimate p_{x_s} for each point x_s. because $n_0 << m$.

For this reason, one can choose functions with continuous densities instead of estimating the probabilities in each points p_{x_s}. But that also does not correspond to the case of the functions with continuous densities which are also a reasonable model cf remark 7.1.21.

Anyway, this is not very important : measures of spaces $(\Omega_1, \mathcal{A}_1, Proba_1)$ are only measures giving an idea of the numbers of models close to a sequence IID.

Moreover, that does not change anything with the ultimate result. The approximation $P\{x_n = k/d^q\}$ is always very fine with one or other space. Moreover, for all the logical models that we studied the result is always the same one: e.g. the case with continuous density. Indeed, the sums associated to $X_n = T_q(Y_n)$ behave as sums taken randomly. That shows well that only a negligible minority of models does not check relations of lemma 8.4.1.

To have probabilities not checking relations of lemma 8.4.1 it is necessary to choose especially selected nonnatural probabilities for the occasion: therefore our results are true for all the logical models: cf chapter 13.

8.5 Use of the T_q^d

In the construction of the random sequences of bits $b^1(n')$ we use the functions T_q^d. twice. The selection criteria of the parameters are not the same.

8.5.1 First use

One considers the transformation T_1^m used in section 11.1.2 : we remind $e_S^3(j) = mT_1^m(e^2(j)/m^1)$ where m and m^1 belong to the Fibonacci sequence.

In order to choose parameters m and m^1, one can use the results studied in section 8.3.5 : $m \leq (m^1)^{3/5}$. These results are sufficient to guarantee that the marginal laws are sufficiently close to the uniform law.

Because $m \leq (m^1)^{3/5} \leq (m^1)^{3/4}$, one supposes only $m \leq (m^1)^{3/4}$. For example, if $m^1 = O(10^{28})$, one can admit $n_{x_s} \leq (m^1)^{3/4} = O(10^{21})$. In fact, this assumption is still too strong. It corresponds for example to the fact that the probability is concentrated in approximately 10^{21} points. Indeed, in section 11.2, one uses a sample $e^2(j)$, $j = 1, 2, ..., J$ where $J \geq 10^7$, where the $e^2(j)$ are well distributed To suppose that for such a sample, the probability is concentrated in 10^{21} points does not have any logical support.

Any way, it is better to choose the parameters in a way which will be good for the *two* methods studied in this chapter in sections 8.2 and 8.3. In the first method one obtains

$$P\big\{X_n = k/d^q\big\} \approx \frac{1}{d^q}\ \Big[1 + \frac{Ob(1).b.\sigma_M}{E_M\sqrt{N(I)}}\Big]\ .$$

Generally, it is a result stronger than that obtained in the second method :

$$P_{X_n}\{X_n = k/d^q\} \approx \frac{1}{m_n^{3/5}}\Big[1 + \frac{Ob(1).b}{m_n^{1/10}}\Big]\ .$$

In the concrete cases, in order to choose well the probability that one wants for Ω_1 or Ω_2, best is to use the two previous studies together. This is why we conclude that the assumption "$m \leq (m^1)^{3/4}$" is quite sufficient.

8.5.2 Second use

One considers the transformation of the h(n) which are obtained by using the XORLT in section 11.1.2 : $x(n) = T_q(k(n))$. This choice depends on the quality of the approximation of an IID sequence which one desired to have for the sequence X(n).

In fact what interests us it is that the assumption studied in section 7.2 is correct :

$$P_Y\{Y = k/m\} = \frac{1}{m}\Big[1 + u_k\Big]\ ,$$

where u_k is a sample of an IID sequence of random variables U_k with $\sigma_U^2 \leq 1$.

But, it is necessary that this assumption is checked by the conditional probabilities $P\{H(j) = h|h_2, ..., h_p\}$, i.e.

$$P\{H(j) = h|h_2, ..., h_p\} = \frac{1}{m}\Big[1 + u_k\Big]\ .$$

One can deduce this result from the hypothesis 5.7.1 which results from the section 5.7.

To confirm this assumption, one will detail here the behavior of the conditional probabilities.

One transforms the H(j)'s by using the functions T_q: to calculate $P\{H(j) = h\}$ (or $P\{H(j) = h|h_2, ..., h_p\}$) amounts to summon the probabilities p_{x_s} associated with the F(i,j)'s. Then according to the corollary 8.1.4, one can always to consider that one has this model. The problem is to prove that one has $\sigma_U^2 \leq 1$.

Probability of G(j) One reminds that $H(j) = \overline{G(j)}$ modulo m : cf section 11.1.2. Then, by proposition 5.6.1, it is enough that this model is correct for the conditional probabilities of the G(j), $P\{G(j) = g|g_2, ..., g_p\}$.

Let us suppose again that one uses text. One understood in section 3.1.3 that, in the worst case, $p_{x_s} \leq 32/32^{r_0}$ as soon as r_0 is large enough. After transformation in $F^*(m)$, that amounts to $p_{x_s} \leq 32/m$. It is noticed that this increase $p_{x_s} \leq 32/m$ is much better than $n_{x_s} \leq m^{1/4}$.

But it is difficult to study a priori all the conditional probabilities associated with a text. One thus will choose a worse increase to be sure of this increase. Finally, by using text, according to various numerical simulations and the logical study carried out in chapter 10, one can admit that $p_{x_s} \leq m^{1/8}$ for $m \geq 10^{10}$.

Distribution of the sums of the conditional probabilities Now, one can apply the results to the laws of sums to the sums of conditional probabilities : cf proposition 5.7.1. One will thus use the sum of the conditional probabilities of the F(i,j)'s : cf section 11.1.2,

One knows that in almost all the cases, as soon as the probability is not concentrated nearly a small number of points, the speed of convergence of the sums to a probability having a density in shape of bell is extremely fast. Now, our study of the conditional probabilities in section 3.1.3 show that the conditional probabilities are not concentrated nearly a small number of points. Moreover, various numerical studies confirm this result.

Finally, the curve of the sums has the shape of bell: cf section 5.4 and 7.1.2.

Preliminary standardization We remind that the F(i,j)'s (cf section 11.1.2) are variables made uniform : $e^3(j) = mT_1^m(e^2(j)/m^1)$. Moreover, by section 8.4, T_1^m makes the $E^2(j)$ independent.

That consolidates the fact that the $P\{G(j) = g|g_2, ..., g_p\}$ are not concentrated nearly a small number of points.

Probability of H(j) The distributions of the H(j)'s have a convergence (to the uniform distribution) even clearer than that of the G(j)'s. One of the reasons, it is that to have the distributions of the H(j)'s, it is necessary to make various mixtures which make uniform the distributions : cf proposition 5.7.2.

Moreover, by proposition 5.2.4, this convergence is a convergence to the uniform distribution. There will be thus

$$P\{H(j) = h \mid H(j + j_s) = h_s\} \approx P\{H(j) = h\} = 1/m \ .$$

In dimension p, there will be

$$P\{H(j + j_s) = h_s | s = 1, 2, ..., p\} \approx 1/m^p \ .$$

Rate of convergence By proposition 5.2.4, $P\{H(j) = h|h_2, ..., h_p\} \approx 1/m$. This result is checked with a rate of convergence extremely fast : cf section 5.5.2.

Conclusion All these results show that, so that the assumption $P\{H(j) = h|h_2, ..., h_p\} = \frac{1}{m}\Big[1 + u_k\Big]$, with $\sigma_U^2 \leq 1$ is not satisfied, it would be necessary that the probabilities $P\{H(j) = h|h_2, ..., h_p\}$ are concentrated nearly a small number of point. Now, even if one used simply sums of F(i,j) not made uniform by T_1^m, it would not be the case. One can thus admit that the assumption "$\sigma_U^2 \leq 1$" is checked.

However, this is, for the moment, only one conjecture because not completely proved. There are thus two solutions.

1) Or one admits this conjecture which is very probably checked

2) Or it is noted that the results are checked with a probability infinitely close to 1 in the set of the possible probabilities : cf section 5.5.

Chapter 9

Empirical Theorems

9.1 Notations and assumptions

We keep the classical notations for the stochastics "O(.)" and "o(.)" (cf notation A.2.1. cf also [42] page 8, section 1.2.5). In particular, $X_n = O_P(1)$ if X_n converges in distribution to a random variable X.

Notations 9.1.1 *A sequence of random variable X_n is bounded in probability, if, for every $\epsilon > 0$, there exists M_ϵ and N_ϵ such that $P\{|X_n| \leq M_\epsilon\} \geq 1 - \epsilon$ for all $n \geq N_\epsilon$. Then, one writes $X_n = O_P(1)$.*

Moreover, we write $X_n = o_P(1)$ for sequence of random variable X_n if X_n converges in probability to 0.

In this chapter, we use the following notations.

Notations 9.1.2 *Let j_s , s=1,2,...,p, $j_s \in \mathbb{Z}$, be an injective sequence such that $j_1 = 0$. Let $d_0 = \big|min(j_s | s = 1, 2, .., p)\big|$.*

Notations 9.1.3 *Let $m \in \mathbb{N}^*$ be a fixed integer. Suppose ϵ fixed and statisfying $0 < \epsilon \leq 1/4$. Let $X_n^0 \in F(m)$, $n \in \mathbb{N}^*$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{A}, P)$.*

One supposes that, for all Borel set $Bo \subset F(m)$, for all $p \in \mathbb{N}^*$, for all sequence j_s, for all $x_2, ..., x_p$, for all $n \in \mathbb{N}^*$, such that $n > d_0$,

$$P\big\{X_n^0 \in Bo \big| X_{n+j_2}^0 = x_2, ..., X_{n+j_p} = x_p\big\} = L(Bo) + Ob(1)\epsilon \ .$$

Notations 9.1.4 *We set $X_n = X_{n+d_0}^0$.*

In this chapter, we shall suppose also that Bo is a fixed Borel set.

Notations 9.1.5 *Let $Bo = Bo_1 \otimes Bo_2 \otimes .... \otimes Bo_p \subset F(m)^p$ be a Borel set where $L(Bo_s) \leq 1/2$ for s=1,...,p.*

We set $1_{Bo}(X_n) = 1_{Bo_1}(X_{n+j_1}) 1_{Bo_2}(X_{n+j_2}) .... 1_{Bo_p}(X_{n+j_p})$.

Then, by proposition 4.2.2, for all Borel set Bo, for all p, for all sequence j_s,

$$E\big\{1_{Bo}(X_n)\big\} = \big[L(Bo_1) + Ob(1)\epsilon\big].....\big[L(Bo_p) + Ob(1)\epsilon\big] \ .$$

We deduce the following property.

Proposition 9.1.1 *For all $p \in \mathbb{N}^*$, we set $\epsilon_p = R_p(\epsilon) = (1/2+\epsilon)^p - (1/2)^p$. Then, for all Borel set Bo such that $L(Bo_s) \leq 1/2$ for s=1,...,p.*

$$E\big\{1_{Bo}(X_n)\big\} = L(Bo) + Ob(1)\epsilon_p \ .$$

Proof One can write

$$E\big\{1_{Bo}(X_n)\big\} - L(Bo) = \big(L(Bo_1) + \epsilon'_1\big).....\big(L(Bo_p) + \epsilon'_p\big) - L(Bo_1)....L(Bo_p)$$

where $|\epsilon'_s| \leq \epsilon$.

Then,

$$\big(L(Bo_1) + \epsilon'_1\big).....\big(L(Bo_p) + \epsilon'_p\big) - L(Bo_1)....L(Bo_p)$$

$$= L(Bo_1)\epsilon'_2....\epsilon'_p + L(Bo_2)\epsilon'_1\epsilon'_3.....\epsilon'_p + ...$$
$$+L(Bo_1)L(Bo_2)\epsilon'_3.....\epsilon'_p + L(Bo_1)L(Bo_3)\epsilon'_2\epsilon'_4.....\epsilon'_p + ...$$
$$+.....$$

$$\leq (1/2)\big[\epsilon'_2....\epsilon'_p + \epsilon'_1\epsilon'_3.....\epsilon'_p + ...\big]$$
$$+(1/4)\Big[\epsilon'_3.....\epsilon'_p + \epsilon'_2\epsilon'_4.....\epsilon'_p + ...\Big]$$
$$+.....$$

$$\leq (1/2)p\epsilon^{p-1} + (1/4)[p(p-1)/2]\epsilon^{p-2} + ... \leq (1/2+\epsilon)^p - (1/2)^p \quad . \blacksquare$$

In this chapter one will study the asymptotic behavior of the empirical probability of a Borel set $Bo \subset F(m)^p$.

Hypothesis 9.1.1 *In the three following sections 9.2 , 9.3 and 9.4, p means an fixed integer. Moreover, Bo means a fixed Borel set. Finally the sequence j_s is fixed.*

9.2 First Theorem

9.2.1 Definition of H(n,q)

At first, we have to define the set H(n,q).

Notations 9.2.1 *We set $j(e) = j_e$ for e=1,2,...,p. We denote by $c(1) < c(2) < .... < c(Q)$, $c(s) \in \mathbb{Z}$, the integers $c(\mu) = j(e) - j(e')$, where e and e' $\in \{1, 2, ..., p\}$.*

That is to say that, if $c(\mu) = m - n$, there exist e and e' such that $X_{n+j(e)} = X_{m+j(e')}$.

Then, there exists e and e' $\in \{1, 2, ..., p\}$ such that $c(\mu) + j(e') = j(e)$. That is to say $n + c(\mu) + j(e') = n + j(e)$. That is to say $m + j(e') = n + j(e)$.

Then, if $m \neq n + c(\mu)$ for all μ, $X_{n+j(e)} \neq X_{m+j(e')}$ for all e and e' $\in \{1, 2, ..., p\}$.

Notations 9.2.2 *For all n, we set $H(n) = \{m \in \mathbb{N}^* | \exists \mu : m = n + c(\mu)\}$, $H^*(n) = \{m \in \mathbb{N}^* | \exists \mu : m = n + c(\mu), m \neq n\}$.*

Notations 9.2.3 *Let $q \in \mathbb{N}^*$. For all n, we set*

$$H(n, q) = \{m \in N^* | \exists\ \mu : |n + c(\mu) - m| \leq q\}$$

and $H^*(n, q) = H(n, q) \setminus H(n)$.

9.2.2 Notations

We suppose that there exists a certain asymptotic independence.

Hypothesis 9.2.1 *For all $n \in \mathbb{N}^*$, let $L_n = L_n(Bo) = E\Big\{1_{Bo}(X_n)\Big\}$. Let $\delta_d = Max_{n \in \mathbb{N}^*} \Big| E\big\{\big(1_{Bo}(X_n) - L_n\big)\big(1_{Bo}(X_{n+d}) - L_{n+d}\big)\big\}\Big|$.*

We suppose that there exists $q \in \mathbb{N}^*$ and $k_B > 0$ such that $\sum_{n+d \notin H(n,q)} \delta_d \leq (k_B/10)\epsilon_p$.

For example, if there is a Qd-dependence, one can choose q = Qd : in this case $\sum_{n+d \notin H(n,q)} \delta_d = 0$.

Notations 9.2.4 *Let $N \in \mathbb{N}^*$. We set*

$$\sigma_1^2 = (1/N)E\Big\{\big[\sum_{n=1}^{N}\big(1_{Bo}(X_n) - L_n\big)\big]^2\Big\} ,$$

$$\sigma_B^2 = \sigma_B^2(Bo) = (1/N)E\Big\{ \big[\sum_{n=1}^{N} \big(1_{Bo}(X'_n) - L(Bo)\big)\big]^2 \Big\} \ ,$$

when X'_n is an IID random sequence which have the uniform distribution on $F(m)$.

For example, if p=1, $\sigma_B^2 = L(Bo)[1 - L(Bo)]$. Remark that $L(Bo) = N(Bo)/m$.

9.2.3 Wording of the first theorem

Now, we can expound the first empirical theorem.

Theorem 9 *Let $A(p) = 1 - (p^2 - p + 1)2^{-p}$. Let $L^N(Bo) = (1/N)\sum_{n=1}^{N} L_n$. Let*

$$\beta_{1,p} = \sqrt{N}[L^N(Bo) - L(Bo)]/\sigma_B(Bo) \ .$$

We set

$$\gamma_{1,p} = \frac{1}{2A(p)L(Bo)}\Big[(p^2-p+1)\Big(\epsilon_p + 2q\epsilon_{2p} + (1+2q)\big[2^{1-p}\epsilon_p + \epsilon_p^2\big]\Big) + k_B 0.1\epsilon_p\Big] \ .$$

Let $P_e \ = \frac{1}{N}\sum_{n=1}^{N} 1_{Bo}(X_n)$. Then, the following inequality holds

$$P\left\{ \sqrt{N}\big|P_e - L(Bo)\big| \geq \sigma_B(Bo)x \right\} \leq K_1\Big(\frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}}x\Big) \ ,$$

where $K_1(x) = P\Big\{ \frac{\sqrt{N}|P_e - L^N(Bo)|}{\sigma_1(Bo)} \geq x \Big\}$.

For example, if X_n is Qd-dependent, $\frac{\sqrt{N}(P_e - L^N(Bo))}{\sigma_1(Bo)}$ has asymptotically the distribution N(0,1).

9.2.4 Proof of theorem 9

First, we suppose ϵ is small enough. Then, the following lemma is needed.

Lemma 9.2.1 *We suppose ϵ is small enough. In particular, we suppose $\epsilon \leq 1/4$. Then, $\epsilon_p \geq \epsilon_{p+1}$ if $p \geq 2$ and $\epsilon_p < \epsilon_{p+1}$ if $p = 1$.*

Proof We have $R_p(\epsilon) = exp\big(p.log(1/2+\epsilon)\big) - exp\big(p.log(1/2)\big)$.
Therefore,

$$\frac{d\big(R_p(\epsilon)\big)}{dp} = log(1/2+\epsilon)exp\big(p.log(1/2+\epsilon)\big) - log(1/2)exp\big(p.log(1/2)\big) \ ,$$

which is equal to 0 if and only if

$$log(1/2+\epsilon)exp\big(p.log(1/2+\epsilon)\big) = log(1/2)exp\big(p.log(1/2)\big) \ ,$$

i.e.

$$exp\big(p.log(1/2+\epsilon) - p.log(1/2)\big) = \frac{log(1/2)}{log(1/2+\epsilon)} \ ,$$

i.e.

$$exp\big(p\big[log(1/2+\epsilon) - log(1/2)\big]\big) = \frac{log(1/2)}{log(1/2+\epsilon)} \ ,$$

which has an alone zero

$$p\big[log(1/2+\epsilon) - log(1/2)\big] = log\Big(\frac{log(1/2)}{log(1/2+\epsilon)}\Big) \ ,$$

that is to say

$$p = \frac{log\Big(\frac{log(1/2)}{log(1/2+\epsilon)}\Big)}{\big[log(1/2+\epsilon) - log(1/2)\big]} = \frac{log\Big(\frac{log(1/2)}{log((1/2)[1+2\epsilon])}\Big)}{\big[log((1/2)[1+2\epsilon]) - log(1/2)\big]}$$

$$\approx \frac{log\Big(\frac{log(1/2)}{Log(1/2)+log(1+2\epsilon)}\Big)}{log(1+2\epsilon)} \approx \frac{log\Big(\frac{log(1/2)}{Log(1/2)+2\epsilon}\Big)}{2\epsilon} \approx \frac{log\Big(\frac{1}{1+2\epsilon/log(1/2)}\Big)}{2\epsilon}$$

$$\approx \frac{-2\epsilon/log(1/2)}{2\epsilon} = 1/log(2) = 1.4427 \ .$$

Now $R_p(\epsilon) \approx \frac{2p\epsilon}{2^p}$ if $p \geq 2$. For example for $p = 3 > 1.4427$, $R_p(\epsilon) \geq R_{p+1}(\epsilon)$. Then, $\frac{d\big(R_p(\epsilon)\big)}{dp} < 0$ if $p > 1.4427$. Then, $R_p(\epsilon)$ is decreasing if $p > 1.4427$ (and also $p > 2$).

If p=2, $R_p(\epsilon) = \epsilon + \epsilon^2$ and $R_{p+1}(\epsilon) = 3\epsilon/2^2 + 3\epsilon^2/2 + \epsilon^3$.

Therefore,

$$R_p(\epsilon) - R_{p+1}(\epsilon) = \epsilon + \epsilon^2 - 3\epsilon/4 - 3\epsilon^2/2 - \epsilon^3$$

$$= (1/4)\epsilon - (1/2)\epsilon^2 - \epsilon^3 = (1/4)\epsilon\big(1 - 2\epsilon - 4\epsilon^2\big) = -(1/4)\epsilon\big(4\epsilon^2 + 2\epsilon - 1\big) \ ,$$

where $4x^2 + 2x - 1 = 0$ has the following zero : $x = \frac{-2\pm\sqrt{4-4(-1*4)}}{2*4} = \frac{-2\pm\sqrt{4+16}}{8} = \frac{-1\pm\sqrt{1+4}}{4} = \frac{-1\pm\sqrt{5}}{4} \approx \frac{-1\pm2,25}{4} = 0{,}312$ or = -0,812.

Therefore, if $0 < \epsilon \leq 0.25 \leq 0.312$, $R_p(\epsilon) - R_{p+1}(\epsilon) \geq 0$ if $p = 2$.

Therefore, $R_p(\epsilon) \geq R_{p+1}(\epsilon)$ if $p \geq 2$. ■

We deduce from notation 9.2.1, the following property.

Lemma 9.2.2 *With the previous notations,*

$$\{c(1), c(2), ..., c(Q)\} = \{j(e) - j(e') | (e, e') \in \{1, 2, ..., p\}^2\} \ ,$$

$$\{c(1), c(2), ..., c(Q) | c(\mu) \neq 0\} = \{j(e) - j(e') | (e, e') \in \{1, 2, ..., p\}^2, e \neq e'\} \ .$$

We deduce the following lemma.

Lemma 9.2.3 *The following inequalities hold :*
$card\big(H^*(n)\big) \leq p^2 - p$,
$card\big(H(n)\big) \leq p^2 - p + 1$,
$card\big(H(n,q)\big) \leq (1 + 2q)(p^2 - p + 1)$,
$card\big(H^*(n,q)\big) \leq 2q(p^2 - p + 1)$.

Lemma 9.2.4 *The following inequality holds :* $\sigma_B^2 \geq A(p)L(Bo)$.

Proof One can write

$$\sigma_B^2 = (1/N) E\Big\{ \Big[\sum_{n=1}^{N} \big[1_{Bo}(X'_n) - L(Bo) \big] \Big]^2 \Big\}$$

$$= (1/N) \sum_{n=1}^{N} \sum_{m \in H(n)} \Big(E\big\{ 1_{Bo}(X'_n) 1_{Bo}(X'_m) \big\} - L(Bo)^2 \Big)$$

$$= (1/N) \sum_{n=1}^{N} \Big(E\big\{ 1_{Bo}(X'_n) \big\} + \sum_{m \in H^*(n)} E\big\{ 1_{Bo}(X'_n) 1_{Bo}(X'_m) \big\} - \sum_{m \in H(n)} L(Bo)^2 \Big)$$

$$\geq (1/N) \sum_{n=1}^{N} \Big(E\big\{ 1_{Bo}(X'_n) \big\} - (p^2 - p + 1) L(Bo)^2 \Big)$$

$$= L(Bo) \Big(1 - (p^2 - p + 1) L(Bo) \Big) \ . \ ■$$

Lemma 9.2.5 *The following inequalities hold :* $\sigma_B^2 \geq A(p)L(Bo) \geq L(Bo)/8$.

Proof At first, $(p^2-p+1)L(Bo) \leq (p^2-p+1)2^{-p}$. Now,

$$\frac{d(p^2-p+1)2^{-p}}{dp} = (2p-1)2^{-p} - Log(2)(p^2-p+1)2^{-p}$$

$$= 2^{-p}\big[(2p-1) - Log(2)p^2 + Log(2)p - Log(2)\big]$$

$$= -2^{-p}\Big[Log(2)p^2 - \big[2+Log(2)\big]p + \big[1+Log(2)\big]\Big]$$

which has the following roots

$$\frac{\big[2+Log(2)\big] \pm \sqrt{\big[2+Log(2)\big]^2 - 4Log(2)\big[1+Log(2)\big]}}{2Log(2)}$$

$$= \frac{2.6931 \pm \sqrt{2.6931^2 - 4*0.6931*1.6931}}{2*0.6931} = \frac{2.6931 \pm \sqrt{2.5586}}{2*0.6931}$$

$$= \frac{2.6931 \pm 1.5996}{2*0.6931} = 0.7888 \quad ou \quad 3.5423.$$

Therefore, $(p^2-p+1)2^{-p}$ decreases and converges to 0 if $p \geq 4$. Moreover, $(p^2-p+1)2^{-p} = 3/4$ if p=2, 7/8 if p=3, 13/16 if p=4, 21/32 if p=5.

Then, $(p^2-p+1)L(Bo) \leq 7/8$. ■

Lemma 9.2.6 *If* $m \in H^*(n,q)$,

$$E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} = L(Bo_1)^2.....L(Bo_p)^2 + Ob(1)\epsilon_{2p},$$

If $m \in H(n)$

$$E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} = E\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\} + Ob(1)\epsilon_p \ .$$

Proof The first equality results from proposition 9.1.1.

Let us study the second equality.
If n=m, that results also from proposition 9.1.1.
If p=1, it is the case n=m : $1_{Bo}(X_n)1_{Bo}(X_m) = 1_{Bo'}(X_n)$, where $Bo' \subset Bo$.

Suppose $p \geq 2$ and $n \neq m$. One can assume $n < m$.
Then, there exists a sequence i_s, s=1,...,p', $p' < 2p$, and a sequence of Borel sets Bo'_s, s=1,...,p', such that

$$1_{Bo}(X_n)1_{Bo}(X_m)$$

$$= 1_{Bo_1}(X_n)1_{Bo_2}(X_{n+j_2}).....1_{Bo_p}(X_{n+j_p})1_{Bo_1}(X_m)1_{Bo_2}(X_{m+j_2}).....1_{Bo_p}(X_{n+j_p})$$
$$= 1_{Bo'_1}(X_n)1_{Bo'_2}(X_{n+i_2}).....1_{Bo'_{p'}}(X_{n+i_{p'}}) \ .$$

Therefore

$$1_{Bo}(X_n)1_{Bo}(X_m) = 1_{Bo'_1}(X_n)1_{Bo'_2}(X_{n+i_2}).....1_{Bo'_{p'}}(X_{n+i_{p'}}) \ . \tag{9.1}$$

Clearly $p \leq p' < 2p$. Therefore, because $p \geq 2$, $E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} = E\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\} + Ob(1)\epsilon_{p'} = E\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\} + Ob(1)\epsilon_p$, considering $\epsilon_p \geq \epsilon_{p'}$ (cf proposition 9.1.1). ■

Lemma 9.2.7 *The following equality holds*

$$L_nL_m = L(Bo_1)^2.....L(Bo_p)^2 + 2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2.$$

Proof We have

$$L_nL_m = \Big[L(Bo_1)....L(Bo_p) + Ob(1)\epsilon_p\Big]\Big[L(Bo_1).....L(Bo_p) + Ob(1)\epsilon_p\Big]$$
$$= L(Bo_1)^2.....L(Bo_p)^2 + 2 * 2^{-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2 \ . \ ■$$

Lemma 9.2.8 *The following equality holds*

$$\sigma_1^2 = \sigma_B^2\big[1 + Ob(1)2\gamma_{1,p}\big].$$

Proof Let X'_n be an IID sequence with uniform distribution. Then,

$$\sigma_1^2 = (1/N)E\Big\{\big[\sum_{n=1}^{N}\big(1_{Bo}(X_n) - L_n\big)\big]^2\Big\}$$

$$= (1/N)E\Big\{\sum_{n=1}^{N}\sum_{m=1}^{N}\big(1_{Bo}(X_n) - L_n\big)\big(1_{Bo}(X_m) - L_m\big)\Big\}$$

$$= (1/N)E\Big\{\sum_{n=1}^{N}\sum_{m\in H(n,q)}\big(1_{Bo}(X_n) - L_n\big)\big(1_{Bo}(X_m) - L_m\big)\Big\}$$

$$+(1/N)E\Big\{\sum_{n=1}^{N}\sum_{m\notin H(n,q)}\big(1_{Bo}(X_n) - L_n\big)\big(1_{Bo}(X_m) - L_m\big)\Big\}$$

$$= (1/N)E\Big\{\sum_{n=1}^{N}\sum_{m\in H(n,q)}\big(1_{Bo}(X_n)-L_n\big)\big(1_{Bo}(X_m)-L_m\big)\Big\}$$

$$+(1/N)E\Big\{\sum_{n=1}^{N}\sum_{n+d\notin H(n,q)}\big(1_{Bo}(X_n)-L_n\big)\big(1_{Bo}(X_{n+d})-L_{n+d}\big)\Big\}$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\}-L_nL_m\Big)$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\in H^*(n,q)}\Big(E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\}-L_nL_m\Big)$$

$$+2Ob(1)\sum_{d=q+1}^{\infty}\delta_d$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(E\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\}-L(Bo)^2\Big)+\Phi$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\in H^*(n,q)}Ob(1)\epsilon_{2p}$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\in H(n,q)}\big[2^{1-p}Ob(1)\epsilon_p+Ob(1)\epsilon_p^2\big]$$

$$+(Ob(1)k_B/10)\epsilon_p\ ,$$

where

$$\Phi = (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big[E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\}-E\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\}\Big]\ . \quad (9.2)$$

Then, $\Phi \leq (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)} Ob(1)\epsilon_p$.

Then,

$$\sigma_1^2$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(E\big\{1_{Bo}(X_n')1_{Bo}(X_m')\big\} - L(Bo)^2\Big)$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\in H(n)} Ob(1)\epsilon_p$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\in H^*(n,q)} Ob(1)\epsilon_{2p}$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\in H(n,q)}\big[2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2\big]$$

$$+(Ob(1)k_B/10)\epsilon_p$$

$$= \sigma_B^2 + (p^2-p+1)(Ob(1)\epsilon_p + (p^2-p+1)(2q)Ob(1)\epsilon_{2p}$$
$$+(p^2-p+1)(1+2q)Ob(1)\big[2^{1-p}\epsilon_p + \epsilon_p^2\big] + (Ob(1)k_B/10)\epsilon_p$$

$$= \sigma_B^2\big[1 + Ob(1)2\gamma_{1,p}A(p)L(Bo)/\sigma_B^2\big] = \sigma_B^2\big[1 + Ob(1)2\gamma_{1,p}\big]\ . \blacksquare$$

Lemma 9.2.9 *The following inequality holds :*

$$\sigma_1 \leq (1+\gamma_{1,p})\sigma_B.$$

Proof Let $e \mapsto \psi(e)$ be the function defined by $\sqrt{1+2e} = 1 + e + \psi(e)e^2$ if $e > 0$, that is $\psi(e) = \frac{\sqrt{1+2e}-1-e}{e^2}$. This function has the following derivative:

$$\frac{\big(2(1/2)(1+2e)^{-1/2}-1\big)e^2 - \big(\sqrt{1+2e}-1-e\big)2e}{e^4}.$$

It is equal to 0 for

$$e(1+2e)^{-1/2} - e - 2\sqrt{1+2e} + 2 + 2e = 0.$$

That is to say

$$e - 2(1+2e) + 2\sqrt{1+2e} + e\sqrt{1+2e} = 0,$$

or

$$(2+e)\sqrt{1+2e} = 2 + 3e \ ,$$

or

$$(4 + e^2 + 4e)(1+2e) = 4 + 9e^2 + 12e,$$

or

$$4 + e^2 + 4e + 8e + 2e^3 + 8e^2 = 4 + 9e^2 + 12e,$$

or

$$2e^3 + 9e^2 + 12e + 4 = 9e^2 + 12e + 4,$$

that is $e = 0$.

Moreover, in $e = 4$, the value of this derivative is $\frac{4/\sqrt{9}-4-2\sqrt{9}+2+8}{4^3} = \frac{4/3}{4^3} > 0$. Therefore, ψ is increasing for $e > 0$. In 4 its value is $\frac{\sqrt{9}-1-4}{4^2} = \frac{-2}{4^3} < 0$. Moreover, $\psi(e) = \frac{\sqrt{1+2e}-1-e}{e^2} \to 0$ as $e \to \infty$. Therefore $\psi < 0$ if $e > 0$. Therefore $\sqrt{1+2e} < 1 + e$.

Therefore, $\sigma_1 = \sigma_B\sqrt{1 + 2Ob(1)\gamma_{1,p}} \leq \sigma_B\sqrt{1+2\gamma_{1,p}} \leq \sigma_B(1+\gamma_{1,p})$. ■

Proof 9.2.10 *We prove now the theorem 9*

The following inequalities hold.

$$P\left\{\left|\sqrt{N}\big[P_e - L(Bo)\right| > \sigma_B x\right\}$$

$$\leq P\left\{\left|\sqrt{N}\big[P_e - L^N(Bo)\big]\right| > \sigma_B x - \sqrt{N}|L^N - L(Bo)|\right\}$$

$$\leq P\left\{\left|\sqrt{N}\big[P_e - L^N(Bo)\big]\right| > \sigma_B x[1 - \beta_{1,p}/x]\right\}$$

$$\leq P\left\{\left|\sqrt{N}\big[P_e - L^N(Bo)\big]\right| > \frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}}\sigma_1 x\right\}$$

$$= K_1\Big(\frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}}x\Big) \ . \ \blacksquare$$

9.3 Second Theorem

9.3.1 Notations and assumptions

In this section, we use a Borel set, $Bo \otimes J$, where $Bo = Bo_1$, $J = J_2 \otimes J_3 \otimes$ ..... $\otimes J_p$, and $J_s = Bo_s$ for s=2,3,...,p.

Notations 9.3.1 *Let $p \in \mathbb{N}^*$. Let $N \in \mathbb{N}^*$. Let $Bo \otimes J = Bo \otimes J_2 \otimes J_3 \otimes$ ..... $\otimes J_p$ where Bo and the J_s, s=2,...,p, are Borel sets. We suppose that $L(J_s) = L(Bo) = 1/2$ for s=2,...,p.*

We set

$1_J(X_{n+j}) = 1_{J_2}(X_{n+j(2)})....1_{J_p}(X_{n+j(p)})$, $1_{Bo\otimes J}(X_n) = 1_{Bo}(X_n)1_J(X_{n+j})$,

$L_n = E\Big\{1_{Bo\otimes J}(X_n)\Big\}$, $l_n = E\Big\{1_J(X_{n+j})\Big\}$,

$L^N(Bo \otimes L) = (1/N)\sum_{n=1}^N L_n$, $l^N(J) = (1/N)\sum_{n=1}^N l_n$.

The following assumptions are admitted.

Hypothesis 9.3.1 *Let $D_n = E\big\{[1_{Bo}(X_n) - L(Bo)]1_J(X_{n+j})\big\}$. Let*

$$\delta'_d = \frac{Max_{n\in N^*}\left\{\left|E\left\{\Big[1_{Bo}(X_n)-L(Bo)\Big]\Big[1_{bo}(X_{n+d})-L(Bo)\Big]1_J(X_{n+j})1_J(X_{n+d+j})\right\}-D_nD_{n+d}\right|\right\}}{l^N(J)}.$$

One supposes that there exists $q \in \mathbb{N}^*$ and $K_B > 0$ such that, for all $n \in \mathbb{N}^*$, $\sum_{n+d\notin H(n,q)} \delta'_d \leq (K_B/10)\epsilon$.

Hypothesis 9.3.2 *Let $\delta_d = Max_{n\in N^*}\Big|E\big\{\big(1_J(X_n) - l_n\big)\big(1_J(X_{n+d}) - l_{n+d}\big)\big\}\Big|$. One supposes that there exists $q \in \mathbb{N}^*$ and $k_B > 0$ such that $\sum_{n+d\notin H(n,q)} \delta_d \leq (k_B/10)\epsilon_{p-1}$.*

Notations 9.3.2 *One simplifies $l^N(J)$ in l : $l = l^N(J)$. We set $\epsilon\mathcal{M}(l) = 1/L(J) - 1/l$.*

Let

$$\xi_p = (1/2)\Bigg[2L(J)\mathcal{M}(l) + \frac{8L(J)}{l}(1 - 2L(I)) + (3/2)(p-1)L(J)\mathcal{M}(l)$$

$$+\frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]} + 13\Big] + \frac{8[p(p-1)+K_B/10]L(J)}{l}$$

$$+\ \big[1+\epsilon\big]\frac{8L(J)[L(J)+\epsilon_{p-1}]}{l}(2q+1)(p^2-p+1)\Bigg]\ .$$

Let $\sigma_B(J)^2 = (1/N)E\Big\{\Big(\sum_{n=1}^N [1_J(X'_n) - L(J)]\Big)^2\Big\}$ *when* X'_n *is IID. Let* $A(p) = 1 - (p^2 - p + 1)2^{-p}$. *Then, we set*

$$\theta_p = \frac{8\sigma_B(J)^2 L(J)}{l^2}$$
$$+(p^2 - 3p + 3)\frac{8\epsilon_{p-1}L(J)}{l^2} + (p^2 - 3p + 3)(2q)\frac{8\epsilon_{2p-2}L(J)}{l^2}$$
$$+(p^2 - 3p + 3)(1 + 2q)\frac{8\big[2^{2-p}\epsilon_{p-1} + \epsilon_{p-1}^2\big]L(J)}{l^2} + \frac{0.8k_B\epsilon_{p-1}L(J)}{l^2} \ .$$

We set

$$2\gamma_{2,p} = \epsilon\Big(2\xi_p + 2[1 + \epsilon\xi_p]\sqrt{\theta_p} + \epsilon\theta_p\Big) \ .$$

Moreover the following assumption will be admitted.

Hypothesis 9.3.3 *One supposes that*

$$NE\big\{(P_e - L(Bo)p_e)^2\big\} \geq (1/4)L(Bo)L(J)$$

if X_n *is IID.*

In fact, this assumption probably holds in all the cases when $X_n \in \{0, 1\}$: cf section 9.6.

9.3.2 Wording of the second theorem

Now, we have the second emprical theorem.

Theorem 10 *Let* $p \in \mathbb{N}^*$. *Let* $N \in \mathbb{N}^*$. *We assume that* ϵ *is small enough and* p *is not too large*

Let $D = L^N(Bo \otimes J) - L(Bo)l^N(J)$. *One supposes that*

$$\sqrt{N}\frac{P_e - L(Bo)p_e - D}{l} = O_P(1) \ ,$$
$$\sqrt{N}(p_e - l^N(J)) = O_P(1) \ ,$$
$$p_e \xrightarrow{P} l \quad as \ N \to \infty \ ,$$
$$l \to L(J) \quad as \ N \to \infty \ ,$$

$$L \to L(Bo)L(J) \ \ as\ N \to \infty\ ,$$
$$P_e - L^N(Bo \otimes J) \xrightarrow{P} 0 \ \ as\ N \to \infty\ .$$

We keep the assumptions of section 9.1. One supposes that the assumptions 9.3.1, 9.3.2 and 9.3.3 hold.

We set $p_e \ = \frac{1}{N}\sum_{n=1}^{N} 1_J(X_n)$, $P_e \ = \frac{1}{N}\sum_{n=1}^{N} 1_{Bo\otimes J}(X_n)$.

We denote by σ_2^2 *the variance of* $\sqrt{N}\Big[\frac{P_e - L(Bo)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big]$.

We denote σ_{cp}^2 *instead of* σ_2^2 *when* X_n *is IID*

Let $\beta_{2,p} = \frac{N^{1/2}D}{\sigma_{cp} l^N(J)} + \eta$ *where* $\eta > 0$.

One supposes that $\sqrt{N}H_e = \sqrt{N}\Big(\frac{P_e - L(Bo)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big)$ *converges in distribution to a random variable which has the distribution function* $F_{H_e}(x) > 0$.

Then, for all $\eta > 0$, *there exists* $N_0 \in \mathbb{N}$ *such that, for all* $N \geq N_0$,

$$P\left\{\sqrt{N}\left|\frac{P_e}{p_e} - L(Bo)\right| > \sigma_{cp}\ x\right\} \leq K_2\Big(\frac{1 - \beta_{2,p}/x}{1 + \gamma_{2,p}}x\Big)\ ,$$

where $K_2(x) = P\Big\{\frac{\sqrt{N}}{\sigma_2}\Big|\frac{P_e - L(Bo)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big| > x\Big\}$.

For example, if the X_n's have the same distribution and if X_n is a sequence Qd-dependent, $\frac{\sqrt{N}}{\sigma_2}\Big[\frac{P_e - L(Bo)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big]$ has asymptotically the distribution N(0,1).

9.3.3 Proof of theorem 10

To clear up the notations, in certain lemmas, one will pose $Bo = I$. In this proof we simplify $L^N(Bo \otimes J)$ in L and $l^N(J)$ in l : $L = L^N(I \otimes J)$, $l = l^N(J)$.

9.3.4 Lemmas of introduction

Notations 9.3.3 *For all n, we denote by*

$$G(n) = \{m \mid \exists\ s, s' : m = n + j_s \ and\ n = m + j_{s'}\}\ .$$

We set $G^*(n) = G(n) \setminus \{0\}$.

Lemma 9.3.1 *The following increase holds : card*$(G^*(n)) \leq p-1$.

Proof If $m \in G^*(n)$, $n = m + j_{s_0}$ and $m = n + j_{s_1}$. Then, $j_{s_0} = -j_{s_1}$. Therefore, for any n, there is at the most p-1 "$m = n + j_s$" such that there exists s_0 et s_1 satisfying $j_{s_0} = -j_{s_1}$ and $s_0 \neq 0$. ■

Lemma 9.3.2 *Let* $n = m + j_{s_0}$ *and* $m = n + j_{s_1}$. *We set*
$J^1 = J_2 \otimes J_3 \otimes ..... \otimes J_{s_1-1} \otimes J_{s_1+1} \otimes ..... \otimes J_p$,
$J^0 = J_2 \otimes J_3 \otimes ..... \otimes J_{s_0-1} \otimes J_{s_0+1} \otimes ..... \otimes J_p$.
We denote by j^1 *the sequence* $j^1(s) = \{j_2, j_3, ..., j_{s_1-1}, j_{s_1+1}, ..., j_p\}$ *and by* j^0 *the sequence* $j^0(s) = \{j_2, j_3, ..., j_{s_0-1}, j_{s_0+1}, ..., j_p\}$.

Then, one sets

$$1_{J^1}(X_{n+j^1}) =$$

$$1_{J_2}(X_{n+j_2})1_{J_3}(X_{n+j_3})....1_{J_{s_1-1}}(X_{n+j_{s_1-1}})1_{J_{s_1+1}}(X_{n+j_{s_1+1}}).....1_{J_p}(X_{n+j_p}).$$

$$1_{J^0}(X_{n+j^0}) =$$

$$1_{J_2}(X_{n+j_2})1_{J_3}(X_{n+j_3})....1_{J_{s_0-1}}(X_{n+j_{s_0-1}})1_{J_{s_0+1}}(X_{n+j_{s_0+1}}).....1_{J_p}(X_{n+j_p}).$$

Then,

$$E\Big\{[1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j})\Big\}$$

$$= \Big[L(I\cap J_{s_0}) - L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1}) - L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+(1/16)Ob(1)\epsilon_{p-1} + Ob(1)(3/8)(1+2\epsilon)L(J) + (3/2)\epsilon Ob(1)l_n \ ,$$

where X'_n *is an IID sequence.*

Proof We have

$$E\Big\{[1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j})\Big\}$$

$$= E\Big\{[1_I(X_n) - L(I)]1_{J_{s_0}}(X_{m+j_{s_0}})[1_I(X_m) - L(I)]1_{J_{s_1}}(X_{n+j_{s_1}})1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$= E\Big\{[1_I(X_n) - L(I)]1_{J_{s_0}}(X_n)[1_I(X_m) - L(I)]1_{J_{s_1}}(X_m)1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$= E\Big\{\Big[1_{I\cap J_{s_0}}(X_n) - L(I)1_{J_{s_0}}(X_n)\Big]\Big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\Big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$= E\Big\{1_{I\cap J_{s_0}}(X_n)\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^2}(X_{m+j^2})\Big\}$$

$$-L(I)E\Big\{1_{J_{s_0}}(X_n)\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$= L(I\cap J_{s_0})E\Big\{\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$+Ob(1)\epsilon E\Big\{\big|\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\big|\Big\}$$

$$-L(I)L(J_{s_0})E\Big\{\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$-L(I)Ob(1)\epsilon E\Big\{\big|\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\big|\Big\}\ ,$$

(by lemma 4.2.1), and where, because $m = n + j_{s_1}$,

$$\begin{aligned}&\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\\ &\le 1_{I\cap J_{s_1}}(X_m)1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\\ &\le 1_{I\cap J_{s_1}}(X_{n+j_{s_1}})1_{J^1}(X_{n+j^1}) \le 1_J(X_{n+j})\ ,\end{aligned}$$

and where

$$\begin{aligned}&-\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\\ &\le L(I)1_{J_{s_1}}(X_m)1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0}) \le L(I)1_J(X_{n+j})\ .\end{aligned}$$

Therefore,

$$\big|\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\big| \le 1_J(X_{n+j})\ .$$

Therefore,

$$E\Big\{\big|\big[1_{I\cap J_{s_1}}(X_m) - L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\big|\Big\} \le l_n\ .$$

Therefore,

$$E\Big\{[1_I(X_n)-L(I)][1_I(X_m)-L(I)]1_J(X_{n+j})1_J(X_{m+j})\Big\}$$

$$= L(I\cap J_{s_0})E\Big\{\big[1_{I\cap J_{s_1}}(X_m)-L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$-L(I)L(J_{s_0})E\Big\{\big[1_{I\cap J_{s_1}}(X_m)-L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+(3/2)\epsilon Ob(1)l_n$$

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]E\Big\{\big[1_{I\cap J_{s_1}}(X_m)-L(I)1_{J_{s_1}}(X_m)\big]1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+(3/2)\epsilon Ob(1)l_n$$

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]E\Big\{1_{I\cap J_{s_1}}(X_m)1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$-\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]E\Big\{L(I)1_{J_{s_1}}(X_m)1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+(3/2)\epsilon Ob(1)l_n$$

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]L(I\cap J_{s_1})E\Big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+Ob(1)\epsilon\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]E\Big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$

$$-\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]L(I)L(J_{s_1})E\Big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+Ob(1)\epsilon\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]L(I)E\Big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+(3/2)\epsilon Ob(1)l_n\ .$$

There exists "t" such that

$$E\big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\big\}\le E\big\{1_{J^1}(X_{n+j^1})1_{J_t}(X_{m+j_t})\big\}$$
$$=\big(L(J_t)+Ob(1)\epsilon\big)E\big\{1_{J^1}(X_{n+j^1})\big\}=\big(1/2+Ob(1)\epsilon\big)2^{2-p}$$
$$=\big(1+Ob(1)2\epsilon\big)L(J)\ .$$

Therefore,

$$E\Big\{[1_I(X_n)-L(I)][1_I(X_m)-L(I)]1_J(X_{n+j})1_J(X_{m+j})\Big\}$$

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+Ob(1)(3/2)\epsilon\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\big(1+2\epsilon\big)L(J)$$
$$+(3/2)\epsilon Ob(1)l_n$$

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X_{n+j^1})1_{J^0}(X_{m+j^0})\Big\}$$
$$+Ob(1)(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon Ob(1)l_n$$

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]Ob(1)\epsilon_{p'-1}$$
$$+Ob(1)(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon Ob(1)l_n$$

(where $p'-1\geq p-1\geq 2$ because $s_0\neq s_1$)

$$= \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+(1/16)Ob(1)\epsilon_{p-1}+Ob(1)(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon Ob(1)l_n \ . \ \blacksquare$$

On the other hand, by using lemma 9.3.1, one obtains the following result.

Lemma 9.3.3 *Let*

$$\mathcal{L}(p)=N(p-1)\frac{(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l}{l^2\sigma^2_{PC}} \ .$$

Then,

$$\sum_n\sum_{m\in G^*(n)}\frac{(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l_n}{l^2\sigma^2_{PC}}\leq\mathcal{L}(p) \ .$$

Lemma 9.3.4 *If ϵ is small enough, if p is not too great,*

$$\mathcal{L}(p) \leq N\frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big] \ .$$

Proof We have

$$l = L(J)+Ob(1)\epsilon_{p-1} \approx L(J)-2(p-1)Ob(1)\epsilon/2^{p-1} = L(J)\big[1-2Ob(1)(p-1)\epsilon\big] \ .$$

Therefore, by hypothesis 9.3.3 ,

$$N(p-1)\frac{(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l}{l^2\sigma_{PC}^2}$$

$$\leq N(p-1)\frac{(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l}{l^2(1/4)L(I)L(J)/L(J)^2}$$

$$\leq 8N(p-1)L(J)\frac{(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l}{l^2}$$

$$\approx 8N(p-1)L(J)\frac{(1/16)(p-1)\epsilon/2^{p-2}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l}{l^2}$$

$$\leq N(p-1)L(J)\epsilon\frac{(p-1)L(J)+3(1+2\epsilon)L(J)+12l}{l^2}$$

$$\approx N(p-1)L(J)\epsilon\Big[\frac{(p-1)L(J)+3(1+2\epsilon)L(J)}{L(J)^2\big[1-2(p-1)\epsilon\big]^2}+\frac{12l}{lL(J)\big[1-2(p-1)\epsilon\big]}\Big]$$

$$\leq N(p-1)\epsilon\Big[\frac{(p-1)+3+6\epsilon}{\big[1-2(p-1)\epsilon\big]^2}+\frac{12}{\big[1-2(p-1)\epsilon\big]}\Big]$$

$$= N\frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+12\Big] \ .$$

Then, one suppresses the approximation with the increase

$$\mathcal{L}(p) \leq N\frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big] \ . \ \blacksquare$$

9.3.5 Study of $E\{(P_e - L(I)p_e - D)^2\}$.

Now, one need some properties. At first, one reminds the following lemma : cf lemma 4.2.1.

Lemma 9.3.5 *Let f be a measurable function defined on* $(X_1, X_2, ..., X_{n-1}, X_{n+1}, ..., X_N)$. *Then,*

$$E\big\{1_J(X_n)f(X)\big\} = L(J)E\big\{f(X)\big\} + Ob(1)\epsilon E\big\{|f(X)|\big\}.$$

We deduce the following lemmas.

Lemma 9.3.6 *The following equality holds :*

$$E\big\{\big(1_J(X_n) - L(J)\big)f(X)\big\} = Ob(1)\epsilon E\big\{|f(X)|\big\}.$$

Lemma 9.3.7 *The following equalities hold :*

$$D_n = Ob(1)\epsilon l_n,$$

$$D = Ob(1)\epsilon l.$$

Proof We have $D_n = E\big\{\big(1_I(X_n) - L(I)\big)1_J(X_{n+j})\big\} = Ob(1)\epsilon E\big\{1_J(X_{n+j})\big\} = Ob(1)\epsilon l_n$.

Moreover $D = (1/N)\sum_n D_n = Ob(1)\epsilon \sum_n l_n = Ob(1)\epsilon l$. ■

Lemma 9.3.8 *The following equality holds :*

$$E\{1_{Bo}(X_n)1_J(X_{n+j}\} = \big[L(Bo) + Ob(1)\epsilon\big]E\{1_J(X_{n+j}\} \ .$$

Proof By lemma 9.3.5,

$$E\{1_{Bo}(X_n)1_J(X_{n+j}\} = \big[L(Bo) + Ob(1)\epsilon\big]E\{1_J(X_{n+j}\} \ . \ ■$$

Lemma 9.3.9 *If* $m \notin H(n)$, *the following equality hold*

$$E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\} = Ob(1)\big[L(J) + \epsilon_{p-1}\big]E\big\{1_J(X_{n+j})\big\}.$$

Proof By proposition 4.2.2 and by lemma 9.3.8,

$$E\big\{1_J(X_{n+j})1_J(X_{n+j})\big\} = (L(J_2) + Ob(1)\epsilon).....(L(J_p) + Ob(1)\epsilon)E\big\{1_J(X_{n+j})\big\} \ .$$

Now,

$$(L(J_2) + Ob(1)\epsilon).....(L(J_p) + Ob(1)\epsilon) = (1/2 + Ob(1)\epsilon).....(1/2 + Ob(1)\epsilon)$$

$$\leq (1/2 + \epsilon)^{p-1} = (1/2)^{p-1} + \epsilon_{p-1} = L(J) + \epsilon_{p-1} \ . \ ■$$

Then, by using the previous lemmas, one obtains the following property.

Lemma 9.3.10 *The following equality holds :*

$$N^2E\{(P_e - L(I)p_e - D)^2\}$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon) + L(I)^2\big][Nl]$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0}) - L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1}) - L(I)L(J_{s_1})\Big] E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon(p^2-p)[Nl]$$

$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)[Nl]$$

$$+N(K_B/10)\epsilon l \ .$$

Proof We know that $NP_e = \sum_n 1_I(X_n)1_J(X_{n+j})$, $Np_e = \sum_n 1_J(X_{n+j})$ and

$$N(P_e - L(I)p_e) = \sum_n \big[1_I(X_n)1_J(X_{n+j}) - L(I)1_J(X_{n+j})\big]$$

$$= \sum_n [1_I(X_n) - L(I)]1_J(X_{n+j}).$$

Then, $N(P_e - L(I)p_e - D) = \sum_n \big([1_I(X_n) - L(I)]1_J(X_{n+j}) - D_n\big)$, where $D_n = E\{[1_I(X_n) - L(I)]1_J(X_{n+j})\}$.

Then, the following equalities hold

$$N^2E\{(P_e - L(I)p_e - D)^2\}$$

$$= E\Big\{\sum_{n,m}\Big(\big\{[1_I(X_n)-L(I)]1_J(X_{n+j})-D_n\big\}\big\{[1_I(X_m)-L(I)]1_J(X_{m+j})-D_m\big\}\Big)\Big\}$$

$$= E\Big\{\sum_{n,m}\big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$= \sum_{n=m} E\Big\{\big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$+ \sum_{n,m\in G^*(n)} E\Big\{\big([1_I(X_n)-L(I)][1_I(X_m)-L(I)]1_J(X_{n+j})1_J(X_{m+j})-D_nD_m\big)\Big\}$$

$$+ \sum_{n,m\in H(n)\setminus G(n)} E\Big\{\big([1_I(X_n)-L(I)][1_I(X_m)-L(I)]1_J(X_{n+j})1_J(X_{m+j})-D_nD_m\big)\Big\}$$

$$+\sum_{n,m,m\in H*(n,q)} E\Big\{\big([1_I(X_n)-L(I)][1_I(X_m)-L(I)]1_J(X_{n+j})1_J(X_{m+j})-D_nD_m\big)\Big\}$$

$$+\sum_{n,m,m\notin H(n,q)} E\Big\{\big([1_I(X_n)-L(I)][1_I(X_m)-L(I)]1_J(X_{n+j})1_J(X_{m+j})-D_nD_m\big)\Big\}$$

$$= E\big\{\sum_n [1_I(X_n)-L(I)]^2 1_J(X_{n+j})\big\}$$

$$+\sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+\sum_{n,m\in G^*(n)} \big[(1/16)Ob(1)\epsilon_{p-1}+ob(1)(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon Ob(1)l_n\big]$$

$$+\sum_{n,m\in H(n)\setminus G(n)} Ob(1)\epsilon E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$+\sum_{n,m,m\in H^*(n,q)} Ob(1)\epsilon E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$-\sum_{n,m\in H(n,q)} D_nD_m$$

$$+\sum_n \sum_{n+d\notin H(n,q)} \delta'_d l$$

$$= \sum_n E\big\{1_I(X_n)-2L(I)1_I(X_n)+L(I)^2]1_J(X_{n+j})\big\}$$

$$+\sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)(p-1)\sum_n \big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l_n\big]$$

$$+Ob(1)\epsilon \sum_{n,m\in H(n)\setminus G(n)} E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$+Ob(1)\epsilon \sum_{n,m,m\in H^*(n,q)} E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$+\sum_{n,m\in H(n,q)} Ob(1)\epsilon^2 E\big\{1_J(X_{n+j})\big\}E\big\{1_J(X_{m+j})\big\}$$

$$+\sum_{n\in\mathbf{N}} (K_B/10)\epsilon l$$

$$= \sum_n E\big\{[(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2]1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon \sum_{n,m\in H(n)\setminus G(n)} E\big\{1_J(X_{n+j}))\big\}$$

$$+Ob(1)\epsilon \sum_{n,m\in H^*(n,q)} E\big\{(L(J)+\epsilon_{p-1})1_J(X_{n+j})\big\}$$

$$+Ob(1)\epsilon^2 \sum_{n,m\in H(n,q)} \big[L(J)+Ob(1)\epsilon_{p-1}\big]E\big\{1_J(X_{n+j})\big\}$$

$$+N(K_B/10)\epsilon l$$

$$= \sum_n [(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2]E\big\{1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon(p^2-p)\sum_n E\big\{1_J(X_{n+j})\big\}$$

$$+Ob(1)\epsilon\big[L(J)+\epsilon_{p-1}\big] \sum_{n,m\in H*(n,q)} E\big\{1_J(X_{n+j})\big\}$$

$$+Ob(1)\epsilon^2\big[L(J)+\epsilon_{p-1}\big] \sum_{n,m\in H(n,q)} E\big\{1_J(X_{n+j})\big\}$$

$$+N(K_B/10)\epsilon l$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2\big] \sum_n E\big\{1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon(p^2-p)[Nl]$$
$$+Ob(1)\epsilon\big[L(J)+\epsilon_{p-1}\big](2q)(p^2-p+1)\sum_n E\big\{1_J(X_{n+j})\big\}$$
$$+Ob(1)\epsilon^2\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)\sum_n E\big\{1_J(X_{n+j})\big\}$$
$$+N(K_B/10)\epsilon l$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2\big][Nl]$$
$$+\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p)[Nl]$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)[Nl]$$
$$+N(K_B/10)\epsilon l\ .\ \blacksquare$$

We deduce a ratio with the variance.

Lemma 9.3.11 *If ϵ is small enough and if p is not too large, the following equality holds :*

$$NE\Big\{\frac{(P_e-L(I)p_e-D)^2}{l^2}\Big\}=\sigma_{cp}^2\big[1+Ob(1)2\epsilon\xi_p\big].$$

Proof We have

$$NE\Big\{\frac{(P_e-L(I)p_e-D)^2}{l^2}\Big\}$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2\big]/l$$
$$+\frac{1}{Nl^2}\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)/l$$

$$+(K_B/10)\epsilon/l$$

$$= \big[(1-2L(I))L(I)+L(I)^2\big]/l + Ob(1)(1-2L(I))\epsilon/l$$
$$+\tfrac{1}{NL(J)^2}\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+\big[\tfrac{1}{Nl^2}-\tfrac{1}{NL(J)^2}\big]\sum_{n,m}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p+K_B/10)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)/l$$

$$= (1/L(J))\text{Ł}(I)[1-L(I)]+\big[1/l-1/L(J)\big]L(I)[1-L(I)]+(Ob(1)/l)(1-2L(I))\epsilon$$
$$+\tfrac{1}{NL(J)^2}\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)\frac{\big[1/L(J)-1/l\big]\big[1/L(J)+1/l\big]}{N}\sum_{n,m\in G^*(n)}(1/16)E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p+K_B/10)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)/l$$

$$= \sigma^2_{cp}+[Ob(1)/4]\big[1/l-1/L(J)\big]+(Ob(1)/l)(1-2L(I))\epsilon$$
$$+Ob(1)\frac{\big[1/L(J)-1/l\big]\big[3/L(J)\big]}{N}\sum_{n,m\in G^*(n)}(1/16)Ob(1)L(J)$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p+K_B/10)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big](2q+1)(p^2-p+1)/l$$

$$= \sigma_{cp}^2 + [Ob(1)/4]\big[1/l - 1/L(J)\big] + (Ob(1)/l)(1 - 2L(I))\epsilon$$
$$+Ob(1)\big[1/L(J) - 1/l\big]\big[3/L(J)\big](p-1)(1/16)Ob(1)L(J)$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2 - p + K_B/10)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J) + \epsilon_{p-1}\big](2q+1)(p^2-p+1)/l$$

$$= \sigma_{cp}^2 + Ob(1)\sigma_{cp}^2\Bigg[\big|1/l - 1/L(J)\big|\frac{1}{4\sigma_{cp}^2} + \frac{1}{l\sigma_{cp}^2}(1-2L(I))\epsilon$$
$$+\big|1/L(J) - 1/l\big|\frac{3}{\sigma_{cp}^2 L(J)}(p-1)(1/16)L(J)$$
$$+\frac{p-1}{\sigma_{cp}^2 l^2}\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+\epsilon\frac{p^2 - p + K_B/10}{l\sigma_{cp}^2}$$
$$+\epsilon\big[1+\epsilon\big]\frac{L(J) + \epsilon_{p-1}}{l\sigma_{cp}^2}(2q+1)(p^2-p+1)\Bigg]$$

$$= \sigma_{cp}^2 + Ob(1)\sigma_{cp}^2\Bigg[\big|1/l - 1/L(J)\big|\frac{1}{L(I)/L(J)} + \frac{1}{l(1/4)L(I)/L(J)}(1-2L(I))\epsilon$$
$$+\big|1/L(J) - 1/l\big|\frac{3}{(1/4)[L(I)/L(J)]L(J)}(p-1)(1/16)L(J)$$
$$+\frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]} + 13\Big]$$
$$+\epsilon\frac{p^2 - p + K_B/10}{l\ (1/4)L(I)/L(J)}$$
$$+\epsilon\big[1+\epsilon\big]\frac{L(J) + \epsilon_{p-1}}{l(1/4)L(I)/L(J)}(2q+1)(p^2-p+1)\Bigg]$$

(by assumptions 9.3.3 and by lemma 9.3.4)

$$= \sigma_{cp}^2 + Ob(1)\sigma_{cp}^2 \Bigg[\mathcal{M}(l)\epsilon \frac{L(J)}{L(I)} + \frac{8L(J)}{l}(1 - 2L(I))\epsilon$$

$$+ \mathcal{M}(l)\epsilon \frac{12}{L(I)}(p-1)(1/16)L(J)$$

$$+ \frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]} + 13\Big]$$

$$+ \epsilon \frac{8[p(p-1) + K_B/10]L(J)}{l}$$

$$+ \epsilon\big[1+\epsilon\big]\frac{8L(J)\big[L(J)+\epsilon_{p-1}\big]}{l}(2q+1)(p^2-p+1)\Bigg]$$

(where, by notation 9.3.2, $\epsilon\mathcal{M}(l) = |1/L(J) - 1/l|$)

$$= \sigma_{cp}^2 + Ob(1)\epsilon\ \sigma_{cp}^2 \Bigg[2L(J)\mathcal{M}(l) + \frac{8L(J)}{l}(1-2L(I)) + (3/2)(p-1)L(J)\mathcal{M}(l)$$

$$+ \frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]} + 13\Big]$$

$$+ \frac{8[p(p-1) + K_B/10]L(J)}{l}$$

$$+ \ \big[1+\epsilon\big]\frac{8L(J)[L(J)+\epsilon_{p-1}]}{l}(2q+1)(p^2-p+1)\Bigg]$$

$$= \sigma_{cp}^2\big[1 + Ob(1)2\epsilon\xi_p\big]\ .\ \blacksquare$$

9.3.6 Study of $E\Big\{\Big[\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2}\Big]^2\Big\}$

Lemma 9.3.12 *We denote by $\sigma_1(J)$ and $\sigma_B(J)$ the variances σ_1 et σ_B defined in notations 9.2.4, but associated to J instead of Bo. Then, the following inequality holds*

$$\frac{\sigma_1(J)^2/l^2}{\sigma_{cp}^2} \leq \theta_p\ .$$

Proof For p'=p-1, $p'^2-p'+1 = (p-1)^2-(p-1)+1 = p^2-2p+1-p+1+1 = p^2-3p+3$.

Therefore, by using hypothesis 9.3.3 and lemma 9.2.8,

$$\frac{\sigma_1(J)^2/l^2}{\sigma_{cp}^2} \leq \frac{\sigma_1(J)^2/l^2}{0.25L(I)/L(J)} = \frac{8\sigma_1(J)^2L(J)}{l^2} \leq \theta_p \ . \ \blacksquare$$

We deduce the following lemma.

Lemma 9.3.13 *The following equality holds :*

$$N.\mathbb{E}\Big\{\frac{[\epsilon(l-p_e)]^2}{l^2\sigma_{cp}^2}\Big\} = Ob(1)\epsilon^2\theta_p.$$

Then, one obtains the following lemma.

Lemma 9.3.14 *The following equality holds*

$$N.\mathbb{E}\Big\{\Big(\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2}\Big)^2\Big\}$$
$$= \sigma_{cp}^2\Big[1 + Ob(1)\epsilon\Big(2\xi_p + 2[1+\epsilon\xi_p]\sqrt{\theta_p} + \epsilon\theta_p\Big)\Big].$$

Proof On the one hand, by the proof of lemma 9.2.9 and by lemma 9.3.11,

$$\sqrt{N.\mathbb{E}\Big\{\frac{(P_e - L(I)p_e - D)^2}{l^2}\Big\}} \leq \sqrt{\sigma_{cp}^2[1+2\epsilon\xi_p]} \leq \sigma_{cp}[1+\epsilon\xi_p].$$

On the other hand, by using lemma 9.3.7, we have $D = Ob(1)\epsilon l$.

Therefore,

$$\frac{Dl - Dp_e}{l^2} = \frac{Ob(1)\epsilon(l-p_e)}{l} \ .$$

Therefore,

$$\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} = \frac{P_e - L(I)p_e - D}{l} + \frac{Ob(1)\epsilon(l-p_e)}{l} \ .$$

By Schwarz Inequality,

$$N.\mathbb{E}\Big\{\Big|\frac{P_e - L(I)p_e - D}{l}\ Ob(1)\frac{\epsilon(l-p_e)}{l}\Big|\Big\}$$

$$\leq \sqrt{N\mathbb{E}\Big\{\frac{(L_e - L(I)l_e - D)^2}{l^2}\Big\}}\ \sqrt{N\mathbb{E}\Big\{\frac{\epsilon^2(l-p_e)^2}{l^2}\Big\}}$$

$$\leq \epsilon\sqrt{\sigma_{cp}^2[1+2\epsilon\xi_p]}\sqrt{\sigma_{cp}^2\theta_p}.$$

$$\leq \epsilon\sigma_{cp}^2[1+\epsilon\xi_p]\sqrt{\theta_p}\ .$$

Therefore,

$$N.\mathbb{E}\Big\{\Big(\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2}\Big)^2\Big\}$$
$$= \sigma_{cp}^2\Big[1 + Ob(1)\epsilon\Big(2\xi_p + 2[1+\epsilon\xi_p]\sqrt{\theta_p} + \epsilon\theta_p\Big)\Big]\ . \blacksquare$$

We deduce the following lemma

Lemma 9.3.15 *The following equality holds*

$$\sigma_2 = \sigma_{cp}\Big[1 + Ob(1)\gamma_{2,p}\Big]\ .$$

Proof By the previous lemma

$$\sigma_2^2 = \sigma_{cp}^2\Big[1 + Ob(1)2\gamma_{2,p}\Big]\ .$$

By the proof of lemma 9.2.9,

$$\sigma_2 = \sigma_{cp}\Big[1 + Ob(1)\gamma_{2,p}\Big]\ . \blacksquare$$

9.3.7 Proof of theorem 10

By our assumptions , $N^{1/2}(p_e - l) = O_P(1)$ and $N^{1/2}(P_e - L(I)p_e - D) = O_P(1)$ and $p_e - l$ converges in probability to 0.

We keep notation 9.1.1 : $o_P(1) \xrightarrow{P} 0$. Then,

$$\sqrt{N}\Big(\frac{P_e}{p_e} - L(I)\Big) = \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{p_e} + \frac{D}{p_e}\Big)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \sqrt{N}\big(P_e - L(I)p_e - D\big)\Big[\frac{1}{p_e} - \frac{1}{l}\Big] + \frac{\sqrt{N}D}{p_e}$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{p_e} + o_P(1)$$

(because $\sqrt{N}\frac{P_e - L(I)p_e - D}{l} = O_P(1)$ and $p_e - l \xrightarrow{P} 0$)

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{l} + \frac{\sqrt{N}D}{p_e} - \frac{\sqrt{N}D}{l} + o_P(1)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{l} + \sqrt{N}\frac{Dl - Dp_e}{lp_e} + o_P(1)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{l} + \sqrt{N}\frac{Dl - Dp_e}{l^2} + \sqrt{N}\frac{Dl - Dp_e}{l}\Big[\frac{1}{p_e} - \frac{1}{l}\Big] + o_P(1)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \sqrt{N}\frac{Dl - Dp_e}{l^2} + \frac{\sqrt{N}D}{l} + o_P(1)$$

(because $p_e - l \xrightarrow{P} 0$).

We set

$$H_e = \frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} \ .$$

Then,

$$P\left\{\sqrt{N}\left|\frac{P_e}{p_e} - L(I)\right| > \sigma_{cp}x\right\}$$

$$= P\left\{\sqrt{N}\left|\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} + \frac{D}{l} + \frac{o_P(1)}{\sqrt{N}}\right| > \sigma_{cp}x\right\}$$

$$\leq P\left\{\sqrt{N}\left|\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2}\right| + \big|o_P(1)\big| > \sigma_{cp}x - \frac{N^{1/2}D}{l}\right\}$$

$$\leq P\left\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x - \frac{N^{1/2}D}{l} - \big|o_P(1)\big|\right\}$$

$$\leq P\left\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1 - \frac{\beta_{2,p}}{x}\big) + \eta\sigma_{cp} - \big|o_P(1)\big|\right\}$$

$$\leq P\left\{\left\{\sqrt{N}\left|H_e\right| > \sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)+\eta\sigma_{cp}-\left|o_P(1)\right|\right\}\cap\left\{\left|o_P(1)\right|\leq\frac{\eta\sigma_{cp}}{2}\right\}\right\}$$

$$+P\left\{\left\{\sqrt{N}\left|H_e\right| > \sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)+\eta\sigma_{cp}-\left|o_P(n)\right|\right\}\cap\left\{\left|o_P(1)\right|>\frac{\eta\sigma_{cp}}{2}\right\}\right\}$$

$$\leq P\left\{\sqrt{N}\left|H_e\right| > \sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)+\eta\sigma_{cp}/2\right\}$$
$$+P\left\{\left|o_P(1)\right|>\frac{\eta\sigma_{cp}}{2}\right\}$$

$$\leq P\left\{\sqrt{N}\left|H_e\right| > \sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)\right\}$$
$$-P\left\{\sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)+\eta\sigma_{cp}/2\geq\sqrt{N}\left|H_e\right| > \sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)\right\}$$
$$+P\left\{\left|o_P(1)\right|>\frac{\eta\sigma_{cp}}{2}\right\}$$

$$\leq P\left\{\sqrt{N}\left|H_e\right| > \sigma_{cp}x\left(1-\frac{\beta_{2,p}}{x}\right)\right\}$$

(for N large enough, considering that $\sqrt{N}H_e$ converges in distribution to a random variable which has a distribution function $F_{H_e}(x) > 0$ and considering the convergence in probability to 0 of $o_P(1)$)

$$= P\left\{\sqrt{N}\left|H_e\right| > \sigma_2 x\frac{1-\beta_{2,p}/x}{1+Ob(1)\gamma_{2,p}}\right\}$$

$$= P\left\{\sqrt{N}\left|\frac{P_e-L(I)p_e-D}{l}+\frac{Dl-Dp_e}{l^2}\right| > \sigma_2 x\frac{1-\beta_{2,p}/x}{1+Ob(1)\gamma_{2,p}}\right\}$$

$$\leq P\left\{\sqrt{N}\left|\frac{P_e-L(I)p_e-D}{l}+\frac{Dl-Dp_e}{l^2}\right| > \sigma_2 x\frac{1-\beta_{2,p}/x}{1+\gamma_{2,p}}\right\}. \quad \blacksquare$$

9.4 Third and Fourth Theorems

Now assume that the hypotheses 9.2.1 and 9.3.1 do not hold or that they are too difficult to compute. Then, one can use more general theorems. It is the aim of this section.

9.4.1 Wordings

Now, we generalize theorem 9.

Theorem 11 *We suppose that the notations of section 9.1 and 9.2 hold except the hypothesis 9.2.1 . We set*

$$\gamma"_{1,p} = \frac{1}{2A(p)L(Bo)}\Big[(p^2-p+1)\epsilon_p + N\Big(2^{1-p}\epsilon_p + \epsilon_p^2\Big) + N\epsilon_{2p}\Big] \ .$$

Then, the following inequality holds

$$P\Big\{\sqrt{N}\big|P_e - L(Bo)\big| \geq \sigma_B(Bo)x\Big\} \leq K_1\Big(\frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}}x\Big) \ ,$$

where $K_1(x) = P\Big\{\frac{\sqrt{N}|P_e - L^N(Bo)|}{\sigma_1(Bo)} \geq x\Big\}$.

Now, we take an interest to theorem 10. First, we generalize notations 9.3.2.

Notations 9.4.1 *We keep the notations of section 9.3. We define* $\xi"_p$, $\theta"_p$ *and* $\gamma"_{2,p}$ *by the following way :*

$$\xi"_p = \frac{1}{2}\Bigg[2L(J)\mathcal{M}(l) + \frac{8L(J)}{l}(1-2L(I)) + (3/2)(p-1)L(J)\mathcal{M}(l) + \frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]} + 13\Big] + \frac{8[p(p-1)]L(J)}{l} + \big[1+\epsilon\big]\frac{8L(J)N[L(J)+\epsilon_{p-1}]}{l}\Bigg]$$

$$\theta"_p = \frac{8\sigma_B(J)^2L(J)}{l^2} + (p^2-3p+3)\frac{8\epsilon_{p-1}L(J)}{l^2} + \frac{8N\big[2^{2-p}\epsilon_{p-1}+\epsilon_{p-1}^2\big]L(J)}{l^2} + \frac{8N\epsilon_{2p-2}L(J)}{l^2} \ .$$

$$2\gamma"_{2,p} = \epsilon\Big(2\xi"_p + 2[1+\epsilon\xi"_p]\sqrt{\theta"_p} + \epsilon\theta"_p\Big) \ .$$

Then, one can generalize theorem 10 by the following way.

Theorem 12 *We keep the notations and assumptions of section 9.3 except the hypothseses 9.3.1 and 9.3.2. Let $p \in \mathbb{N}^*$. Let $N \in \mathbb{N}^*$. One supposes that*

$$\sqrt{N}\frac{P_e - L(Bo)p_e - D}{l} = O_P(1) \ ,$$
$$\sqrt{N}(p_e - l^N(J)) = O_P(1) \ ,$$
$$p_e \xrightarrow{P} l \quad as \ N \to \infty \ ,$$
$$l \to L(J) \quad as \ N \to \infty \ ,$$
$$L \to L(Bo)L(J) \quad as \ N \to \infty \ ,$$
$$P_e - L^N(Bo \otimes J) \xrightarrow{P} 0 \quad as \ N \to \infty \ .$$

One supposes that $\sqrt{N}H_e = \sqrt{N}\Big(\frac{P_e - L(Bo)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big)$ converges in distribution to a random variable which has the distribution function $F_{H_e}(x) > 0$.

Then, for all $\eta > 0$, there exists $N_0 \in \mathbb{N}$ such that, for all $N \geq N_0$,

$$P\left\{\sqrt{N}\left|\frac{P_e}{p_e} - L(Bo)\right| > \sigma_{cp} \ x\right\} \leq K_2\Big(\frac{1-\beta_{2,p}/x}{1+\gamma"_{2,p}}x\Big) \ ,$$

where $K_2(x) = P\Big\{\frac{\sqrt{N}}{\sigma_2}\Big|\frac{P_e - L(Bo)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big| > x\Big\}$.

9.4.2 Proof of theorem 11

First, we generalize lemma 9.2.8.

Lemma 9.4.1 *The following equality holds*

$$\sigma_1^2 = \sigma_B^2\big[1 + Ob(1)2\gamma"_{1,p}\big].$$

Proof Let X'_n be an IID sequence with uniform distribution. We recall the equalities (cf lemmas 9.2.6 and 9.2.7) :

$$if\ m \notin H(n),\ \ \mathbb{E}\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} = L(Bo)^2 + Ob(1)\epsilon_{2p},$$

$$if\ m \in H(n),\ \ \mathbb{E}\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} = \mathbb{E}\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\} + Ob(1)\epsilon_p\ ,$$

$$L_nL_m = L(Bo)^2 + 2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2\ .$$

Then,

$$\sigma_1^2 = (1/N)\mathbb{E}\Big\{\big[\sum_{n=1}^{N}\big(1_{Bo}(X_n) - L_n\big)\big]^2\Big\}$$

$$= (1/N)\mathbb{E}\Big\{\sum_{n=1}^{N}\sum_{m=1}^{N}\big(1_{Bo}(X_n) - L_n\big)\big(1_{Bo}(X_m) - L_m\big)\Big\}$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m=1}^{N}\Big(\mathbb{E}\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} - L_nL_m\Big)$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(\mathbb{E}\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} - L_nL_m\Big)$$

$$+(1/N)\mathbb{E}\Big\{\sum_{n=1}^{N}\sum_{m\notin H(n)}\Big(E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} - L_nL_m\Big)$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(\mathbb{E}\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\} + Ob(1)\epsilon_p\Big)$$

$$-(1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(L(Bo)^2 + 2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2\Big)$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\notin H(n)}\Big(L(Bo)^2 + Ob(1)\epsilon_{2p} - L(Bo)^2 - 2^{1-p}Ob(1)\epsilon_p - Ob(1)\epsilon_p^2\Big)$$

$$= (1/N)\sum_{n=1}^{N}\sum_{m\in H(n)}\Big(\mathbb{E}\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\} - L(Bo)^2\Big)$$

$$-(1/N)\sum_{n=1}^{N}\sum_{m\in H(n)} Ob(1)\epsilon_p$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m\notin H(n)} Ob(1)\epsilon_{2p}$$

$$+(1/N)\sum_{n=1}^{N}\sum_{m=1}^{N}\Big(2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2\Big)$$

$$= \sigma_B^2 + (p^2-p+1)Ob(1)\epsilon_p$$
$$+\big[N-(p^2-p+1)\big]Ob(1)\epsilon_{2p}$$
$$+N\Big(2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2\Big)$$

$$= \sigma_B^2 + \sigma_B^2\Big[\frac{(p^2-p+1)Ob(1)\epsilon_p}{\sigma_B^2}$$
$$+\frac{NOb(1)\epsilon_{2p}}{\sigma_B^2} + \frac{N\Big(2^{1-p}Ob(1)\epsilon_p + Ob(1)\epsilon_p^2\Big)}{\sigma_B^2}\Big]$$

$$= \sigma_B^2 + 2\frac{\sigma_B^2 Ob(1)}{2A(p)L(Bo)}\Big[(p^2-p+1)\epsilon_p + N\epsilon_{2p} + N\Big(2^{1-p}\epsilon_p + \epsilon_p^2\Big)\Big]\ . \blacksquare$$

Proof 9.4.2 *We prove now the theorem 11*

The following inequalities hold.

$$P\left\{\left|\sqrt{N}\big[P_e - L(Bo)\right| > \sigma_B x\right\}$$

$$\leq P\left\{\left|\sqrt{N}\big[P_e - L^N(Bo)\big]\right| > \sigma_B x - \sqrt{N}|L^N - L(Bo)|\right\}$$

$$\leq P\left\{\left|\sqrt{N}\big[P_e - L^N(Bo)\big]\right| > \sigma_B x[1-\beta_{1,p}/x]\right\}$$
$$\leq P\left\{\left|\sqrt{N}\big[P_e - L^N(Bo)\big]\right| > \frac{1-\beta_{1,p}/x}{1+\gamma"_{1,p}}\sigma_1 x\right\}$$
$$= K\Big(\frac{1-\beta_{1,p}/x}{1+\gamma"_{1,p}}x\Big)\ .\ \blacksquare$$

9.4.3 Proof of theorem 12

At first, we generalize lemma 9.3.10.

Lemma 9.4.3 *The following equality holds :*

$$N^2 E\{(P_e - L(I)p_e - D)^2\}$$
$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon) + L(I)^2\big][Nl]$$
$$+\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p)[Nl]$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big][N^2 l]\ .$$

Proof We recall the properties of section 9.3.5 :

$$\mathbb{E}\big\{1_J(X_n)f(X)\big\} = L(J)\mathbb{E}\big\{f(X)\big\} + Ob(1)\epsilon\mathbb{E}\big\{|f(X)|\big\}\ ,$$
$$\mathbb{E}\big\{\big(1_J(X_n)-L(J)\big)f(X)\big\} = Ob(1)\epsilon E\big\{|f(X)|\big\}\ ,$$
$$D_n = Ob(1)\epsilon l_n\ ,$$
$$D = Ob(1)\epsilon l\ ,$$
$$E\{1_I(X_n)1_J(X_{n+j}\} = \big[L(I)+Ob(1)\epsilon\big]\mathbb{E}\{1_J(X_{n+j}\}\ ,$$
$$If\ m\notin H(n),\ \mathbb{E}\big\{1_J(X_{n+j})1_J(X_{m+j})\big\} = \big[L(J)+\epsilon_{p-1}\big]E\big\{1_J(X_{n+j})\big\}\ ,$$
$$NP_e = \sum_n 1_I(X_n)1_J(X_{n+j})\ ,\quad Np_e = \sum_n 1_J(X_{n+j})\ ,$$
$$N(P_e - L(I)p_e) = \sum_n [1_I(X_n)-L(I)]1_J(X_{n+j})\ ,$$

$$N(P_e - L(I)p_e - D) = \sum_n \big([1_I(X_n) - L(I)]1_J(X_{n+j}) - D_n\big) \ .$$

Moreover, we use notations 9.3.3 and lemmas 9.3.2 and 9.3.3. Then, the following equalities hold

$$N^2 E\{(P_e - L(I)p_e - D)^2\}$$

$$= E\Big\{ \sum_{n,m} \Big(\big\{[1_I(X_n) - L(I)]1_J(X_{n+j}) - D_n\big\}\big\{[1_I(X_m) - L(I)]1_J(X_{m+j}) - D_m\big\}\Big)\Big\}$$

$$= E\Big\{ \sum_{n,m} \big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$= \sum_{n=m} E\Big\{ \big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$+ \sum_{n,m\in G^*(n)} E\Big\{ \big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$+ \sum_{n,m\in H(n)\setminus G(n)} E\Big\{ \big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$+ \sum_{n,m,m\notin H(n)} E\Big\{ \big([1_I(X_n) - L(I)][1_I(X_m) - L(I)]1_J(X_{n+j})1_J(X_{m+j}) - D_nD_m\big)\Big\}$$

$$= E\big\{ \sum_n [1_I(X_n) - L(I)]^2 1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0}) - L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1}) - L(I)L(J_{s_1})\Big] E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+ \sum_{n,m\in G^*(n)} \big[(1/16)Ob(1)\epsilon_{p-1} + Ob(1)(3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon Ob(1) l_n\big]$$

$$+ \sum_{n,m\in H(n)\setminus G(n)} Ob(1)\epsilon E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$+ \sum_{n,m,m\notin H(n)} Ob(1)\epsilon E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$-\sum_{n,m} D_n D_m$$

$$= \sum_n E\big\{1_I(X_n) - 2L(I)1_I(X_n) + L(I)^2]1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0}) - L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1}) - L(I)L(J_{s_1})\Big] E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+ Ob(1)(p-1)\sum_n \big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l_n\big]$$

$$+ Ob(1)\epsilon \sum_{n,m\in H(n)\setminus G(n)} E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$+ Ob(1)\epsilon \sum_{n,m\notin H(n)} E\big\{1_J(X_{n+j})1_J(X_{m+j})\big\}$$

$$+ \sum_{n,m} Ob(1)\epsilon^2 E\big\{1_J(X_{n+j})\big\} E\big\{1_J(X_{m+j})\big\}$$

$$= \sum_n E\big\{[(1-2L(I))(L(I) + Ob(1)\epsilon) + L(I)^2]1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0}) - L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1}) - L(I)L(J_{s_1})\Big] E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+ Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$

$$+ Ob(1)\epsilon \sum_{n,m\in H(n)\setminus G(n)} E\big\{1_J(X_{n+j}))\big\}$$

$$+ Ob(1)\epsilon \sum_{n,m\notin H(n)} E\big\{(L(J) + \epsilon_{p-1})1_J(X_{n+j})\big\}$$

$$+ Ob(1)\epsilon^2 \sum_{n,m} \big[L(J) + Ob(1)\epsilon_{p-1}\big] E\big\{1_J(X_{n+j})\big\}$$

$$= \sum_n [(1-2L(I))(L(I) + Ob(1)\epsilon) + L(I)^2] E\big\{1_J(X_{n+j})\big\}$$

$$+ \sum_{n,m\in G^*(n)} \Big[L(I\cap J_{s_0}) - L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1}) - L(I)L(J_{s_1})\Big] E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+ Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon(p^2-p)\sum_n E\big\{1_J(X_{n+j})\big\}$$

$$+Ob(1)\epsilon\big[L(J)+\epsilon_{p-1}\big]\sum_{n,m\notin H(n)} E\big\{1_J(X_{n+j})\big\}$$

$$+Ob(1)\epsilon^2\big[L(J)+\epsilon_{p-1}\big]\sum_{n,m} E\big\{1_J(X_{n+j})\big\}$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2\big]\sum_n E\big\{1_J(X_{n+j})\big\}$$

$$+\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon(p^2-p)[Nl]$$

$$+Ob(1)\epsilon\big[L(J)+\epsilon_{p-1}\big]N\sum_n E\big\{1_J(X_{n+j})\big\}$$

$$+Ob(1)\epsilon^2\big[L(J)+\epsilon_{p-1}\big]N\sum_n E\big\{1_J(X_{n+j})\big\}$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon)+L(I)^2\big][Nl]$$

$$+\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$

$$+Ob(1)N(p-1)\big[(1/16)\epsilon_{p-1}+(3/8)(1+2\epsilon)\epsilon L(J)+(3/2)\epsilon l\big]$$

$$+Ob(1)\epsilon(p^2-p)[Nl]$$

$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big][N^2l]\ .\ \blacksquare$$

We deduce a ratio with the variance.

Lemma 9.4.4 *If ϵ is small enough, the following equality holds :*

$$NE\Big\{\frac{(P_e-L(I)p_e-D)^2}{l^2}\Big\} = \sigma_{cp}^2\big[1+Ob(1)2\epsilon\xi"_p\big].$$

Proof We have

$$NE\Big\{\frac{(P_e - L(I)p_e - D)^2}{l^2}\Big\}$$

$$= \big[(1-2L(I))(L(I)+Ob(1)\epsilon) + L(I)^2\big]/l$$
$$+\frac{1}{Nl^2}\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big]N/l$$

$$= \big[(1-2L(I))L(I) + L(I)^2\big]/l + Ob(1)(1-2L(I))\epsilon/l$$
$$+\frac{1}{NL(J)^2}\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+\big[\frac{1}{Nl^2}-\frac{1}{NL(J)^2}\big]\sum_{n,m}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big]N/l$$

$$= (1/L(J))\text{Ł}(I)[1-L(I)] + \big[1/l - 1/L(J)\big]L(I)[1-L(I)] + (Ob(1)/l)(1-2L(I))\epsilon$$
$$+\frac{1}{NL(J)^2}\sum_{n,m\in G^*(n)}\Big[L(I\cap J_{s_0})-L(I)L(J_{s_0})\Big]\Big[L(I\cap J_{s_1})-L(I)L(J_{s_1})\Big]E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)\frac{\big[1/L(J)-1/l\big]\big[1/L(J)+1/l\big]}{N}\sum_{n,m\in G^*(n)}(1/16)E\Big\{1_{J^1}(X'_{n+j^1})1_{J^0}(X'_{m+j^0})\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2-p)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J)+\epsilon_{p-1}\big]N/l$$

$$= \sigma_{cp}^2 + [Ob(1)/4]\big[1/l - 1/L(J)\big] + (Ob(1)/l)(1 - 2L(I))\epsilon$$
$$+Ob(1)\frac{\big[1/L(J) - 1/l\big]\big[3/L(J)\big]}{N} \sum_{n,m \in G^*(n)} (1/16)Ob(1)L(J)\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2 - p)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J) + \epsilon_{p-1}\big]N/l$$

$$= \sigma_{cp}^2 + [Ob(1)/4]\big[1/l - 1/L(J)\big] + (Ob(1)/l)(1 - 2L(I))\epsilon$$
$$+Ob(1)\big[1/L(J) - 1/l\big]\big[3/L(J)\big](p-1)(1/16)Ob(1)L(J)\Big\}$$
$$+Ob(1)(p-1)(1/l^2)\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+Ob(1)\epsilon(p^2 - p)/l$$
$$+Ob(1)\epsilon\big[1+\epsilon\big]\big[L(J) + \epsilon_{p-1}\big]N/l$$

$$= \sigma_{cp}^2 + Ob(1)\sigma_{cp}^2\Bigg[\big|1/l - 1/L(J)\big|\frac{1}{4\sigma_{cp}^2} + \frac{1}{l\sigma_{cp}^2}(1 - 2L(I))\epsilon$$
$$+\big|1/L(J) - 1/l\big|\frac{3}{\sigma_{cp}^2 L(J)}(p-1)(1/16)L(J)\Big\}$$
$$+\frac{p-1}{\sigma_{cp}^2 l^2}\big[(1/16)\epsilon_{p-1} + (3/8)(1+2\epsilon)\epsilon L(J) + (3/2)\epsilon l\big]$$
$$+\epsilon\frac{p^2 - p}{l\sigma_{cp}^2}$$
$$+\epsilon\big[1+\epsilon\big]\frac{N(L(J) + \epsilon_{p-1})}{l\sigma_{cp}^2}\Bigg]$$

$$= \sigma_{cp}^2 + Ob(1)\sigma_{cp}^2\Bigg[\big|1/l - 1/L(J)\big|\frac{1}{L(I)/L(J)} + \frac{1}{l(1/4)L(I)/L(J)}(1 - 2L(I))\epsilon$$
$$+\big|1/L(J) - 1/l\big|\frac{3}{(1/4)[L(I)/L(J)]L(J)}(p-1)(1/16)L(J)$$

$$+\frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$
$$+\epsilon\frac{p^2-p}{l\ (1/4)L(I)/L(J)}$$
$$+\epsilon\big[1+\epsilon\big]\frac{N(L(J)+\epsilon_{p-1})}{l(1/4)L(I)/L(J)}\Bigg]$$

(by assumptions 9.3.3 and by lemma 9.3.4)

$$=\sigma_{cp}^2+Ob(1)\sigma_{cp}^2\Bigg[\mathcal{M}(l)\epsilon\frac{L(J)}{L(I)}+\frac{8L(J)}{l}(1-2L(I))\epsilon$$
$$+\mathcal{M}(l)\epsilon\frac{12}{L(I)}(p-1)(1/16)L(J)$$
$$+\frac{(p-1)\epsilon}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$
$$+\epsilon\frac{8[p(p-1)]L(J)}{l}$$
$$+\epsilon\big[1+\epsilon\big]\frac{8L(J)N\big[L(J)+\epsilon_{p-1}\big]}{l}\Bigg]$$

(where, by notation 9.3.2, $\epsilon\mathcal{M}(l)=|1/L(J)-1/l|$)

$$=\sigma_{cp}^2+Ob(1)\epsilon\ \sigma_{cp}^2\Bigg[2L(J)\mathcal{M}(l)+\frac{8L(J)}{l}(1-2L(I))+(3/2)(p-1)L(J)\mathcal{M}(l)$$
$$+\frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$
$$+\frac{8[p(p-1)]L(J)}{l}$$
$$+\ \big[1+\epsilon\big]\frac{8L(J)N[L(J)+\epsilon_{p-1}]}{l}\Bigg]$$

$$=\sigma_{cp}^2\big[1+Ob(1)2\epsilon\xi"_p\big]\ .\ \blacksquare$$

Lemma 9.4.5 *The following inequality holds*

$$\frac{\sigma_1(J)^2/l^2}{\sigma_{cp}^2} \leq \theta_p \ .$$

Proof For p'=p-1, $p'^2 - p' + 1 = p^2 - 3p + 3$. Therefore, by using hypothesis 9.3.3 and lemma 9.2.8,

$$\frac{\sigma_1(J)^2/l^2}{\sigma_{cp}^2} \leq \frac{\sigma_1(J)^2/l^2}{0.25L(I)/L(J)} = \frac{8\sigma_1(J)^2 L(J)}{l^2} \leq \theta"_p \ . \ \blacksquare$$

We deduce the following lemmas.

Lemma 9.4.6 *The following equality holds :*

$$N.\mathbb{E}\Big\{ \frac{[\epsilon(l - p_e)]^2}{l^2 \sigma_{cp}^2} \Big\} = Ob(1)\epsilon^2 \theta"_p.$$

Lemma 9.4.7 *The following equality holds*

$$N.\mathbb{E}\Big\{ \Big(\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} \Big)^2 \Big\}$$

$$= \sigma_{cp}^2 \Big[1 + Ob(1)\epsilon \Big(2\xi"_p + 2[1 + \epsilon\xi"_p]\sqrt{\theta"_p} + \epsilon\theta"_p \Big) \Big].$$

Proof On the one hand, by the proof of lemma 9.2.9 and by lemma 9.4.4,

$$\sqrt{N.\mathbb{E}\Big\{ \frac{(P_e - L(I)p_e - D)^2}{l^2} \Big\}} \leq \sqrt{\sigma_{cp}^2[1 + 2\epsilon\xi"_p]} \leq \sigma_{cp}[1 + \epsilon\xi"_p].$$

On the other hand, by using lemma 9.3.7, we have $D = Ob(1)\epsilon l$.

Therefore,

$$\frac{Dl - Dp_e}{l^2} = \frac{Ob(1)\epsilon(l - p_e)}{l} \ .$$

Therefore,

$$\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} = \frac{P_e - L(I)p_e - D}{l} + \frac{Ob(1)\epsilon(l - p_e)}{l} \ .$$

By Schwarz Inequality,

$$N.\mathbb{E}\Big\{\Big|\frac{P_e - L(I)p_e - D}{l}\ Ob(1)\frac{\epsilon(l-p_e)}{l}\Big|\Big\}$$

$$\leq \sqrt{N\mathbb{E}\Big\{\frac{(L_e - L(I)l_e - D)^2}{l^2}\Big\}}\ \sqrt{N\mathbb{E}\Big\{\frac{\epsilon^2(l-p_e)^2}{l^2}\Big\}}$$

$$\leq \epsilon\sqrt{\sigma_{cp}^2[1+2\epsilon\xi"_p]}\sqrt{\sigma_{cp}^2\theta"_p}.$$

$$\leq \epsilon\sigma_{cp}^2[1+\epsilon\xi"_p]\sqrt{\theta"_p}\ .$$

Therefore,

$$N.\mathbb{E}\Big\{\Big(\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2}\Big)^2\Big\}$$

$$= \sigma_{cp}^2\Big[1 + Ob(1)\epsilon\Big(2\xi"_p + 2[1+\epsilon\xi"_p]\sqrt{\theta"_p} + \epsilon\theta"_p\Big)\Big] = \sigma_{cp}^2\Big[1 + Ob(1)2\gamma"_{2,p}\Big]\ . \blacksquare$$

By the proof of lemma 9.2.9, we deduce the following lemma

Lemma 9.4.8 *The following equality holds*

$$\sigma_2 = \sigma_{cp}\Big[1 + Ob(1)\gamma"_{2,p}\Big]\ .$$

Proof 9.4.9 *Now, we prove theorem 12*

By our assumptions , $N^{1/2}(p_e - l) = O_P(1)$ and $N^{1/2}(P_e - L(I)p_e - D) = O_P(1)$ and $p_e - l$ converges in probability to 0.

We keep notation 9.1.1 : $o_P(1) \xrightarrow{P} 0$. Then,

$$\sqrt{N}\Big(\frac{P_e}{p_e} - L(I)\Big) = \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{p_e} + \frac{D}{p_e}\Big)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \sqrt{N}\big(P_e - L(I)p_e - D\big)\Big[\frac{1}{p_e} - \frac{1}{l}\Big] + \frac{\sqrt{N}D}{p_e}$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{p_e} + o_P(1)$$

(because $\sqrt{N}\frac{P_e-L(I)p_e-D}{l} = O_P(1)$ and $p_e - l \xrightarrow{P} 0$)

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{l} + \frac{\sqrt{N}D}{p_e} - \frac{\sqrt{N}D}{l} + o_P(1)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \frac{\sqrt{N}D}{l} + \sqrt{N}\frac{Dl - Dp_e}{lp_e} + o_P(1)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big)+\frac{\sqrt{N}D}{l}+\sqrt{N}\frac{Dl - Dp_e}{l^2}+\sqrt{N}\frac{Dl - Dp_e}{l}\Big[\frac{1}{p_e}-\frac{1}{l}\Big]+o_P(1)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e - D}{l}\Big) + \sqrt{N}\frac{Dl - Dp_e}{l^2} + \frac{\sqrt{N}D}{l} + o_P(1)$$

(because $p_e - l \xrightarrow{P} 0$).

We set

$$H_e = \frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} \ .$$

Then,

$$P\left\{\sqrt{N}\left|\frac{P_e}{p_e} - L(I)\right| > \sigma_{cp}x\right\}$$

$$= P\left\{\sqrt{N}\left|\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2} + \frac{D}{l} + \frac{o_P(1)}{\sqrt{N}}\right| > \sigma_{cp}x\right\}$$

$$\le P\left\{\sqrt{N}\left|\frac{P_e - L(I)p_e - D}{l} + \frac{Dl - Dp_e}{l^2}\right| + \big|o_P(1)\big| > \sigma_{cp}x - \frac{N^{1/2}D}{l}\right\}$$

$$\le P\left\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x - \frac{N^{1/2}D}{l} - \big|o_P(1)\big|\right\}$$

$$\le P\left\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1 - \frac{\beta_{2,p}}{x}\big) + \eta\sigma_{cp} - \big|o_P(1)\big|\right\}$$

$$\le P\left\{\Big\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1 - \frac{\beta_{2,p}}{x}\big) + \eta\sigma_{cp} - \big|o_P(1)\big|\Big\} \cap \Big\{\big|o_P(1)\big| \le \frac{\eta\sigma_{cp}}{2}\Big\}\right\}$$

$$+P\left\{\Big\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1 - \frac{\beta_{2,p}}{x}\big) + \eta\sigma_{cp} - \big|o_P(n)\big|\Big\} \cap \Big\{\big|o_P(1)\big| > \frac{\eta\sigma_{cp}}{2}\Big\}\right\}$$

$$\leq P\Big\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1-\frac{\beta_{2,p}}{x}\big)+\eta\sigma_{cp}/2\Big\}$$
$$+P\Big\{\big|o_P(1)\big| > \frac{\eta\sigma_{cp}}{2}\Big\}$$

$$\leq P\Big\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1-\frac{\beta_{2,p}}{x}\big)\Big\}$$
$$-P\Big\{\sigma_{cp}x\big(1-\frac{\beta_{2,p}}{x}\big)+\eta\sigma_{cp}/2 \geq \sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1-\frac{\beta_{2,p}}{x}\big)\Big\}$$
$$+P\Big\{\big|o_P(1)\big| > \frac{\eta\sigma_{cp}}{2}\Big\}$$

$$\leq P\Big\{\sqrt{N}\big|H_e\big| > \sigma_{cp}x\big(1-\frac{\beta_{2,p}}{x}\big)\Big\}$$

(for N large enough, considering that $\sqrt{N}H_e$ converges in distribution to a random variable which has a distribution function $F_{H_e}(x) > 0$ and considering the convergence in probability to 0 of $o_P(1)$)

$$= P\Big\{\sqrt{N}\big|H_e\big| > \sigma_2 x\frac{1-\beta_{2,p}/x}{1+Ob(1)\gamma"_{2,p}}\Big\}$$

$$\leq P\Big\{\sqrt{N}\Big|\frac{P_e-L(I)p_e-D}{l}+\frac{Dl-Dp_e}{l^2}\Big| > \sigma_2 x\frac{1-\beta_{2,p}/x}{1+\gamma"_{2,p}}\Big\}. \ \blacksquare$$

9.5 Practical applications

In this section, we study how the assumptions of the previous theorems are used concretely and how one can apply them.

We are interested by the sequence of random bits $b^1(n')$ built in section 11.2 : we assume that $B^1(n')$ is a sequence of random bits satisfying

$$P\big\{B^1(n') = b \ \big| \ B^1(n'+j_2) = b_2, ..., B^1(n'+j_p) = b_p\big\}$$

$$= 1/2 + Ob(1)\epsilon = 1/2 + \frac{Ob(1)\alpha}{\sqrt{Nq_0}} \ ,$$

where $\alpha = 0.02$, N= 1.000.000 and $q_0 = 57 : \epsilon = \alpha/\sqrt{Nq_0} = 0.000002649$.

To simplify the notations, in this section we will replace Nq_0 by N. We keep the notations of the previous sections: we study the sequence X_n introduced in section 9.1 when it equal to the sequence $B^1(n')$. Thus, we have the following notations.

Notations 9.5.1 *In this section we assume that $X_n = B^1(n')$ is a sequence of random bits such that*

$$P\big\{B^1(n') = b \ \big| \ B^1(n'+j_2) = b_2, ..., B^1(n'+j_p) = b_p\big\} = 1/2 + \frac{Ob(1)\alpha}{\sqrt{N}} \ ,$$

with $\alpha = 0.02$, N= 57.000.000 : $\epsilon = \alpha/\sqrt{N} = 0.000002649$.

Now, the sequence C(j) defined in section 11.2 is Qd-dependent with Qd=22. Therefore $K_B = k_B = 0$ for $q \geq Qd$.

Obviously, all Borel set Bo_s are intervals I if p=1. Then, in this section we impose the following notations.

Notations 9.5.2 *In this section, we keep the notations of this chapter. But, we assume that Bo_1 is an interval I : $Bo_1 = I$.*

Then,
1) One can write $\epsilon_p = (1/2+\epsilon)^p - (1/2)^p \approx \frac{p\epsilon}{2^{p-1}} = \frac{2p\epsilon}{2^p}$.
2) On suppose $p \leq Log(N)/log(2)$.
Indeed, the study of P_e has none sense if $p \geq Log(N)/log(2)$: it is known that it is useleess to study $P\{\sqrt{N}|P_e - L(Bo)| \geq \sigma_B x\}$ for IID sequences when L(Bo) is too small.
It is no possible to test the independence of p variables in this case: the samples are too small. For example a sample with 3 elements do not allow to study the dependence between two random variables. Thus, if one uses the chi-squared test, a sample with size $N_0 \geq 2^p$ is necessary.

Example 9.5.1 *Assume that $N = 57.000.000$, $p \leq log(N)/log(2) < 25,7$.*

9.5.1 Study of Theorem 9

Now, we study the assumptions of theorem 9.

Hypothesis 9.5.1 *We assume also that the sequence $B^1(n')$ satisfies the theorem 9.*

Then, we have the following properties.

1) Obviously, $L^N(Bo) = (1/N)\sum_{n=1}^N L_n = \frac{1}{N}(\sum_{n=1}^N (L(Bo) + Ob(1)\epsilon_p) = L(Bo) + Ob(1)\epsilon_p$.

2) We remind $\sigma_B^2 \geq A(p)L(Bo)$.

3) Then, $\beta_{1,p} = \sqrt{N}[L^N(Bo) - L(Bo)]/\sigma_B(Bo) \leq \frac{\sqrt{N}\epsilon_p}{\sqrt{A(p)L(Bo)}}$.

Now, $\frac{\sqrt{N}\epsilon_p}{\sqrt{A(p)L(Bo)}} \approx \frac{2\sqrt{N}p\epsilon}{\sqrt{A(p)L(Bo)2^p}} \approx \frac{2\sqrt{N}p\epsilon}{A(p)^{1/2}2^{p/2}} \approx \frac{2p\alpha}{A(p)^{1/2}2^{p/2}}$.

Then, generally

$$\beta_{1,p} \leq \frac{2p\alpha}{A(p)^{1/2}2^{p/2}} \ . \tag{9.3}$$

4) Moreover,

$$\gamma_{1,p} = \frac{1}{2A(p)L(Bo)}\Big[(p^2-p+1)\Big(\epsilon_p + 2q\epsilon_{2p} + (1+2q)\big[2^{1-p}\epsilon_p + \epsilon_p^2\big]\Big) + k_B 0.1\epsilon_p\Big].$$

Then, if $k_B = 0$,

$$\gamma_{1,p} \approx \frac{(p^2-p+1)2^p}{2A(p)}\Big[\frac{2p\epsilon}{2^p} + 2q\frac{2(2p)\epsilon}{2^{2p}} + (1+2q)\big[2^{1-p}\frac{2p\epsilon}{2^p} + \frac{4p^2\epsilon^2}{2^{2p}}\big]\Big]$$
$$\approx \frac{(p^2-p+1)\epsilon 2^p}{2A(p)2^p}\Big[2p + 2q\frac{2(2p)}{2^p} + (1+2q)\big[2\frac{2p}{2^p} + \frac{4p^2\epsilon}{2^p}\big]\Big]$$
$$\approx \frac{(p^2-p+1)\alpha}{2A(p)\sqrt{N}}\Big[2p + (1+4q)\frac{4p}{2^p} + (1+2q)\frac{4p^2\epsilon}{2^p}\Big] \ .$$

Remark 9.5.2 *We impose the condition "$\epsilon = \alpha/\sqrt{N}$" because $\beta_{1,p} \leq \frac{\sqrt{N}\epsilon_p}{\sqrt{A(p)L(Bo)}}$.*

Therefore,

$$\frac{(p^2-p+1)\alpha}{2A(p)\sqrt{N}}\Big[2p + (1+4q)\frac{4p}{2^p} + (1+2q)\frac{4p^2\epsilon}{2^p}\Big]$$
$$\leq \frac{0.0000026(p^2-p+1)}{2A(p)}\Big[2p + (1+4q)\frac{4p}{2^p} + (1+2q)\frac{4p^2 0.0000026}{2^p}\Big]$$
$$\leq 0.0413 \ .$$

Remark that if $B^1(n')$ is Qd-dependent, under simple assumptions, $\frac{\sqrt{N}(P_e - L^N(Bo))}{\sigma_1(Bo)}$ has asymptotically a normal distribution. It is the case under the assumptions of our data.

Moreover, the $1_I(B^1(n'))$'s satisfy condition H_{mS} (cf Notation 5.1.5 and section 11.2.2). At last, by definition, $E\big\{\big[\sqrt{N}(P_e - L^N(I))\big]^2\big\} = \sigma_1(I)^2$ is indeed equal to $\sigma(u)^2/N$ introduced in the CLT : cf notations 5.1.4. Therefore, $\frac{\sqrt{N}(P_e - L^N(Bo))}{\sigma_1(Bo)}$ has asymptotically the N(0,1) distribution.

Moreover, if B_n is IID, $\sqrt{N}\frac{P_e - L(Bo)}{\sigma_B}$ has asymptotically the standard normal distribution.

We want that B_n behaves like an IID sequence.
We thus increase $P\Big\{\frac{\sqrt{N}|P_e-L(Bo)|}{\sigma_B} \geq x\Big\}$ by using theorem 9 :

$$P\Big\{\sqrt{N}\big|P_e - L(Bo)\big| \geq \sigma_B(Bo)x\Big\} \leq K_1\Big(\frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}}x\Big) \ .$$

Let H1 be the hypothesis 9.5.1. Under assumptions IID and H1, one has the following tables of increases of $P\Big\{\sqrt{N}\big|P_e - L(Bo)\big| \geq \sigma_B(Bo)x\Big\}$ regarded as function of (x,p).

(x,p)	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,10)	(1,15)	(1,20)
Under IID	0.317	0.317	0.317	0.317	0.317	0.317	0.317	0.317
Under H1	0.334	0.359	0.356	0.357	0.346	0.340	0.361	0.380

(x,p)	(2,1)	(2,2)	.(2,3)	(2,4)	(2,5)	(2,10)	(2,15)	(2,20)
Under IID	0.030	0.030	0.030	0.030	0.030	0.030	0.030	0.030
Under H1	0.049	0.052	0.051	0.052	0.053	0.050	0.061	0.073

9.5.2 Study of Theorem 10

We keep the notations of section 9.3 with $Bo_1 = I$.

General case

First, we study the assumptions of theorem 10.

Hypothesis 9.5.2 *We assume that the sequence $B^1(n')$ satisfies theorem 10.*

Then, the following properties hold.

Lemma 9.5.3 *The following equalities hold*
$D = L^N(I \otimes J) - L(I)l^N(J) = Ob(1)\epsilon L(J)$,
$D = L^N(I \otimes J) - L(I)l^N(J) = Ob(1)\epsilon l^N(J)$.

Proof By the lemma 4.2.1,
$\Big|E\{\big[1_I(X_n) - L(I)\big]1_J(X_{n+j})\}\Big| \leq \epsilon E\{1_J(X_{n+j}\}$.

Then, the following relations hold :

$$\begin{aligned}
&\left|L^N(I\otimes J) - L(I)l^N(J)\right| \\
&= \left|(1/N)\sum E\{1_I(X_n)1_J(X_{n+j})\} - (1/N)\sum L(I)E\{1_J(X_{n+j})\}\right| \\
&\leq (1/N)\left|\sum E\{\big[1_I(X_n) - L(I)\big]1_J(X_{n+j})\}\}\right|. \\
&\leq (1/N)\left|\sum \epsilon E\{1_J(X_{n+j}\}\right| \leq \epsilon(1/N)\sum E\{1_J(X_{n+j}\} = \epsilon l^N(J). \ \blacksquare
\end{aligned}$$

Lemma 9.5.4 *Under the hypothesis 9.3.3,* $\sigma_{cp}^2 \geq \frac{1/8}{L(J)}$.

Proof By definition σ_2^2 is the variance of $\sqrt{N}\Big[\frac{P_e - L(I)p_e - D}{l^N(J)} + \frac{Dl^N(J) - Dp_e}{l^N(J)^2}\Big]$ where $D = L^N(I\otimes J) - L(I)l^N(J)$. We denote σ_{cp}^2 instead of σ_2^2 when X_n is IID .

In this case, $L^N(I\otimes J) = L(I)L(J)$ and $l^N(J) = L(J)$. Therefore $D = 0$. Therefore σ_{cp}^2 is the variance of $\sqrt{N}\Big[\frac{P_e - L(I)p_e}{L(J)}\Big]$

Therefore, by our hypothesis 9.3.3, $\sigma_{cp}^2 \geq \frac{(1/4)L(I)L(J)}{L(J)^2} = \frac{(1/4)L(I)}{L(J)} = \frac{1/8}{L(J)}$. $\blacksquare$.

Lemma 9.5.5 *If* η *is small, approximately* $\beta_{2,p} \approx \frac{Ob(1)16\alpha}{2^p}$.

Proof We have $\beta_{2,p} = \frac{N^{1/2}D}{\sigma_{cp}l^N(J)} + \eta$ where $\eta > 0$.

Now,

$$\begin{aligned}
&\frac{N^{1/2}D}{\sigma_{cp}l^N(J)} \leq \frac{N^{1/2}\epsilon l^N(J)}{\sigma_{cp}l^N(J)} \leq \frac{N^{1/2}\epsilon}{\sigma_{cp}} \\
&\leq \frac{N^{1/2}\epsilon}{(1/8)/L(J)} \leq \frac{8N^{1/2}\alpha L(J)}{N^{1/2}} \leq \frac{16\alpha}{2^p} \ . \ \blacksquare
\end{aligned}$$

Lemma 9.5.6 *The following equalities hold :* $l^N(J) = L(J) + Ob(1)\epsilon_{p-1} = L(J)[1 + Ob(1)2\epsilon(p-1)]$.

Proof We have

$$l^N(J) = (1/N)\sum_n l_n = (1/N)\sum_n E\{1_J(X_{n+j}) = (1/N)\sum_N [L(J) + Ob(1)\epsilon_{p-1}]$$

$$= L(J) + Ob(1)\epsilon_{p-1} = L(J) + Ob(1)\frac{2\epsilon(p-1)}{2^{p-1}} = L(J)[1 + Ob(1)2\epsilon(p-1)] \ . \ \blacksquare$$

Lemma 9.5.7 *One supposes that* B_n *is Qd-dependent. Then,*

$$\xi_p \approx \frac{Ob(1)}{2}\Big[(p-1)[11p+16] + \frac{16(2q+1)(p^2-p+1)}{2^p}\Big] \ .$$

Proof We have

$$\epsilon\mathcal{M}(l) = 1/L(J) - 1/l = 1/L(J) - \frac{1}{L(J)[1+2Ob(1)\epsilon(p-1)]}$$

$$= [1/L(J)]\Big[1 - \frac{1}{1+2Ob(1)\epsilon(p-1)}\Big] = [1/L(J)]\frac{1+2Ob(1)\epsilon(p-1)-1}{1+2Ob(1)\epsilon(p-1)}$$

$$= Ob(1)[1/L(J)]\frac{2\epsilon(p-1)}{1-2\epsilon(p-1)} \ .$$

Then, $\mathcal{M}(l) = [1/L(J)]\frac{2Ob(1)(p-1)}{1-2\epsilon(p-1)}$.

For $q = Qd = 57$, we have $K_B = 0$. then

$$\xi_p = (1/2)\Bigg[2L(J)\mathcal{M}(l) + \frac{8L(J)}{l}(1-2L(I)) + (3/2)(p-1)L(J)\mathcal{M}(l)$$

$$+\frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$

$$+\frac{8[p(p-1)+K_B/10]L(J)}{l}$$

$$+\ \big[1+\epsilon]\frac{8L(J)[L(J)+\epsilon_{p-1}]}{l}(2q+1)(p^2-p+1)\Bigg]$$

$$= (1/2)\Bigg[2L(J)\frac{2(p-1)Ob(1)}{L(J)[1-2\epsilon(p-1)]} + 0 * \frac{8L(J)}{L(J)[1-2\epsilon(p-1)]}$$

$$+(3/2)(p-1)Ob(1)L(J)\frac{2(p-1)}{L(J)[1-2\epsilon(p-1)]}$$

$$+\frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$

$$+\frac{8[p(p-1)]L(J)Ob(1)}{L(J)[1-2\epsilon(p-1)]}$$

$$+\ \big[1+\epsilon]\frac{8L(J)[L(J)+\epsilon_{p-1}]Ob(1)}{L(J)[1-2\epsilon(p-1)]}(2q+1)(p^2-p+1)\Bigg]$$

$$= (1/2)\Bigg[\frac{4(p-1)Ob(1)}{1-2\epsilon(p-1)} + \frac{3(p-1)^2Ob(1)}{1-2\epsilon(p-1)}$$
$$+\frac{(p-1)}{[1-2(p-1)\epsilon]}\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$
$$+\frac{8p(p-1)Ob(1)}{1-2\epsilon(p-1)}$$
$$+\ \big[1+\epsilon]\frac{8[L(J)+\epsilon_{p-1}]Ob(1)}{1-2\epsilon(p-1)}(2q+1)(p^2-p+1)\Bigg]$$

$$= \frac{1}{2\big(1-2(p-1)\epsilon\big)}\Bigg[4(p-1)Ob(1)+3(p-1)^2Ob(1)$$
$$+(p-1)\Big[\frac{p+2+6\epsilon}{[1-2(p-1)\epsilon]}+13\Big]$$
$$+8p(p-1)Ob(1)$$
$$+[1+\epsilon]Ob(1)\big[8[L(J)+\epsilon_{p-1}]](2q+1)(p^2-p+1)\Bigg]$$

$$\approx \frac{1}{2}\Big[4(p-1)Ob(1)+3(p-1)^2Ob(1)+(p-1)[(p+2)+13]\Big]$$
$$+8p(p-1)Ob(1)+8Ob(1)L(J)(2q+1)(p^2-p+1)\Big]$$

$$\approx \frac{Ob(1)}{2}\Big[(p-1)\big[4+3(p-1)+[p+15]+8p\big]+\frac{16(2q+1)(p^2-p+1)}{2^p}\Big]$$

$$\approx \frac{Ob(1)}{2}\Big[(p-1)[12p+16]+\frac{16(2q+1)(p^2-p+1)}{2^p}\Big]$$

$$\approx \frac{Ob(1)}{2}\Big[(p-1)[12p+16]+\frac{1840(p^2-p+1)}{2^p}\Big]\ .\ \blacksquare$$

Lemma 9.5.8 *One supposes that B_n is Qd-dependent. Then,*

$$\theta_p \approx \frac{8\sigma_B(J)^2}{L(J)} + 16(p^2-3p+3)(p-1)\epsilon\Big[1+\frac{2}{2^{p-1}}+\frac{8q}{2^{p-1}}\Big] \ .$$

Proof We set $q = Qd = 57$. We have $K_B = k_B = 0$. Then, the following relations hold :

$$\theta_p = \frac{8\sigma_B(J)^2 L(J)}{l^2}$$
$$+(p^2-3p+3)\frac{8\epsilon_{p-1}L(J)}{l^2} + (p^2-3p+3)(2q)\frac{8\epsilon_{2p-2}L(J)}{l^2}$$
$$+(p^2-3p+3)(1+2q)\frac{8\big[2^{2-p}\epsilon_{p-1}+\epsilon_{p-1}^2\big]L(J)}{l^2} + \frac{0.8k_B\epsilon_{p-1}L(J)}{l^2}$$

$$\approx \frac{8\sigma_B(J)^2}{L(J)}$$
$$+\frac{8(p^2-3p+3)[2(p-1)\epsilon/2^{p-1}]}{L(J)} + (p^2-3p+3)(2q)\frac{8[2(2p-2)\epsilon/2^{2p-2}]}{L(J)}$$
$$+(p^2-3p+3)(1+2q)\frac{8\big[2^{2-p}[2(p-1)\epsilon/2^{p-1}]+[2(p-1)\epsilon/2^{p-1}]^2\big]}{L(J)}$$

$$\approx \frac{8\sigma_B(J)^2}{L(J)}$$
$$+16(p^2-3p+3)(p-1)\epsilon + (p^2-3p+3)(2q)\frac{8[2(2p-2)\epsilon]}{2^{p-1}}$$
$$+(p^2-3p+3)(1+2q)\frac{8\big[2[2(p-1)\epsilon]+[2(p-1)]^2\epsilon^2\big]}{2^{2(p-1)}L(J)}$$

$$\approx \frac{8\sigma_B(J)^2}{L(J)}$$
$$+16(p^2-3p+3)(p-1)\epsilon + (p^2-3p+3)(2q)\frac{32(p-1)\epsilon}{2^{p-1}}$$
$$+(p^2-3p+3)(1+2q)\frac{32(p-1)\epsilon+32(p-1)^2\epsilon^2\big]}{2^{p-1}}$$

$$\approx \frac{8\sigma_B(J)^2}{L(J)}$$

$$+(p^2-3p+3)(p-1)\epsilon\Big[16+\frac{32(2q)}{2^{p-1}}+(1+2q)\frac{32+32(p-1)\epsilon\big]}{2^{p-1}}\Big]$$

$$\approx \frac{8\sigma_B(J)^2}{L(J)}+16(p^2-3p+3)(p-1)\epsilon\Big[1+\frac{4q}{2^{p-1}}+\frac{2(1+2q)}{2^{p-1}}\Big]$$

$$\approx \frac{8\sigma_B(J)^2}{L(J)}+16(p^2-3p+3)(p-1)\epsilon\Big[1+\frac{2}{2^{p-1}}+\frac{8q}{2^{p-1}}\Big]\ .\ \blacksquare$$

Then, one will adopt the following assumption : cf section 9.7.

Hypothesis 9.5.3 *One supposes that* $\sigma_B(J)^2 \leq 4L(J)$.

Study under the assumptions of section 11.2

Now, we study the numerical results obtained with the data used in section 11.2: $\epsilon = \alpha/\sqrt{N} = 0.000002649$, $L(J) = 2^{1-p}$.

Then, $2\gamma_{2,p} = \epsilon\Big(2\xi_p + 2[1+\epsilon\xi_p]\sqrt{\theta_p} + \epsilon\theta_p\Big)\Big]$ and $\beta_{2,p} \approx \frac{16\alpha_B}{2^p}$.

Then, for $p \geq 3$, we have the following increases.

p	3	5	10	15	20	32
ξ_p	1176	987.3	681.5.	1275.3	2242.5.	2959.1
θ_p	32	33	40	58.3	98.8	136.4
$\gamma_{2,p}$	0.018	0.015	0.010	0.0023	0.035	0.047

Remark that, if B_n is Qd-dependent, in many cases, the distribution of de $\frac{\sqrt{N}(\frac{P_e}{p_e}-L(I))}{\sigma_2(I)}$ is close to a normal distribution. It is the case under the assumptions of our data.

By refering to the proof of theorem 10 in section 9.3.7, we deduce of it that $\frac{\sqrt{N}(\frac{P_e}{p_e}-L(I))}{\sigma_2(I)}$ has asymptotically a normal distribution if $\sqrt{N}\Big(\frac{P_e-L(I)p_e-D}{l}\Big)+\sqrt{N}\frac{Dl-Dp_e}{l^2}$ has asymptotically a normal distribution.

It is the same when B_n is IID.

Proposition 9.5.1 *Suppose that B_n is IID. Then, $\sqrt{N}\frac{\frac{P_e}{p_e}-L(I)}{\sigma_2(I)}$ has asymptotically the standard normal distribution*

Proof By our assumptions, $D = 0$.

We keep notations 9.1.1 : $o_P(1) \xrightarrow{P} 0$. Then,

$$\sqrt{N}\Big(\frac{P_e}{p_e} - L(I)\Big) = \sqrt{N}\Big(\frac{P_e - L(I)p_e}{p_e}\Big)$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e}{l}\Big) + \sqrt{N}\big(P_e - L(I)p_e\big)\Big[\frac{1}{p_e} - \frac{1}{l}\Big]$$

$$= \sqrt{N}\Big(\frac{P_e - L(I)p_e}{l}\Big) - \sqrt{N}(p_e - l)\ \frac{P_e - L(I)p_e}{p_e l}\ .$$

By our assumptions $\sqrt{N}(p_e - l)$ is asymptotically normal. Moreover, $P_e - L(I)p_e \xrightarrow{P} 0$ and $p_e l \xrightarrow{P} l^2$.

Then,

$$\sqrt{N}\Big(\frac{P_e}{p_e} - L(I)\Big) = \sqrt{N}\Big(\frac{P_e - L(I)p_e}{l}\Big) + o_P(1).$$

Moreover, $\sqrt{N}\Big(\frac{P_e - L(I)p_e}{\sigma_{cp} l}\Big)$ is asymptotically normally distributed. ■

Then, if B_n is IID, $\sqrt{N}\frac{\frac{P_e}{p_e}-L(I)}{\sigma_{cp}}$ has asymptotically the standard normal distribution.

We want that B_n behaves like an IID sequence.

Then, we increase $P\Big\{\frac{\sqrt{N}|P_e/p_e - L(I)|}{\sigma_{cp}} \geq x\Big\}$ by using theorem 10 :

$$P\Big\{\sqrt{N}\Big|\frac{P_e}{p_e} - L(I)\Big| > \sigma_{cp}\ x\Big\} \leq K_2\Big(\frac{1 - \beta_{2,p}/x}{1 + \gamma_{2,p}}x\Big)$$

if N is large enough.

Then, we have the following table of $P\Big\{\frac{\sqrt{N}|P_e/p_e - L(I)|}{\sigma_{cp}} \geq x\Big\}$ regarded as function of (p,x)

(x,p))	(1,2)	(1,3)	(1,4)	(1,5)	(1,10)	(1,15)	(1,20)
Sous IID	0.317	0.317	0.317	0.317	0.317	0.317	0.317
Sous H1	0.332	0.334	0.330	0.324	0.319	0.335	0.346

(x,p)	(2,2)	(2,3)	(2,4)	(2,5)	(2,10)	(2,15)	(2,20)
Sous IID	0.030	0.030	0.030	0.030	0.030	0.030	0.030
Sous H1	0.053	0.055	0.052	0.049	0.049	0.050	0.054

9.6 First assumption about variances

In this section one wants to know if assumption 9.3.3 holds in all the cases. One thus places oneself under the following assumptions.

Hypothesis 9.6.1 *One supposes that X_n is IID, $X_n \in \{0,1\}$. Then, one supposes that Bo is an interval I : Bo=I .*

We set $V2 = NE\big\{(P_e - L(I)p_e)^2\big\}$.

Thus, one wants to know if the following assumption is satisfied.

$$V2 = NE\big\{(P_e - L(I)p_e)^2\big\} \geq (1/4)L(I)L(J) \ .$$

First, this assumption is always satisfied for some sequences j_s , for example if j_s is increasing : cf 9.6.2 below.

As a matter of fact, this assumption is probably all the time satisfied. But to prove it risks to be complicated considering the complexity of each case as soon as p increases a little. For better revealing the reasons why assumption 9.3.3 probably holds always, one uses the following lemma.

Lemma 9.6.1 *Under the previous assumptions,*

$$N^2E\big\{(P_e - L(I)p_e)^2\big\}$$

$$= \sum_n E\big\{[1_I(X_n) - L(I)]^2 1_J(X_{n+j})\big\}$$

$$+\sum_n \sum_{s:\ s\neq 1, m=n+j_s, n=m+j_t} E\Big\{\Big[\prod_{u\neq s} 1_J(X_{n+j_u})\Big]\Big[\prod_{v\neq t} 1_J(X_{m+j_v})\Big]\Big\}.....$$

$$.....E\Big\{\big[1_I(X_n) - L(I)\big]1_{J_s}(X_n)\Big\}E\Big\{\big[1_I(X_m) - L(I)\big]1_{J_t}(X_m)\Big\} \ .$$

Proof We have the following equalities

$$N^2E\big\{(P_e - L(I)p_e)^2\big\}$$
$$= N^2E\big\{\big[(1/N)\sum_n 1_I(X_n)1_J(X_{n+j}) - L(I)(1/N)\sum_n 1_J(X_{n+j})\big]^2\big\}$$

$$= E\Big\{ \Big[\sum_n [1_I(X_n) - L(I)] 1_J(X_{n+j}) \Big]^2 \Big\}$$

$$= \sum_{n,m} E\Big\{ [1_I(X_n) - L(I)] 1_J(X_{n+j}) [1_I(X_m) - L(I)] 1_J(X_{m+j}) \Big\}$$

$$= \sum_n E\Big\{ [1_I(X_n) - L(I)]^2 1_J(X_{n+j}) \Big\}$$

$$+ \sum_n \sum_{s:\ s \neq 1, m = n + j_s, n = m + j_t} E\Big\{ \Big[\prod_{u \neq s} 1_J(X_{n+j_u}) \Big] \Big[\prod_{v \neq t} 1_J(X_{m+j_v}) \Big] \Big\} .....$$

$$..... E\Big\{ [1_I(X_n) - L(I)] 1_{J_s}(X_n) \Big\} E\Big\{ [1_I(X_m) - L(I)] 1_{J_t}(X_m) \Big\} . \blacksquare$$

Therefore, hypothesis 9.3.3 is statisfied if j_s is increasing.

Lemma 9.6.2 *One supposes that the sequence j_s is increasing. Then*

$$N E\Big\{ (P_e - L(I) p_e)^2 \Big\} = L(I)[1 - L(I)] L(J)$$

Proof We have $\{ s \neq 1 :\ m = n + j_s, n = m + j_t \} = \emptyset$. Therefore,

$$N^2 E\Big\{ (P_e - L(I) p_e)^2 \Big\} = \sum_n E\Big\{ [1_I(X_n) - L(I)]^2 1_J(X_{n+j}) \Big\}$$

$$= L(J) \sum_n E\Big\{ [1_I(X_n) - L(I)]^2 \Big\} = L(J) \sum_n \Big(L(I) - L(I)^2 \Big) . \blacksquare$$

In order to study the general case, one rewrites the sequence j_s by the following way.

Notations 9.6.1 *One rewrites the sequence j_s by using to sequences $j'_s > 0$ and $j"_t > 0$, where $\{j_s\} = \{-j"_t\} \cup \{j'_s\}$: that is the sequence $\{j_s\} = \{...., -j"_2, -j"_1, j_1, j'_1, j'_2, ...\}$.*

One rewrites the sequence J_s by using two sequences J'_s and $J"_t$ corresponding to the sequences j'_s and $-j"_t$: $\{.....J"_{-2}, J"_{-1}, J_0, J'_1, J'_2, ...\}$.

With these notations, one has the following property.

Lemma 9.6.3 *The following equality hold :*

$$E\Big\{ [1_I(X_n) - L(I)] 1_{J"}(X_{m+j_t}) \Big\} E\Big\{ [1_I(X_m) - L(I)] 1_{J'}(X_{n+j_s}) \Big\} = \delta(1/16) ,$$

where $\delta = 1$ si $J' = J"$ and $\delta = -1$ if $J" \neq J'$.

Proof We have

$$E\Big\{[1_I(X_n) - L(I)]1_{J"}(X_{m+j_t})\Big\}E\Big\{[1_I(X_m) - L(I)]1_{J'}(X_{n+j_s})\Big\}$$

$$= E\Big\{[1_I(X_n) - L(I)]1_{J"}(X_n)\Big\}E\Big\{[1_I(X_m) - L(I)]1_{J'}(X_m)\Big\} \ .$$

Moreover,

$$E\Big\{[1_I(X_n) - L(I)]1_{J"}(X_n)\Big\}$$

$$= E\Big\{[1_{I\cap J"}(X_n) - L(I)L(J")\Big\}$$

$$= E\Big\{[1_{I\cap J"}(X_n) - 1/4\Big\}$$

= 0 - 1/4 if $I \cap J" = \emptyset$, i.e. if $I \neq J"$
= 1/2 -1/4 if $I = J"$.

We deduce the result. ■

Therefore,

$$NE\big\{(P_e - L(I)p_e)^2\big\}$$

$$= (1/4)L(J) + \sum_n \sum_{s:s\neq, m=n+j_s, n=m+j'_t} (\delta_s/16)\mathcal{H}(s)L(J) \ ,$$

where $\mathcal{H}(s)$ is a function of s.

If $\mathcal{H}(s) > 0$, that increases the possibility that $NE\big\{(P_e - L(I)p_e)^2\big\} > (1/4)L(I)L(J)$.

If $\mathcal{H}(s) < 0$, that decreases it. It will be the case when $J' \neq J"$.

In all the examples which we studied, we have obtained $N'E\big\{(P_e - L(I)p_e)^2\big\} \geq (1/4)L(I)L(J)$. It is thus very probable that this assumption is always satisfied.

To support this assertion, one now will study some examples. One will represent them by graphs with the following definitions.

Symbols 9.6.4 *In the following graphs, for n fixed, the first line represents the $n + j_s$, s=1,2,...,p, around point n.*

The points "x" represent the points $t \neq n + j_s$, s=1,2,...,p.

The symbol "O" represents the point $n : n = n + j_1 = n$ avec $I = \{0\}$

The symbol "0" represent the points $n + j_s : J_s = \{0\}$,

The symbol "1" represent the points $n + j_s : J_s = \{1\}$.

The second line represents, for a "m" fixed, the $m + j_s$, s=1,2,...,p, around the point m with the same notations : e.g. the points "x" of the second line represent the points $t \neq m + j_s$, s=1,2,...,p.

We set Mini = L(I)L(J)/4 = L(J)/8.

In the first examples, $J'_t = \{1\}$: that corresponds to the j_s positive. Moreover, $J"_r = \{0\}$ that corresponds to the j_s negative. Finally one supposes $I = \{0\}$. In example 9.6.6 $j_s = 0 \pm 1$. Therefore, p=3.

In theses example we compute $(\delta_s/16)\mathcal{H}(s)L(J)$ for each case, i.e. for each values of "s".

Example 9.6.5 *It is the case where j_s is increasing. In this case, the variance V2 is equal to L(J)/4 .*

Example 9.6.6 *One supposes $\{j_s\} = \{-1, 0, 1\}$, $I = \{0\}$, $J"_t = \{0\}$ for t = -1, $J'_r = \{1\}$ for r=1 .*

Therefore the points of each line are represented by x 0 O 1 x .
Moreover, $L(J) = 2^{-2}$, $L(I)L(J) = 2^{-3}$, $(1/4)L(I)L(J) = 2^{-5}$.
Case 1 : Result = $(-1/16)2^{-2}$.

x	0	O	1
0	O	1	x

Case 2 : Result = (1/4)L(J)

0	O	1
0	O	1

Therefore, altogether : $V2 = (1/4)L(J) - 2 * (1/16)2^{-2} = (1/16) - (1/32) = 1/32 = (1/4)L(I)L(J) = Mini.$ ■

Example 9.6.7 *One supposes $\{j_s\} = \{-2, -1, 0, 1, 2\}$, $I = \{0\}$, $J"_t = \{0\}$ for t = -2,-1 $J'_r = \{1\}$ for r=1,2 .*

Therefore the points of each line are represented by x 0 0 O 1 1 x .
Moreover, $L(J) = 2^{-4}$, $L(I)L(J) = 2^{-5}$, $(1/4)L(I)L(J) = 2^{-7}$.
Case 1 : Result = 0

x	x	0	0	O	1	1
0	0	O	1	1	x	x

Case 2 : Result = $(-1/16)2^{-4}$.

x	0	0	O	1	1
0	0	O	1	1	x

Case 3 : Result = (1/4)L(J)

0	0	O	1	1
0	0	O	1	1

Therefore, altogether : $V2 = (1/4)L(J) - 2 * (1/16)2^{-4} = L(J)[(1/4) - (1/8)] = L(J)[(1/8)] = L(I)L(J)/4 = Mini.$ ■

Example 9.6.8 *One supposes* $\{j_s\} = \{-2, -1, 0, 1, 2\}$, $I = \{0\}$, $J"_{-2} = J'_1 = \{0\}$ *and* $J"_{-1} = J'_2 = \{1\}$.

Therefore the points of each line are represented by x 0 1 O 0 1 x
Moreover, $L(J) = 2^{-4}$, $L(I)L(J) = 2^{-5}$, $(1/4)L(I)L(J) = 2^{-7}$.

Case 1 : Result = 0

x	x	0	1	O	0	1
0	1	O	0	1	x	x

Case 2 : Result = 0

x	0	1	O	0	1
0	1	O	0	1	x

Case 3 : Result = (1/4)L(J)

0	1	O	0	1
0	1	O	0	1

Case 4 : Result = 0

0	1	O	0	1	x
x	0	1	O	0	1

Case 5 : Result = 0

0	1	O	0	1	x	x
x	x	0	1	O	0	1

Therefore, altogether : $V2 = (1/4)L(J) > Mini.$ ■

Example 9.6.9 *One supposes* $\{j_s\} = \{-2, -1, 0, 1, 2\}$, $I = \{0\}$, $J"_{-2} = J'_2 = \{0\}$ *and* $J"_{-1} = J'_1 = \{1\}$.

Therefore the points of each line are represented by x 0 1 O 1 0 x . Moreover, $L(J) = 2^{-4}$, $L(I)L(J) = 2^{-5}$, $(1/4)L(I)L(J) = 2^{-7}$.

Case 1 : Result = (1/16)(1/2)L(J)

x	x	0	1	O	1	0
0	1	O	1	0	x	x

Case 2 : Result = 0

x	0	1	O	1	0
0	1	O	1	0	x

Case 3 : Result = (1/4)L(J)

0	1	O	1	0
0	1	O	1	0

Case 4 : Result = 0

0	1	O	1	0	x
x	0	1	O	1	0

Case 5 : Result = (1/16)(1/2)L(J)

0	1	O	1	0	x	x
x	x	0	1	O	1	0

Therefore, altogether : $V2 = (1/4)L(J) + (1/16)L(J) > Mini.$ ■

Example 9.6.10 *One supposes* $\{j_s\} = \{-3, -1, 0, 1, 2\}$, $I = \{0\}$, $J"_{-2} = J'_2 = \{0\}$ *and* $J"_{-1} = J'_1 = \{1\}$.

Therefore the points of each line are represented by par x 0 x 1 O 1 0 x .
Moreover $L(J) = 2^{-4}$, $L(I)L(J) = 2^{-5}$, $(1/4)L(I)L(J) = 2^{-7}$.

Case 1 : Result = 0

x	x	x	0	x	1	O	1	0
0	x	1	O	1	0	x	x	x

Case 2 : Result = 0

x	x	0	x	1	O	1	0
0	x	1	O	1	0	x	x

Case 3 : Result = 0

x	0	x	1	O	1	0
0	x	1	O	1	0	x

Case 4 : Result = (1/4)L(J)

0	x	1	O	1	0
0	x	1	O	1	0

Case 5 : Result = 0

0	x	1	O	1	0	x
x	0	x	1	O	1	0

Case 6 : Result = 0

0	x	1	O	1	0	x	x
x	x	0	x	1	O	1	0

Case 7 : Result = 0

0	x	1	O	1	0	x	x	x
x	x	x	0	x	1	O	1	0

Therefore, altogether : $V2 = (1/4)L(J) > Mini.$ ■

Example 9.6.11 *One supposes* $\{j_s\} = \{-2, -1, 0, 1, 2\}$, $I = \{0\}$, $J"_{-2} = J"_{-1} = J'_2 = \{0\}$ *and* $J'_1 = \{1\}$.

Therefore the points of each line are represented by x 0 0 O 1 0 x . Moreover, $L(J) = 2^{-4}$, $L(I)L(J) = 2^{-5}$, $(1/4)L(I)L(J) = 2^{-7}$.

Case 1 : Result = 0

x	x	0	0	O	1	0
0	0	O	1	0	x	x

Case 2 : Result = 0

x	0	0	O	1	0
0	0	O	1	0	x

Case 3 : Result = (1/4)L(J)

0	0	O	1	0
0	0	O	1	0

Case 4 : Result =0

0	0	O	1	0	x
x	0	0	O	1	0

Case 5 : Result =0

0	0	O	1	0	x	x
x	x	0	0	O	1	0

Therefore, altogether : $V2 = (1/4)L(J) \geq Mini.$ ■

Example 9.6.12 *One supposes* $\{j_s\} = \{-3, -2, -1, 0, 1, 2, 3\}$, $I = \{0\}$, $J"_t = \{0\}$ *for t = -3,-2,-1,* $J'_r = \{1\}$ *for r=1,2,3 .*

Therefore the points of each line are represented by x x x 0 0 0 O 1 1 1 .

Moreover, $L(J) = 2^{-6}$, $L(I)L(J) = 2^{-7}$, $(1/4)L(I)L(J) = 2^{-9}$.

Case 1 : Result = 0.

x	x	x	0	0	0	O	1	1	1
0	0	0	O	1	1	1	x	x	x

Case 2 : Result = 0.

x	x	0	0	0	O	1	1	1
0	0	0	O	1	1	1	x	x

Case 3 : Result $=(1/4)(-1/4)2^{-6} = -(1/16)L(J)$

x	0	0	0	O	1	1	1
0	0	0	O	1	1	1	x

Case 4 : Result = $(1/4)L(J)$

0	0	0	O	1	1	1
0	0	0	O	1	1	1

Therefore, altogether : $V2 = (1/4)L(J) - 2 * (1/16)L(J) = (1/8)L(J) = (1/4)L(I)L(J)$ = Mini .■

Example 9.6.13 *One supposes* $\{j_s\} = \{-4, -3, -1, 0, 1, 2, 3\}$, $I = \{0\}$, $J"_t = \{0\}$ *for t = -3,-2,-1* $J'_r = \{1\}$ *for r=1,2,3 .*

Therefore the points of each line are represented by x x x 0 0 x 0 O 1 1 1 .

Moreover, $L(J) = 2^{-6}$.

Case 1 : Result = 0.

x	x	x	0	0	x	0	O	1	1	1
0	0	x	0	O	1	1	1	x	x	x

Case 2 : Result = 0

x	x	0	0	x	0	O	1	1	1
0	0	x	0	O	1	1	1	x	x

Case 3 : Result = $(1/16)2^{-7} = -(1/16)L(J)/2$

x	0	0	x	0	O	1	1	1
0	0	x	0	O	1	1	1	x

Case 4 : Result = $(1/4)L(J)$

0	0	x	0	O	1	1	1
0	0	x	0	O	1	1	1

Case 25 : Result = $(1/16)2^{-7} = -(1/16)L(J)/2$

0	0	x	0	O	1	1	1	x
x	0	0	x	0	O	1	1	1

Case 6 : Result = 0

0	0	x	0	O	1	1	1	x	x
x	x	0	0	x	0	O	1	1	1

Case 7 : Result = 0

0	0	x	0	O	1	1	1	x	x	x
x	x	x	0	0	x	0	O	1	1	1

Therefore, altogether : $V2 = L(J)/4 - 2 * (1/16)L(J)/2 = L(J)/4 - L(J)/16 = 3L(J)/16 = (3/2)[L(I)L(J)/4] > Mini.$ ■

Example 9.6.14 *One supposes* $\{j_s\} = \{-4, -3, -1, 0, 1, 3, 4\}$, $I = \{0\}$, $J"_t = \{0\}$ *for t = -3,-2,-1* $J'_r = \{1\}$ *for r=1,2,3 .*

Therefore the points of each line are represented by x x x 0 0 x 0 O 1 x 1 1 .

Moreover, $L(J) = 2^{-6}$.

Case 1 : Result =0

x	x	x	x	0	0	x	0	O	1	x	1	1
0	0	x	0	O	1	x	1	1	x	x	x	x

Case 2 : Result $= -(1/16)2^{-8} = -(1/16)L(J)/4$

x	x	x	0	0	x	0	O	1	x	1	1
0	0	x	0	O	1	x	1	1	x	x	x

Case 3 : Result = 0

x	x	0	0	x	0	O	1	x	1	1
0	0	x	0	O	1	x	1	1	x	x

Case 4 : Result $= -(1/16)2^{-8} = -(1/16)L(J)/4$

x	0	0	x	0	O	1	x	1	1
0	0	x	0	O	1	x	1	1	x

Case 5 : Result $= L(J)/4$

0	0	x	0	O	1	x	1	1
0	0	x	0	O	1	x	1	1

Therefore, altogether : $V2 = L(J)/4 - 4 * (1/16)L(J)/4 = L(J)/4 - L(J)/16 = 3L(J)/16 > Mini.$ ■

Example 9.6.15 *One supposes* $\{j_s\} = \{-6, -3, -1, 0, 1, 3, 6\}$, $I = \{0\}$, $J"_t = \{0\}$ *for t = -3,-2,-1* $J'_r = \{1\}$ *for r=1,2,3 .*

Therefore the points of each line are represented by
x x x 0 x x 0 x 0 O 1 x 1 x x 1 .

Moreover, $L(J) = 2^{-6}$.

Case 1 : Result = 0

x	x	x	x	x	x	0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1	x	x	x	x	x	x

Case 2 : Result = 0 (because O)

x	x	x	x	x	0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1	x	x	x	x	x

Case 3 : Result = 0 (because O)

x	x	x	x	0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1	x	x	x	x

Case 4 : Result = $-(1/16)2^{-8} = -(1/16)L(J)/4$

x	x	x	0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1	x	x	x

Case 5 : Result = 0 (because O)

x	x	0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1	x	x

Case 6 : Result = $-(1/16)2^{-10} = -(1/16)L(J)/16$

x	0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1	x

Case 7 : Result = $L(J)/4$

0	x	x	0	x	0	O	1	x	1	x	x	1
0	x	x	0	x	0	O	1	x	1	x	x	1

Therefore, altogether : $V2 = L(J)/4 - 2*(1/16)L(J)/4 - 2*(1/16)L(J)/16 = L(J)/4 - (5/8)L(J)/16 = (L(J)/4)[1 - (5/32)] = (L(J)/4)(27/32) > Mini.$ ■

Example 9.6.16 *One supposes* $\{j_s\} = \{-3, -2, -1, 0, 1, 2, 3\}$, $I = \{0\}$, $J"_t = \{0\}$ *for t = -3,-2,* $J"_t = \{1\}$ *for t = -1,* $J'_r = \{0\}$ *for r=1 ,* $J'_r = \{1\}$ *for r=2,3 .*

Therefore the points of each line are represented by x x x 0 0 1 O 0 1 1 .

Moreover, $L(J) = 2^{-6}$, $L(I)L(J) = 2^{-7}$, $(1/4)L(I)L(J) = 2^{-9}$.

Case 1 : Result = $(-1/16)2^{-8}$

x	x	x	0	0	1	O	0	1	1
0	0	1	O	0	1	1	x	x	x

Case 2 : Result = 0.

x	x	0	0	1	O	0	1	1	x
0	0	1	O	0	1	1	x	x	x

Case 3 : Result = 0

x	0	0	1	O	0	1	1	x	x
0	0	1	O	0	1	1	x	x	x

Case 4 : Result =$(1/4)L(J)$

0	0	1	O	0	1	1	x	x
0	0	1	O	0	1	1	x	x

Therefore, altogether : $V2 = L(J)/4 - 2 * (-1/16)2^{-8} = L(J)[(1/4) - (1/32)] = [L(J)L(I)/4][2 - (1/4)] > Mini.$ ■

Example 9.6.17 *One supposes* $\{j_s\} = \{-3, -2, -1, 0, 1, 2, 3\}$, $I = \{0\}$, $J"_t = \{0\}$ *for t = -3,-2,-1,* $J'_r = \{0\}$ *for r=1,* $J'_r = \{1\}$ *for r=2,3* .

Therefore the points of each line are represented by x x x 0 0 0 O 0 1 1 .

Moreover, $L(J) = 2^{-6}$

Case 1 : Result = 0

x	x	x	0	0	0	O	0	1	1
0	0	0	O	0	1	1	x	x	x

Case 2 : Result = 0

x	x	0	0	0	O	0	1	1
0	0	0	O	0	1	1	x	x

Case 3 : Result = 0

x	0	0	0	O	0	1	1
0	0	0	O	0	1	1	x

Case 4 : Result = $(1/4)L(J)$

0	0	0	O	0	1	1
0	0	0	O	0	1	1

Case 5 : Result = 0

0	0	0	O	0	1	1	x
x	0	0	0	O	0	1	1

Case 6 : Result = 0

0	0	0	O	0	1	1	x	x
x	x	0	0	0	O	0	1	1

Case 7 : Result = 0

0	0	0	O	0	1	1	x	x	x
x	x	x	0	0	0	O	0	1	1

Therefore, altogether : $V2 = L(J)/4 > Mini.$ ■

Example 9.6.18 *Example 9.6.6 is made more complex.*

In order to make more complex this example, one adds "0" and "1".

The worst case is that where it is necessary to subtract $2*L(J)/16 = L(J)/8$. Then, it is understood that it is necessary to add 0 on the left and 1 on the right in order to remain in the case where $V2$ is minimal. If not, the result of this addition will be null. On the other hand, if one adds n "x" between the 0 or the 1, one will decrease the value which one has to substract: one will obtain $[2 * L(J)/16]/2^{n'}$, $1 \leq n' \leq n$.

Finally, one always obtains $V2 \geq (1/8)L(J)$.

But this is not a complete proof. If one wants to be surer about used result, one can choose increases weaker, for example $V2 \geq (1/16)L(J)$.

9.7 Second assumption about variances

9.7.1 Study of variances

Now we study the hypothesis 9.5.3. We keep the notations of section 9.3 : X'_n is an IID sequence of bits and $J = J_1 \otimes .... \otimes J_{p'} \subset \{0,1\}^{p'}$ where p'=p-1. One reminds that $\sigma_B(J)^2 = (1/N)E\Big\{\Big(\sum_{n=1}^{N}(1_J(X'_n) - L(J)]\Big)^2\Big\}$.

The strongest increase of the assumption 9.5.3 takes place when all the J_i's are identical, which we will thus suppose.

Then, in this section one wants to prove the following assumption.

Hypothesis 9.7.1 *Let p'=p-1. One supposes that $J_s = \{b_s\}$, $b_s \in \{0,1\}$. Then, one supposes that $\sigma_B(J)^2 \leq 4L(J)$.*

In this section, one will understand why this assumption seems correct.

First, by the proof of lemma 9.2.4, we know that
$\sigma_B^2(J) = (1/N)\sum_{n=1}^{N}\Big(\sum_{m\in H(n)} E\big\{1_J(X'_n)1_J(X'_m)\big\} - \sum_{m\in H(n)} L(J)^2\Big)$.

Moreover, by lemma 9.2.3 , $\sigma_B(J)^2 \leq (p^2 - p + 1)L(J)$.

Of course the hypothesis 9.7.1 is a much better increase. One will understand why it holds.

First, it holds for p'=1, $\sigma_B(J)^2 = L(J)[1 - L(J)] = L(J)/2$.

In the general case, there is first the following proposition.

Proposition 9.7.1 *Let $l_0 \in \mathbb{N}^*$. Suppose that $j_s = (s-1)l_0$ for $s = 1, 2, ..., p'$. Then*

$$\sum_{m \in H(n)} E\big\{1_J(X'_n)1_J(X'_m)\big\} \leq 3L(J) \ .$$

Proof Let $\mathcal{G}_{n,m} = \{n + j_t \mid n + j_t = m + j_s,\ s = 1, .., p';\ t = 1, .., p'\}$. Then, for $n \leq m$,

if $m = n + (p'-1)l_0$, $\mathcal{G}_{n,m} = \{n + (p'-1)l_0\}$,
if $m = n + (p'-2)l_0$, $\mathcal{G}_{n,m} = \{n + (p'-2)l_0, n + (p-1')l_0\}$
.....
if $m = n + rl_0$, $0 \leq r < p'$, $\mathcal{G}_{n,m} = \{n + rl_0, n + (r+1)l_0, ..., n + (p'-1)l_0\}$
if $m \neq n + rl_0$, $0 \leq r < p'$ $\mathcal{G}_{n,m} = \emptyset$.

Suppose $J_s = \{b\}$ for s=1,2,...,p', $b \in \{0, 1\}$. Then, because $L(J) = 1/2^{p'}$,
if $m = n + (p'-1)l_0$, $E\big\{1_J(X'_n)1_J(X'_m)\big\} = L(J)/2^{p'-1}$
if $m = n + (p'-2)l_0$, $E\big\{1_J(X'_n)1_J(X'_m)\big\} = L(J)/2^{p'-2}$
.....
if $m = n + rl_0$, $E\big\{1_J(X'_n)1_J(X'_m)\big\} = L(J)/2^r$
if $m \neq n + rl_0$, $E\big\{1_J(X'_n)1_J(X'_m)\big\} = L(J)/2^{p'}$.

if $n \geq m$, we obtain the same type of results.

Then,

$$\sum_{m \in H(n)} \big[E\big\{1_J(X'_n)1_J(X'_m)\big\} - L(J)^2\big]$$
$$\leq L(J) + 2L(J)\big[1/2 + 1/2^2 + ... + 1/2^{p'-1}\big]$$
$$\leq L(J) + 2L(J) = 3L(J) \ .\blacksquare$$

In this case, $card(H(n)) = 2p' - 1$. One can understand by numerical studies that it is the case where $\sum_{m \in H(n)} E\big\{1_J(X'_n)1_J(X'_m)\big\}$ is maximum.

It is understood easily that this result becomes widely usable in the following way.

Proposition 9.7.2 *Suppose that $j_s = (s-1)l_0$ for $s = -p_1, -p_1 + 1, ... - 1, 0, 1, 2, ..., p_2$ where $p_S > 0$. Then*

$$\sum_{m \in H(n)} E\big\{1_J(X'_n)1_J(X'_m)\big\} \leq 3L(J) \ .$$

Now, we study the minimal case.

Proposition 9.7.3 *Let $p' \geq 2$. Suppose $J_s = \{b\}$ for s=1,2,...,p. Suppose $card(H^*(n)) = p'^2 - p'$. In this case,*

$$\sum_{m \in H(n)} E\{1_J(X'_n)1_J(X'_m)\} \leq [L(J)/2](p'^2 - p')/2^{p'}) + L(J) \leq (11/8)L(J).$$

Proof If $card(H^*(n)) = p'^2 - p'$, then, $E\{1_J(X'_n)1_J(X'_m)\} = L(J)^2/2$: otherwise $card(H^*(n)) < p'^2 - p'$ (simplest to understand it is to make graphs).

Therefore,

$$\sum_{m \in H^*(n)} E\{1_J(X'_n)1_J(X'_m)\}$$

$$= (p'^2 - p')L(J)^2/2 = [(p'^2 - p')/2^{p'}]L(J)/2 \leq (3/8)L(J).$$

Therefore,

$$\sum_{m \in H(n)} E\{1_J(X'_n)1_J(X'_m) = (p'^2 - p')/2^{p'})(L(J)/2) + L(J) \leq (11/8)L(J). \blacksquare$$

It is the case where $\sum_{m \in H(n)} E\{1_J(X'_n)1_J(X'_m)\}$ is minimal.

Finally, in the numerical studies, one understands that the maximum is reached when $card(H(n)) = 2p' - 1$: $\sigma_B(J)^2 \leq 3L(J)$.

9.7.2 Study of Φ

In some cases, $\gamma_{1,p}$ defined in theorem 9 is too big. Indeed, in equation 9.2 in the proof of lemma 9.2.8, the increase $\Phi \leq (1/N) \sum_{n=1}^{N} \sum_{m \in H(n)} Ob(1)\epsilon_p$ is not fine enough.

Indeed, study

$$\Phi = (1/N) \sum_{n=1}^{N} \sum_{m \in H(n)} \left[E\{1_{Bo}(X_n)1_{Bo}(X_m)\} - E\{1_{Bo}(X'_n)1_{Bo}(X'_m)\} \right] .$$

Now we suppose that X_n is a sequence of bits. Therefore, $Bo_s = \{b\}$ where b=0 or b=1.

First, let us use again the equation 9.1 in the proof of lemma 9.2.6 : there exists a sequence i_s s=1,...,p', $p' < 2p$ and a sequence of Borel sets Bo'_s, s=1,...,p' such that

$$1_{Bo}(X_n)1_{Bo}(X_m) = 1_{Bo'_1}(X_n)1_{Bo'_2}(X_{n+i_2}).....1_{Bo'_{p'}}(X_{n+i_{p'}}) .$$

Therefore, $|E\{1_{Bo}(X_n)1_{Bo}(X_m)\} - E\{1_{Bo}(X'_n)1_{Bo}(X'_m)\}| \leq \epsilon_{p'}$.

To calculate the p', we should use results of the same type that those of section 9.7.1. Of course, for sequences of bits, the case which it should be studied is that where $Bo_s = \{b\}$.

First, let us interest in the case where ϕ is maximum.

Proposition 9.7.4 *Let $l_0 \in \mathbb{N}^*$. Suppose $Bo_s = J = \{b\}$, for s=1,2,...,p. Let $\zeta(p) = \frac{2p}{2^p} + 2\sum_{r=1}^{p-1} \frac{2(p+r)}{2^{p+r}}$. Suppose that $j_s = (s-1)l_0$ pour $s = 1, 2, ..., p$. Then*

$$\Phi \leq c(\epsilon)\epsilon\zeta(p) \ ,$$

where $c(\epsilon) > 1$ and $c(\epsilon) \approx 1$.

Proof One uses again the technique used in proposition 9.7.1.
If $m = n$, $1_{Bo}(X_n)1_{Bo}(X_m) = 1_J(X_n)1_J(X_{n+l_0})$.....$1_J(X_{n+(p-1)l_0})$. Then, p'=p.
If $m = n$+1, $1_{Bo}(X_n)1_{Bo}(X_m) = 1_J(X_n)1_J(X_{n+l_0})$.....$1_J(X_{n+pl_0})1_J(X_{n+(p+1-1)l_0})$. Then, p'=p+1. It is the same for m=n-1.

.....

If m=n+r, $1_{Bo}(X_n)1_{Bo}(X_m) = 1_J(X_n)1_J(X_{n+l_0})$.....$1_J(X_{n+(p+r-1)l_0})$. Then, p'=p+r. It is the same for m=n-r.

If m=n+p-1, $1_{Bo}(X_n)1_{Bo}(X_m) = 1_J(X_n)1_J(X_{n+l_0})$.....$1_J(X_{n+(2p-1-1)l_0})$. Then, p'=2p-1. It is the same for m=n-(p-1).

Then,

$$\Phi \leq \epsilon_p + 2[\epsilon_{p+1} + \epsilon_{p+2} + ... + \epsilon_{p+p-1}]$$
$$\approx \epsilon\Big[\frac{2p}{2^p} + 2\Big(\frac{2(p+1)}{2^{p+1}} + \frac{2(p+2)}{2^{p+2}} + ... + \frac{2(p+p-1)}{2^{p+p-1}}\Big)\Big]$$
$$\leq \epsilon\zeta(p) \ .\blacksquare$$

Lemma 9.7.1 *The following inequality holds*

$$\zeta(p) \leq \frac{10p-4}{2^p} \ .$$

Proof The following inequalities hold

$$\zeta(p) \leq \frac{2p}{2^p} + 2\Big(\frac{2(p+1)}{2^{p+1}} + \frac{2(p+2)}{2^{p+2}} + ... + \frac{2(p+p-1)}{2^{p+p-1}}\Big)$$
$$\leq \frac{2p}{2^p} + 4(2p-1)\Big(\frac{1}{2^{p+1}} + \frac{1}{2^{p+2}} + ... + \frac{1}{2^{p+p-1}}\Big)$$

$$\leq \frac{2p}{2^p} + \frac{4(2p-1)}{2^{p+1}}\Big(2\Big)$$
$$\leq \frac{2p}{2^p} + \frac{4(2p-1)}{2^p} \leq \frac{2p+8p-4}{2^p} \leq \frac{10p-4}{2^p} \ .\blacksquare$$

Now, let us interest in the case where Φ is minimal.

Proposition 9.7.5 *Let $p' \geq 2$. Suppose that $J_s = \{b\}$, for s=1,2,...,p. Suppose $card(H^*(n)) = p^2 - p$. In this case,*

$$\Phi \leq \frac{c(\epsilon)2(p^2-p)(2p-1))\epsilon}{2^{2p-1}} \ ,$$

where $c(\epsilon) > 1$ and $c(\epsilon) \approx 1$.

Proof One uses again the technique used in proposition 9.7.3. If $card(H^*(n)) = p'^2 - p'$, then $1_{Bo}(X_n)1_{Bo}(X_m) = 1_{Bo'_1}(X_n)1_{Bo'_2}(X_{n+i_2})\ldots\ldots 1_{Bo'_{p'}}(X_{n+i_{p'}})$, with p'=2p-1 : otherwise $card(H^*(n)) < p'^2 - p'$.

Therefore,

$$\Big| \sum_{m \in H^*(n)} \Big[E\big\{1_{Bo}(X_n)1_{Bo}(X_m)\big\} - E\big\{1_{Bo}(X'_n)1_{Bo}(X'_m)\big\}\Big]\Big| \leq (p^2-p)\epsilon_{2p-1} \ .$$

Therefore,

$$\Phi \leq c(\epsilon)(p^2-p)\frac{2(2p-1)\epsilon}{2^{2p-1}} \leq \frac{c(\epsilon)2(p^2-p)(2p-1)\epsilon}{2^{2p-1}} \ \blacksquare$$

Finally, in the numerical studies, one understands that the maximum is reached when $card(H(n)) = 2p-1 : \Phi \leq c(\epsilon)\epsilon\zeta(p)$.

Then, one can write the following proposition.

Lemma 9.7.2 *Let $\zeta'(p) = 2p + 2\sum_{r=1}^{p-1} \frac{2(p+r)}{2^r}$. We keep the notations of the proof of lemma 9.2.8 .*

Let

$$\gamma'_{1,p} = c(\epsilon)\frac{\zeta'(p)\epsilon}{2A(p)} + \frac{(p^2-p+1)(2q)\epsilon_{2p}}{2A(p)L(Bo)}$$
$$+\frac{(p^2-p+1)(1+2q)}{2A(p)L(Bo)}\big[2^{1-p}\epsilon_p + \epsilon_p^2\big] + \frac{(k_B/10)\epsilon_p}{2A(p)L(Bo)}\Big] \ .$$

Then,

$$\sigma_1^2 = \sigma_B^2\big[1 + Ob(1)2\gamma'_{1,p}\big] \ .$$

Proof In this case, with the notations of the proof of lemma 9.2.8 and if X_n is q-dependent,

$$\sigma_1^2$$

$$= \sigma_B^2 + \Phi + (p^2 - p + 1)(2q)Ob(1)\epsilon_{2p}$$
$$+(p^2 - p + 1)(1 + 2q)Ob(1)\big[2^{1-p}\epsilon_p + \epsilon_p^2\big] + (Ob(1)k_B/10)\epsilon_p$$

$$= \sigma_B^2\Big[1 + \frac{\Phi}{\sigma_B^2} + \frac{(p^2 - p + 1)(2q)Ob(1)\epsilon_{2p}}{\sigma_B^2}$$
$$+\frac{(p^2 - p + 1)(1 + 2q)}{\sigma_B^2}Ob(1)\big[2^{1-p}\epsilon_p + \epsilon_p^2\big] + \frac{(Ob(1)k_B/10)\epsilon_p}{\sigma_B^2}\Big]$$

$$= \sigma_B^2\Big[1 + \frac{\Phi}{A(p)L(Bo)} + \frac{(p^2 - p + 1)(2q)Ob(1)\epsilon_{2p}}{A(p)L(Bo)}$$
$$+\frac{(p^2 - p + 1)(1 + 2q)}{A(p)L(Bo)}Ob(1)\big[2^{1-p}\epsilon_p + \epsilon_p^2\big] + \frac{(Ob(1)k_B/10)\epsilon_p}{A(p)L(Bo)}\Big]$$

$$= \sigma_B^2\big[1 + Ob(1)2\gamma'_{1,p}\big] \ . \ \blacksquare$$

Then, we have the following approximation.

Lemma 9.7.3 *We assume that X_n is a sequence of q-dependent bits satisfying $\epsilon = \frac{\alpha}{\sqrt{q_0 N}}$. Then,*

$$\gamma'_{1,p} \approx \frac{\zeta'(p)\alpha}{2A(p)\sqrt{q_0 N}} + \frac{(p^2 - p + 1)\alpha}{2A(p)\sqrt{q_0 N}}\Big[(1 + 4q)\frac{4p}{2^p} + (1 + 2q)\frac{4p^2\epsilon}{2^p}\Big] \ .$$

Proof By our assumptions, $L(Bo) = 1/2^p$, $k_B = 0$ considering X_n is q-dependent. Then,

$$2\gamma'_{1,p} = \frac{c(\epsilon)\zeta'(p)\epsilon}{A(p)} + \frac{(p^2 - p + 1)(2q)\epsilon_{2p}}{A(p)L(Bo)}$$
$$+\frac{(p^2 - p + 1)(1 + 2q)}{A(p)L(Bo)}\big[2^{1-p}\epsilon_p + \epsilon_p^2\big] + \frac{(k_B/10)\epsilon_p}{A(p)L(Bo)}\Big]$$

$$\approx \frac{c(\epsilon)\zeta'(p)\epsilon}{A(p)} + \frac{(p^2 - p + 1)(2q)2(2p)\epsilon}{A(p)L(Bo)2^{2p}}$$

$$+\frac{(p^2-p+1)(1+2q)}{A(p)L(Bo)}\big[2^{1-p}\frac{2p\epsilon}{2^p}+\frac{4p^2\epsilon^2}{2^{2p}}\big]\Big]$$

$$=\frac{c(\epsilon)\zeta'(p)\epsilon}{A(p)}+\frac{(p^2-p+1)(2q)(4p)\epsilon}{A(p)2^p}$$
$$+\frac{(p^2-p+1)(1+2q)}{A(p)}\big[\frac{4p\epsilon}{2^p}+\frac{4p^2\epsilon^2}{2^p}\big]\Big]$$

$$=\frac{c(\epsilon)\zeta'(p)\epsilon}{A(p)}+\frac{(p^2-p+1)\epsilon}{A(p)}\Big[(1+4q)\frac{4p}{2^p}+(1+2q)\frac{4p^2\epsilon}{2^p}\Big]$$

$$\approx\frac{\zeta'(p)\alpha}{A(p)\sqrt{q_0N}}+\frac{(p^2-p+1)\alpha}{A(p)\sqrt{q_0N}}\Big[(1+4q)\frac{4p}{2^p}+(1+2q)\frac{4p^2\epsilon}{2^p}\Big]\ .\ \blacksquare$$

Chapter 10

Study of some files

10.1 Introduction

In this chapter, we study the data resulting from certain electronic files, especially from texts. By a study of these data based on logic, we will understand that one will be able to conclude that they behave like asymptotically independent sequences (and even Qd-dependent sequences).

We understood that, to build x_n, we use a sequence y_n which one can regard as a realization of a sequence of random variables.

In this chapter, we denote by y_n, a such sequence and by Y_n the associated sequence of random variables defined an a probability space (Ω, Δ, P) : there exists $\omega \in \Omega$ such that, for all n, $y_n = Y_n(\omega)$ for all n=1,...,N .

We do not impose that the Y_n are independent or identically distributed. But it is useful that the CLT is satisfied.

As a matter of fact, there are many data which are appropriate to obtain the sequences $b^1(n')$ built in chapter 11 . But we want to be sure that the previous hypotheses holds. That restricts the numbers of possible data.

10.2 Existence of satisfactory datas

10.2.1 Definition

At first, we had to know when a sequence y_n can be regarded as a realization of a sequence of really random variables : $y_n = Y_n(\omega)$ for all i=1,...,N .

First, any sequences of reals numbers can be regarded as a realization of a sequence of random variable of a certain type (completely deterministic, IID, etc) : this sequence of random variable is the model. But this model is correct with a some probability.

Then, to suppose "$y_n = Y_n(\omega)$" is a traditional scientific assumption if the y_n represents a physical phenomenon. One wants thus to show in an unquestionable way that it is also the case when y_n is resulting from certain electronic files

As a matter of fact a such sequence is simply a not-determinist sequence : that is to say, a sequence such that it is impossible to predicte fully y_{n+p}, when, one knows $y_1, y_2, ..., y_n$.

Now in order that the CLT holds, we impose that there is an asymptotic independence. Of course, a such sequence is non-determinist. Indeed, in this case, the sequence is not completely predictable.

10.2.2 Objections

But is what such sequences y_n exist? It is a physical question. It is almost a philosophical question. As a matter of fact, some people claimed that there does not exist finite random sequences : e.g. cf [1] page 167.

It is due partly so that any sample of a sequence of random variables can be regarded as fully determinist. Indeed the following proposition is obvious.

Proposition 10.2.1 *Let x_n, n=1,...,N, a sequence of real numbers. Then, there exists a function $g : \{1, 2, ..., N\} \to \mathbb{R}$ such that for all $n \in \mathbb{N}$, $x_n = g(n)$. Moreover, there exists p and a function $g : \mathbb{R}^p \to \mathbb{R}$ such that for all $n \in \{1, 2, ..., N - p\}$, $x_{n+p} = f(x_n, x_{n+1}, ..., x_{n-p+1})$.*

Moreover, some philosophies claim that all is given. For example, meteorology would be fully determined by all data of earth (all temperatures in all point of earth, all the atmospheric pressures, etc).

In the same way, actions of the men would be fully determined by the context in which they live and by the cells of their brains. Then, a book is fully determined before his writing by theses events.

But, in this case, the quantum theory is rejected. For that, one can call upon various reasons: 1) it is valid only for the infinitely small. 2) It is only a theory 3) It implies inadmissible contradictions for some people (Schrodinger cat).

But, all theses objections are false. In order to prove that, we use a counterexample : one can exhib a finite unpredictable sequence.

10.2.3 A finite random sequence

Let P(x) = $(x - x_1)(x - x_2)....(x - x_{2N})$ where $0 \leq x_1 < x_2 < .... < x_{2N} < 1$, $x_{j+1} - x_j \geq 1/4N$ for j=1,2,...,2N-1. Let $z_1, z_2, ..., z_N$ be a pseudo-random sequence with values in [0,1] obtainded by a good pseudo-random generator. Let $y_i = P(z_i)$ for i=1,2,....,N.

Then, it is no possible to predict y_{n+p} , $n \leq n+p \leq N$ if one knows only $y_1, y_2, ..., y_n$.

Indeed, even if one knew $z_1, z_2, ..., z_{n+p}$, it would not possible because any polynomial Q such that $deg(Q) = 2N$ and $y_n = Q(z_n)$ for n=1,2,...,p is a correct prediction of P. Then, all $y^*_{n+p} = Q(z_{n+p})$ is a correct prediction of y_{n+p}.

Now there exists an infinite number of possible polynomials Q.

Then, it is no possible to predict y_{n+p} even if one knew the sequence z_n and if one had an infinite computing power.

One can visualize that on the following example.

Example 10.2.1 *The following sample is considered*
$z^1_n = [-7.92, -2.70.2, 4.5, 8.99]$,
$y^1_n = Q(z_n) = [-0.2, -2.6, -0.45, -0.2, -0.3] * 10^7$.

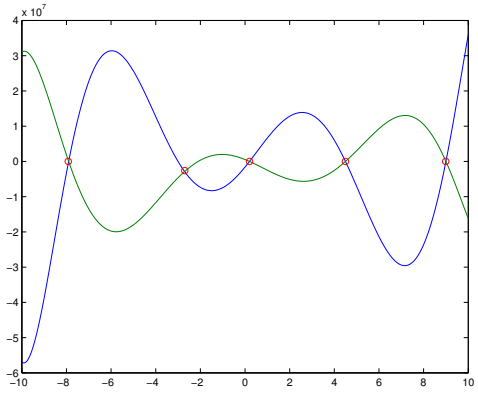

Figure 10.1: Marginal distribution near to the uniform distribution

Study The two polynomials P_1 and P_2 whose graphs appear in the figure 10.1 are two correct estimates for P. It verifies both $y^1_n = P_1(z^1_n)$ et $y^1_n = P_2(z^1_n)$. ■

Now there is no reasons that the $Y'_n s$ have the same distribution, ($y_n = Y_n(\omega)$). But is is not important because the philosophical objections are that the sequence is not independent.

Anyway, one can build a sequence y'_n where the $Y'_n s$ have the uniform distribution : one uses $y'_n = F^{-1}(y_n)$, where F is the distribution function of P(X) when X has the uniform distribution: $F^{-1}(P(X))$ has also the uniform distribution.

There is another reason that it no possible to predict y_{n+p} . In order to estimate P, it would be necessary to compute all the polynomial correlation coefficient of order smaller than 2N (cf [10]) .

It would thus be necessary to calculate the empirical orthogonal polynomials P_j^N of order J smaller than 2N associated with $z_1, z_2, ..., z_N$. However $P_j^N \equiv 0$ if $j > N$: the empirical polynomials of a order larger than the sample size are impossible to estimate.

Then, it is no possible to predict (even a little) y_{n+p} , $n \leq 2N$ if one knows $y_1, y_2, ..., y_n$. That is to say, y_{n+p} appears independent of the $y'_n s$.

Moreover, it is not surprising that y_n is unpredictable : indeed P depends on more parameters than N. As matter of fact, many simple function using more than N parameter z_n can be appropriate to obtain unpredictable sequence. For example if $y_i = Q(k_1, k_2, ..., k_N, k'_1, k'_2, ..., k'_n, z_i)$.

Indeed, in order to estimate the $k'_i s$ and the $k'_i{}' s$, one has to resolve the N equations : $y_i = Q(k_1, k_2, ..., k_N, k'_1, k'_2, ..., k'_n, z_i)$ for i=1,2,...,N, that is there are more parameters than equations.

Then, all sequence $y_1, y_2, ..., y_n$. which depend more parameters than N may be an unpredictable sequence.

10.2.4 Consequence 1

Then, the sequence y_n is random : a sequence whose it is impossible to predict the future, it is obligatorily random. It is even an independent sequence.

Then, the philosophy which affirms that there does not exists finite random sequences x_n , n=1,...,n, does not corresponds to reality : a sequence which one cannot predict is obligatorily random. To say the opposite is illogical.

10.2.5 Consequence 2

In order to obtain sequences which satisfy concretely some asymptotical independence assumptions, we shall use sequences which depend on a number of parameters much many larger than the size of sample.

it is always possible because the size of IID sequences which one can use on computer is quite lower than the number of parameters potentially usable : those are provided by the physical universe. This number is thus close to the infinity.

Thus to have finite random sequences, we use data which depend a priori on a very great number of parameters

10.3 Practical example

The sequence $b^1(n')$ which we have built in section 11.2 has been obtained by using text. in this case, the $y'_n s$, $n = 1, 2, ..., 10^7$ represent letters or punctuations which we write modulo $\kappa = 32$. Moreover we took these texts in various languages. In certain cases we have removed the introduction. A text on two

was rewritten in the opposite direction : $y_n^s = y'^s_{N_s-n}$ where N_s is the size of obtained sequence : cf section 11.2.4.

Concretely, one has used the following texts (because they are large) in various language : dictionary, Encyclopaedia, Bible, Theological sum, etc.

The dictionaries and the encyclopaedias are very good examples: the definitions which are consecutive in a dictionary generally represent independent facts : for example "decibel" is followed by "decide" in some dictionaries. The numbers which correspond to them are thus extracted from independent random sequences.

There are other books having equivalent properties. Thus, the book of history form quasi-independent sequence : it is difficult to predict the associated sequences of numbers because it is a question of predicting human behaviors.

In order to convert this files in number, one can use various programs. For example, in MATLAB 2005, one can use the function :

fid=fopen('Name of file');
fread(fid);

For example, in Matlab 2005,

[ABCDEFGHIJKLMNOPQRSTUVWXYZ abcdefghijklmnopqrstuvwxyz]
= [65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 10 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 10 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122]

In the same way, the key of the computer "return" = 10.

10.3.1 Use of text

In the majority of the sequences obtained from texts, it is reasonable to admit asymptotic independence: one can admit this assumption because of the following logical arguments.

1) The writing of a book depend of a very large number of parameters. To write a book is an human phenomenon. One can think that, in order to predict fully a book, it would at least be necessary to know the contents of cells of the brain of its writer and a great number of the events which he lived in his life. Normally, the number of parameters whose the content of the book depend will be always larger than the sample size of the example. One thus finds the argumentation introduced in section 10.2.3.

2) When they write a book many authors do not know what they will write exactly one page later. Concretely they would not predicte exactly what words he will use 100 words later. It will be even more difficult for letters. Then the dependence is weaker between more distant lines. That is, there is asymptotical independence.

3) It is more difficult to predicte the letters used for the people who are not the author of the book. For that, it would initially be necessary to know many things, for example its style, which is difficult with only one book.

4) Let us take the example of a novel. If the first pages of a book are known, it is sometimes possible to predict about what is written some pages after. But, a novel is left with the imagination of an author and sometimes these ones like to surprise. In fact if the beginning of a novel is known, there is a very great number of possible alternatives for the continuation of the history. Even for each alternative, there is a very great number of possible texts.

5) Not only, it quasi-impossible to predict about a text. But it is even more difficult to envisage the letters used.

6) To predict logically what is written in a book, it should initially be known that it is written in a certain language. It is not sure that one can arrive at this conclusion. Thus, one is unable to even currently decipher some languages. One could have deciphered the Egyptian hieroglyphs if the Rosetta Stone had not been written in several languages.

In addition, it has to be known that this text is written with an alphabet of 26 letters for example. If the same book is written in Chinese, one has an alphabet much more important. If this book were written in a rational written form, but not yet invented by men, it would be still other matter. Then, it is not at all certain that, even with means of infinite calculations, it is possible to know that the sequences of numbers obtained has a meaning as a text of English language, this more especially, as to build $b^1(n')$, we mix various texts and use also texts written with backward.

Then, in most of texts it is very clear that it is many more difficult to predict what words will used 200 words later than 100 words later. That is, there is asymptotical independence.

On the other hand, some texts have particular characteristics.

Let us take a dictionary or an encyclopaedia: in this case, there has no logic in this text, besides the first letters of the definite words (this characteristic disappears when the words are turned in numbers)

All these facts mean that logic implies that the files obtained starting from texts are asymptotically independent. One thus obtains a result concerning the first step of our method of construction of the random bits $b^1(n')$. That is logically surer than if one uses sequences supposed random provided by machines always prone to possible dysfunctions: if certain electronic files are used, there are certain assumptions because of logical reasoning.

10.3.2 Other data

One can use other datas in order to obtain the sequences of random numbers. Some give results less sure than text but easily usable.

One can also use software etablished on computer. Anyway, it is necessary to study by logical reasonning each type of files : programs, musics , etc..

For example, study mathematical texts

Of course, there is a certain logic in these texts.

But in order to predict what is written it would be necessary to know which theorems, among the multitude of possible theorems, the author discovered. That seems impossible especially if one takes care to remove the introduction.

Moreover it should be known which is the manner of writing of the author, if there are errors, which are true results and the best way of proving them. It also should be known if there are no errors and even if this mathematical text is not based on an error. These are informations that one cannot know a priori when one studies a file as a source of numbers.

Now study softwares. Then, much of subprogram can be used. Often, they can be regarded as independent from each other.

One cannot detail all that here more especially as the important thing is that in conclusion, the XORLT holds. However probably that arrives in much case since it does not require asymptotic independence : cf section 5.2.1.

Moreover, the number obtainded in chapter 11 satisfies all these tests of randomness.

10.3.3 Several files

One can use several files, for example, a dictionary and a software. Those are often completely independent from each other. The sequences of numbers which they provide are thus also independent.

It is the technique which we used to build the sequence $b^1(n')$ in section 11.2.

10.3.4 Conclusion

For this type of files, one can assume that y_n is a realization of a sequence of random variable Y_n : $y_n = Y_n(\omega)$ where $\omega \in \Omega$ and $Y_n \in \{0, 1/\kappa, 2/\kappa, ..., \kappa/\kappa\}$ where $\kappa = 32$. Moreover, there is asymptotical independence, and the CLT holds : often there is Qd-dependence.

10.4 Numerical Study

We study numerically various types of electronic files. If they are text's files, we study the data used in section 11.2 to build the sequences of random bits $b^1(n')$.

10.4.1 Independence of the $D(j)$'s

First, we study numerically the asymptotic independence of the $D(j) = D_S(j)$ defined in section 11.1.2 with S=10. In this section, for conveniences of notations, we pose also $D_j = D(j)$.

Numerically these data behave as if there were Qd-dependence, that is there exists Qd such that $(....D_{j-2}D_{j-1}, D_j)$ and $(D_{j+Qd}, D_{j+Qd+1}, D_{j+Qd+2}....)$ are independent : cf [21], page 369.

10.4.2 Texts

Test of the linear correlation coefficients

One will test the coefficient of linear correlation of $\{D_j\}$ and $\{D_{j+p_0}\}$, i.e. the correlation of the samples of size J , d_j and d_{j+p_0}, j=1,2,...,J .

Let $\rho^J(d_j, d_{j+p_0})$ be the empirical linear correlation coefficient associated to a sample (d_j, d_{j+p_0}).

One takes an interest in the table of $\sqrt{J}\rho^J(d_j, d_{j+p_0})$ for p_0 varying between 1 and 200. According to the following lemma, if the sample is independent, these values must have asymptotically the N(0,1) distribution.

Lemma 10.4.1 *Let X_n be an IID sequence of random variables with $E\{X_1\} = 0$. Let $\sigma^2(X)$ be the variance of X_1. Let $t \in \mathbb{N}^*$.*

Then $\frac{1}{\sqrt{n-t}} \sum_{s=1}^{n-t} \frac{X_s X_{s+t}}{\sigma^2(X)}$ has asymptotically the N(0,1) distribution.

Proof Let $Z_s = X_s X_{s+t}$. First $E\{Z_s\} = E\{X_s\}E\{X_{s+t}\} = 0$.

Moreover Z_s is (t+1)-dependent. Let $\sigma^2(Z) = E\{(Z_1 + ... + Z_{n-t})^2\}$. Then, $\sum_{s=1}^{n-t} \frac{Z_s}{\sigma(Z)}$ has asymptotically the N(0,1) distribution.

Now, $\sigma^2(Z) = \sum_{s=1}^{n-t} E\{Z_s^2\} + \sum_{s \neq s'} E\{Z_s Z_{s'}\} = \sum_{s=1}^{n-t} E\{Z_s^2\}$
$= \sum_{s=1}^{n-t} E\{X_s^2 X_{s+t}^2\} = (n-t)E\{X_1^2\}^2 = (n-t)\sigma^2(X)^2$. ■

For the texts, one has the following table meaning the "p_0" and $Cor_{p_0} = \sqrt{N}\rho^N(d_j, d_{j+p_0})$.

p_0	1	2	3	4	5	6	7	8	9
Cor_{p_0}	-9.71	-4.78	-0.29	-4.44	1.73	-0.50	-1.81	-0.46	-0.23
p_0	10	11	12	13	14	15	16	17	18
Cor_{p_0}	1.03	-1.01	0.87	-0.27	-0.56.	-0.21	1.45	0.29	-0.63
p_0	19	20	21	22	23	24	25	26	27
Cor_{p_0}	-0.93	-1.62	1.19	0.27	-0.61	-0.66	-1.02	-0.12	0.70

These results show that independence between D_j and D_{j+20} is a plausible assumption. They are confirmed by the following curve representing the empirical linear correlation coefficients for the texts. In this case, the correlation tends very quickly to 0 : cf figure 10.2.

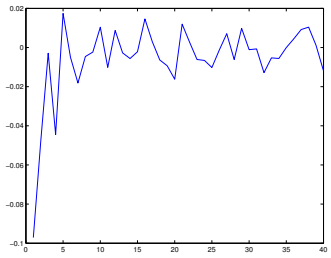

Figure 10.2: Correlation Coefficients :Text

Distance L^2

One will consider as measures of dependence between two random variables X and Y the distance L^2 to independence : for a partition $\{I_s\}$ of $[0,1]$ in N intervals with the same length, one sets

$$DL2(X,Y,n)$$

$$= \sum_{t=1}^{n}\sum_{s=1}^{n} P(X \in I_s)P(Y \in I_t)\big[P\{(X,Y) \in I_s \otimes I_t\} - P(X \in I_s)P(Y \in I_t)\big]^2.$$

One transforms the values of the two variables X and Y in [0,1] by translation and homothecy. Then, for the distance $DL2(D_j, D_{j+p_0}, 10)$, there is the graph of figure 10.3.

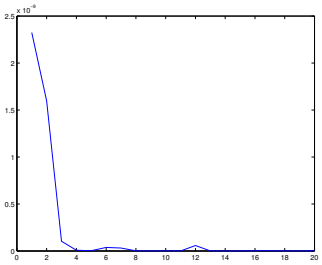

Figure 10.3: Distance L^2 : Texts

10.4.3 Mathematical text

One is interested now in the texts of mathematical papers or reports : one obtains also independence between D_j and D_{j+20} .

First, one studies the linear correlation coefficient. We use the samples (D_j, D_{j+5p_0}) , j=1,....,10000, with $p_0 = 1, 2, ..., 60$: One obtains the graph of

figure 10.4.

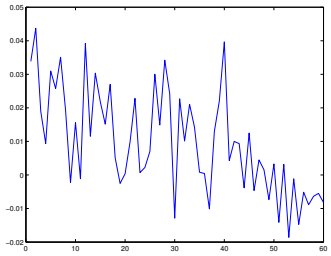

Figure 10.4: Correlation Coefficient : Mathematical text

With regard to distance DL2, one notices that only the first term is too large: for a partition (7,7), one obtains the following table representing different $10^{12} DL2(C_j, C_{j+p_0}, 7)$, j=1,....,10000.

0.2472	0.0013	0.0025	0.0024	0.0032	0.0008

10.4.4 Programs

One tests various sequences obtained thanks to some Matlab programs. One uses the chi-squared Test of independence : cf proposition A.1.1.

One use the statistics $\sqrt{2\chi_I^2} - \sqrt{2.d-1}$. It is known that it has roughly the distribution N(0,1) if there is independence : cf proposition A.1.2 .

Various values of these statistics are in the following table for the samples (d_j, d_{j+5p_0}), n=1 2,.....,10000, $p_0 = 1, 2, ...$

13.5736	8.0426	7.0507	5.4497	6.0559	4.9351	7.7882	0.4296
3.0867	2.1431	2.5712	3.9886	1.4108	1.6074	2.6225	0.1329
2.2573	1.6054	0.2588	0.9548	-0.1447	0.4380	1.4387	1.9938
-0.4956	1.3421	0.8168	0.2529	0.8284	-0.3554	0.5218	0.1666
0.1172	0.3091	-0.0522	-0.5690	-0.3577	1.6076	1.7870	1.0123

On this table, there is independence when $5p_0 \geq 18$.

The correlation coefficients are in figure 10.5 for the samples (D_j, D_{j+5p_0}) j=1 2,.....,10000.

We study distance DL2(X,Y,7), between D_j and D_{j+5p_0}, j=1 2,.....,10000. The graph of $10^{12} DL2(X, Y, 7)$ is in figure 10.6 .

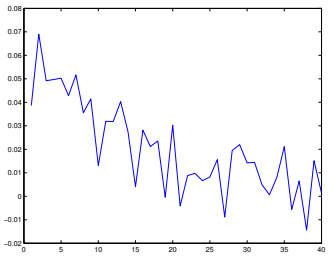

Figure 10.5: Correlation coefficients : Matlab programs

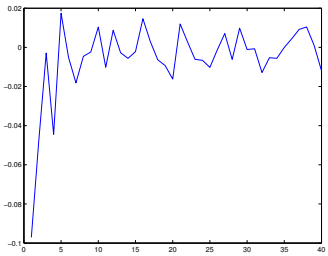

Figure 10.6: L^2 Distance : Matlab Programs

Remark 10.4.2 *One carried out much more tests than those used here. All give the same result : there is Qd-dependence.*

10.4.5 Multidimensional Tests

One has just carried out the tests of independence between D_j and D_{j+p_0}. But in order that there id Qd-dependence, one needs independence between $(D_j D_{j+1}, ..., D_{j+p_0})$ and $(D_{j+n_1+p_0}, D_{j+n_1+p_0+1}, ...D_{j+n_1+p_0+p_1})$ with $n_1 \geq Qd$.

In order to test this independence, simplest is to test independence between

$$U_j = a_0(D_j)^{b_0} + a_1(D_{j+1})^{b_1} + ... + a'_p(D_{j+p_0})^{b_p}$$

and

$$U'_{j+n_1+p_0} = a'_0(D_{j+n_1+p_0})^{b'_0} + a'_1(D_{j+n_1+p_0+1})^{b'_1} + ... + a_{p'}(D_{j+n_1+p_0+p_1})^{b'_{p'}}$$

for all the possibles sequences a_s, a'_s, b_s and b'_s.

Of course, there is a very great number of possible tests. We thus have carries out very many tests for various valeures of p_0, p_1, a_s, a'_s, b_s, and b'_s.

We tested the linear correlation coefficients by the chi-squared independence test and the L2 distance. The results have the same order that for the independence of X_n and X_{n+j} : there is Qd-dependence. Now, we give some results of these tests.

Linear correlation coefficient tests

The following table represents the values $Cor_{p_0} = \sqrt{J}\rho^J(U_j, U'_{j+n_1+p_0})$ for j=1,2,... Each line represents successively, mathematical texts, texts, programs. The values chosen are the following ones: $n_1 = 1 + i$, $a_s = a'_s = 1$, $b_s = b'_s = 1$ for $p_0 = p_1 = 5$, when i increases

15.03	12.56	5.54	-6.91	-2.52	2.46	-1.71	+0.52	-0.33
-12.23	10.35	-8.65	-7.06	-4.55	3.01	2.84	1.11	-0.98
11.66	10.08	-6.54	-5.77	-3.09	2.21	1.01	-0.87	1.12

In the case of independence, the asymptotic distribution is the N(0,1) distribution.

Lemma 10.4.3 *Let X_n be an IID sequence of random variables with $E\{X_1\} = 0$. Let $Y_s = X_s + X_{s+1} + ... + X_{s+p}$. Let $q > p$. Then,*

$$\frac{Y_1Y_{1+p+q} + Y_2Y_{2+p+q} + ...Y_nY_{n+p+q}}{\sqrt{n}E\{X_1^2\}\sqrt{(p+1)^2 + 2\sum_{t=1}^{p} t^2}}$$

has asymptotically the N(0,1) distribution.

Proof First, $Z_s = Y_sY_{s+p+q}$ is Qd-dependent.
Moreover, $E\{Z_s\} = E\{Y_s\}E\{Y_{s+p+q}\} = 0$.
On the other hand

$$E\{Z_s^2\}$$
$$= E\{(X_s + X_{s+1} + ... + X_{s+p})^2(X_{s+p+q} + X_{s+p+q+1} + ... + X_{s+p+q+p})^2\}$$
$$= E\{(X_s + X_{s+1} + ... + X_{s+p})^2\}^2 = E\{Y_s^2\}^2 = (p+1)^2E\{X_s^2\}^2 .$$

Moreover

$$E\{Z_sZ_{s+t}\}$$
$$= E\{(X_s + ... + X_{s+p})(X_{s+p+q} + ... + X_{s+2p+q})..... .....(X_{s+t} + ... + X_{s+t+p})(X_{s+t+p+q} + ... + X_{s+t+2p+q})\} .$$

Therefore, $E\{Z_s Z_{s+t}\} = 0$ si $p < t$.

If $t \leq p$,

$$E\{Z_s Z_{s+t}\}$$

$$= E\{(X_s + ... + X_{s+p})(X_{s+t} + ... + X_{s+t+p}).....$$

$$.....(X_{s+p+q} + ... + X_{s+2p+q})(X_{s+t+p+q} + ... + X_{s+t+2p+q})\}$$

$$= E\{(X_s + ... + X_{s+p})(X_{s+t} + ... + X_{s+t+p})\}.....$$

$$.....E\{(X_{s+p+q} + ... + X_{s+2p+q})(X_{s+t+p+q} + ... + X_{s+t+2p+q})\}$$

$$= E\{(X_{s+t} + ... + X_{s+p})^2\}E\{(X_{s+t+p+q} + ... + X_{s+2p+q})^2\}$$
$$= (p - t + 1)^2 E\{X_{s+1}^2\}^2.$$

Therefore,

$$E\{(Z_1 + Z_2 + ... + Z_n)^2\} = \sum_{s=1}^{n} E\{Z_s^2\} + \sum_{s=1}^{n} \sum_{s' \neq s} E\{Z_s Z_{s'}\}$$

$$= \sum_{s=1}^{n} E\{Z_s^2\} + \sum_{s=p+1}^{n-p} \sum_{s \neq s'} E\{Z_s Z_{s'}\} + O(1)$$

$$= \sum_{s=1}^{n} E\{Z_s^2\} + 2 \sum_{s=p+1}^{n-p} \sum_{s < s'} E\{Z_s Z_{s'}\} + O(1)$$

$$= \sum_{s=1}^{n} E\{Z_s^2\} + 2 \sum_{s=p+1}^{n-p} \sum_{t=1}^{p} E\{Z_s Z_{s+t}\} + O(1)$$

$$= nE\{X_1^2\}^2(p+1)^2 + 2 \sum_{s=p+1}^{n-p} E\{X_1^2\}^2 \sum_{t=1}^{p} (p - t + 1)^2 \ + O(1)$$

$$= nE\{X_1^2\}^2(p+1)^2 + 2 \sum_{s=p+1}^{n-p} E\{X_1^2\}^2 \sum_{t=1}^{p} t^2 + O(1)$$

$$= nE\{X_1^2\}^2(p+1)^2 + 2(n-2p)E\{X_1^2\}^2 \sum_{t=1}^{p} t^2 + O(1)$$

$$= nE\{X_1^2\}^2\Big[(p+1)^2 + 2\sum_{t=1}^{p} t^2 + O(1)/n\Big] .$$

Then, $\frac{Z_n+....+Z_n}{\sqrt{n}E\{X_1^2\}\sqrt{(p+1)^2+2\sum_{t=1}^{p} t^2+O(1)/n}}$ has asymptotically the N(0,1) distribution .

Then, $\frac{Z_n+....+Z_n}{\sqrt{n}E\{X_1^2\}\sqrt{(p+1)^2+2\sum_{t=1}^{p} t^2}}$ has asymptotically the N(0,1) distribution. ■

Chi-squared independence test

One carries out the chi-squared independence test with a partition in (10,10) intervals: i.e one uses the statistics $\sqrt{2\chi_I^2} - \sqrt{2.9^2 - 1}$ where $\chi_I^2 = \chi_{X,Y}^2 - \chi_X^2 - \chi_Y^2$. We know that, under the hypothesis of independence, $\sqrt{2\chi_I^2} - \sqrt{2.9^2 - 1}$ has about the N(0,1) distribution : cf propositions A.1.1 and A.1.2 .

One carries out the test between U_n and $U'_{n+n_1+p_0}$ with the following parameters : $p_1 = 25$ $n_1 = 1+j$, $a_s = a'_s = 0$ except $a_{5s} = a'_{5s} = 1$ for s=1,2,3,4,5, $b_5 = b'_5 = 2$, $b_{10} = b'_{10} = 4$, $b_{15} = b'_{15} = 1$, $b_{20} = b'_{20} = 3$, $b_{25} = b'_{25} = 4$, when j increases.

For the text's files, we obtain

15.03	12.56	5.54	-6.91	-2.52	2.46	-1.71	+0.52	-0.33

For the files of mathematical texts, we obtain

-12.23	10.35	-8.65	-7.06	-4.55	3.01	2.84	1.11	-0.98

For the Matlab programs, we obtain

11.66	10.08	-6.54	-5.77	-3.09	2.21	1.01	-0.87	1.12

All these results - as all those which we obtained in other tests - thus conclude that there is Qd-dependence.

10.4.6 Study of the dependence between different files

When files of the different type are used - for example, a text file and a software - these files are independent in the usual meaning of the English language. It is the case also for files of comparable nature but of different sources, for example, two different books.

It is logical to think that they are also independent in the statistical meaning. All the tests which we have carry out confirm this result.

Indeed, one obtains results as those which we have just described, for example in section 10.4.5 for the three types of used tests.

For example, let us carry out the chi-squared test of independence which uses $\sqrt{2\chi_I^2} - \sqrt{2.9^2 - 1}$ where $\chi_I^2 = \chi_{X,Y}^2 - \chi_X^2 - \chi_Y^2$ and which are associated with a partition in (10,10) squares.

The following table gives the results for independence between two texts of mathematical reports relating, one to the CLT, the other on cryptography.

Let X_n and Y_n be the random variables associated with the CLT and the cryptography, respectively. Then, one tests independence between the variables

$$2X_n^3 - 4.5X_{n+1}^2 - 0.5X_{n+2}^2 - 2X_{n+3}^3 + 7X_{n+4} - 3.2X_{n+5}^4 - X_{n+6}^2 + 8X_{n+7}^5$$

and

$$2Y_{n+h}^3 - 4.5Y_{n+1+h}^2 - 0.5Y_{n+2+h}^2 - 2Y_{n+3+h}^3 + 7Y_{n+4+h} - 3.2Y_{n+5+h}^4 - Y_{n+6+h}^2 + 8Y_{n+7+h}^5$$

for different h.

Then, the following results have been obtained for h=0,10,20.....

-1.12	0.80	-0.45	-2.01	1.23	0.78	-1.37	0.61	0.98

Problems for some files

It is necessary to pay attention during the conversion of the files in numbers. Indeed, for some systems, it can have certain problems there.

Thus if one uses files text AppleWork, there is much more zeros than the text lets appear normally. Thus for a file of $Y \in \{0, 1, ..., 31\}$, of size 10623940, there were 2466752 zeros, which is of course too much. In this case, one can find that there is even dependence between files of different books.

It is thus necessary to study correctly the behavior of the files translated in number before using them in the construction of the random numbers.

10.4.7 Uniformity

In the previous sections, we saw that there is asymptotic independence. To apply the theorem 4, it remains to prove H_S, $H_{ms}(4)$ and that $E\{(S_u)^2\} - E\{(S'_u)^2\} \rightarrow 0$.

In fact, there is a result much more strong : one can conclude to the uniformity of distributions.

Indeed, for the sequences D_n, it seems already reasonable to admit that there is an approximation of the stationarity which check H_S. It is confirmed by all the tests.

But in fact, it is even truer for sequences $e^2(j)$ considering one adds modulo m^1 a sequence of pseudo-random numbers $rand_0(j)$ whose uniformity was abundantly tested : $e^2_S(j) = \overline{e^1_S(j) + rand_0(j)}$: cf section 11.1.2. Let us recall that it is not the uniformity of the pseudo-random numbers which generally pose problem, but the number of independence to be checked : cf section 2.1.5.

Reasonably, it is thus logical to think that one has the uniformity of the distributions. Besides, it is what all the tests which one could carry out confirm absolutely. In fact what one tests, it is again the uniformity of the used random pseudo-generators. It is not surprising that one concludes to the uniformity.

There remains a condition so that theorem holds 4 holds : $E\{\xi^2_u\} \rightarrow 0$. So that happens, it is enough to choose suitably sequence $\kappa(n)$ so that $\kappa(n) \rightarrow \infty$. Thus all the assumptions so that the theorem 4 holds are checked.

.

10.4.8 Conclusion : CLT

The previous numerical results have all the same conclusion: there are Qd-dependence in all the texts, mathematical texts and softwares which we tested.

Let us notice also that we developed in this section numerical, logical and philosophical arguments. These arguments are perfectly reasonable.

One can reasonably conclude from all these facts that there is asymptotic independence. With regard to theorem 4, all the assumptions are checked. We deduce that one is in the case of the CLT.

But, the fact which interests us, it is that the curve of the probabilities have the shape of bell. The fact that one can apply the CLT and numerical studies carried out in the chapter 5.1 (in particular in section 5.3.3) show that this result is probably checked for all the files that we studied here.

10.4.9 XORLT

Since one can apply the CLT, one can thus also apply the XORLT which is weaker (cf proposition 5.2.3). But, one does not know yet all the assumptions under which it is satisfied. To have a more definite idea, it is necessary to study the H_n's used in the section 11.1.2.

Indeed, if there is Qd-dependence on the D_n, there is independence between $.....H_{n-2}, H_{n-1}, H_n$ and $.....H_{n+Qd}, H_{n+Qd+1}, H_{n+Qd+2}, ...$: cf also proposition 5.2.4 .

As a matter of fact, the H_n behave even as being independent. That is confirmed by the following numerical studies.

Independence between the H_n's : numerical results

One carried out many tests on the independence of the H_n's. In particular, one studied the dependences between $\sum_{s=0}^{P} a_s (H_{n+s})^{b_s}$ and $\sum_{t=0}^{Q} a'_s (H_{n+p_0+t+j})^{b'_s}$.

The a_s ,a'_s , b_s, b'_s have been chosen randomly : $a = \lfloor 10 * rand(1, P) \rfloor$, $b = \lfloor 10 * rand(1, Q) \rfloor$, where rand is a Matlab pseudo-random generator.

For example, we have used the chi-squared independence test with P=Q and various sequences a_s ,a'_s , b_s, b'_s. Then the following results have been obtained for the statistics $\sqrt{\chi^2_I} - \sqrt{2d-1}$ (cf proposition A.1.2) :

P=1	-0.7019	2.9104	0.952
P=2	-2.4783	2.010	1.0245
P=3	-0.9502	0.854	-1.254
P=10	-0.2449	0.548	-1.987

Many of other tests were carried out: one can conclude to the independence of the H_n's.

One also has tested the uniformity. For the same reasons that for the CLT, one concludes that uniformity holds.

Conclusion : XORLT

All the previous results confirm it to us: with our data, one can apply the CLT and the XORLT. With all the arguments which we developed one can to conclude that reasonably, there are all the chances that one can regard the H_n as Qd-dependent. In fact, they are even probably independent.

Chapter 11

Building of a random sequence

11.1 General method

11.1.1 Choice of data

Notations of data

It is supposed that one has a sequence of data a(j) translated in number: a(j) , $j = 1, 2, ..., N_3$, $a(j) \in \{0, 1, ..., Ka - 1\}$. One will transform it into a sequence of random bits $b^0(n')$.

One supposes that Ka is small enough : for example, $Ka \leq 1000$. If it is not the case, one can break up the a(j)'s in order to have Ka small enough.

It is supposed that a(j) can be regarded as a sample of a sequence of random variables A(j) defined over a probability space (Ω, Δ, P) : $a(j) = A(j)(\omega)$ where $\omega \in \Omega$.

We write with CAPITAL letters the sequences of random variables defined over (Ω, Δ, P) and with small letters the realization of these sequences (cf Notations 1.2.1). For example, $c(j) = a(j) - \kappa \lfloor a(j)/\kappa \rfloor$ and $C(j) = A(j) - \kappa \lfloor A(j)/\kappa \rfloor$: cf below.

Study of data

It is supposed that there is some asymptotic independence : cf sections 11.1.1 and 11.2.9. If these conditions are not satisfied, it is maybe still possible to apply the method described here. But that depends on the properties of the XORLT.

One checks this asymptotic independence by logical and numerical studies. For example, one studied data according to the method of the chapter 10. In

particular, it is necessary that, for all Borel set Bo,

$$\Big| E\big\{ \big(1_{Bo}[A(j)] - L_j\big)\big(1_{Bo}[A(j+d)] - L_{j+d}\big)\big\}\Big| \leq \phi(d) \ ,$$

where $\sum_d \phi(d) < \infty$ with $L_j = E\big\{1_{Bo}[A(j)]\big\}$.

Of course, because the sequence a(j) is finite, one has always such a result. But, which one wants, it is to be able to consider that there is well convergence of $\sum_d \phi(d)$. One will be able to admit this convergence if one finds for example, $\phi(d) < C_0/d^{1+t}$, where $t > 0$ and where $C_0 \in \mathbb{R}_+$ is a suitably chosen constant.

One can also check asymptotic independence with the sequences transformed of a(j), for example c(j), d(j) or $e^2(j)$: cf definitions below.

11.1.2 Description of the method

Shortening of the a(j)'s

Let $\kappa \in \mathbb{N}^*$. We set $c(j) = a(j) - \kappa\lfloor a(j)/\kappa \rfloor$.
Comment 1 $c(j) = a(j) \mod \kappa$.
Comment 2 One chooses κ in order to obtain a sequence c(j) such as, for all $t \in \{0, 1, ., \kappa - 1\}$, $P'_e\{C(j) = t\} > 0$ where P'_e is the empirical probability associated with $c(1), c(2), ..., c(N_3)$.

Choice of the parameters

a) We choose $\alpha \in \mathbb{R}_+$ such that $\alpha \leq 0.02$.
Comment We choose α (and therefore ϵ : cf section 11.1.4) according to the quality of the desired approximation: in fact, one chooses α according to $\beta_{1,p}$ and $\beta_{2,p}$: cf section 11.1.4.

b) One choose first S=10.

c) One chooses now q_0 and $r_0 \in \mathbb{N}^*$. For that, we use the following notation.

Notations 11.1.1 *Let fi_n be the Fibonacci sequence $\{fi_n\}$ (cf definition 1.3.2). For all $x \geq 2$, we set $m^F(x) = fi_{n_0-1}$ where $fi_{n_0-1} \leq x < fi_{n_0}$.*

One chooses q_0 and $r_0 \in \mathbb{N}^*$ such that :
a) q_0/r_0 is maximum
b) $m_S = m^F([m^F(\kappa^{r_0})]^{3/4})$ is sufficiently large but not too (cf section 8.5.1 and remark 11.1.7)
c) They satisfy the conditions

$$m_S/2^{q_0} \geq 1001 \ ,$$

$$\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2^S) \leq \frac{2\alpha\sqrt{S}}{\sqrt{N_3}}\sqrt{r_0 m_S}\ ,$$

where $a_2^S = \Gamma\Big(\Gamma^{-1}(4^{-q_0})\sqrt{\lfloor m_S/2^{q_0}\rfloor/(m_S/2^{q_0}) + 2^{q_0}/m_S}\Big)$ and where Γ is defined in notation 1.3.6 .

d) One checks that the following assumption holds.

Hypothesis 11.1.1 *The following approximation holds :* $a_2^S \approx 1/4^{q_0}$. *In particular,* $0.9/4^{q_0} \leq 4^{-q_0} \leq a_2^S \leq 1.1/4^{q_0}$.

That amounts supposing that the event $|\epsilon_{I_k}| \leq \frac{a_2\sqrt{N_{Iel}}}{m_S}$ has a probability about $1.1/4^{q_0}$ to be carried out whereas one has at the most a sample of 2^{q_0} possible intervals $I_k = [k/2^{q_0}, (k+1)/2^{q_0}[$: cf section 7.2; cf also section 11.1.3 and lemma 11.1.4 .

To make uniform the marginal distribution

a) We set $d_S(j) = \sum_{r=1}^{r_0} c(r_0(j-1)+r)\kappa^{r-1}$ for $j = 1, 2, ..., \lfloor N_3/r_0\rfloor$.
Comment Then, $d_S(j) \in \{0, 1, ..., \kappa^{r_0} - 1\}$: $d_S(j)$ is a number basis κ written with r_0 digits and obtained by joining the c(j) . For example, the writting of $d_S(1)$ is $d_S(1) = \overline{\overline{c(r_0)c(r_0-1).....c(2)c(1)}}$.

b) We set $e_S^1(j) = \Big\lfloor d_S(j)[m_S^1/\kappa^{r_0}]\Big\rfloor$ for $j = 1, 2, ..., \lfloor N_3/r_0\rfloor$ where $m_S^1 = m^F(\kappa^{r_0})$.
Comment The $d_S(j) \in \{0, 1, ..., \kappa^{r_0} - 1\}$ are transformed in integers $e_S^1(j) \in \{0, 1, ..., m_S^1 - 1\}$.

c) We set $e_S^2(j) = \overline{e_S^1(j) + rand_0(j)}$ mod m_S^1 for $j = 1, ..., \lfloor N_3/r_0\rfloor$ where $rand_0(j)$ is a pseudo-random generator with values in $F^*(m_S^1)$ and period m_S^1 or $k_4.m_S^1$, $k_4 \in \mathbb{N}^*$.
Comment : We want that, for all $k \in F^*(m_S^1)$, it is logical to admit $P\{e_S^2(j) = k\} > 0$. We want that all $k \in F^*(m_S^1)$ have a reasonable probability to be realized.

d) For $j = 1, 2, ..., \lfloor N_3/r_0\rfloor$, we set $e_S^3(j) = m_S T_1^{m_S}(e_S^2(j)/m_S^1)$ where $T_1^{m_S}$ is the Fibonacci function modulo m_S^1 : cf definition 1.3.5.
Comment 1 We remark that $e_S^3(j) \in F^*(m_S)$.
Comment 2 One makes independent the $e_S^2(j)$'s and one makes uniform the marginal distributions of the $e_S^2(j)$'s.

Remark 11.1.1 *In some cases, one can remove the transformation* $T_1^{m_S}$ *defined in the step d) for the last numbers. For example, one can set* $e_S^3(j) =$

$\left\lfloor e_S^2(j) m_S / m_S^1 \right\rfloor$ *for the last* $j \in \{n_d, n_d+1, ..., \lfloor N_3/r_0 \rfloor\}$, *where* n_d *is about* $(S-1)\lfloor N_3/[Sr_0] \rfloor$. *This is carried out with the aim described in remark 11.1.2.*

To avoid losing data, one can make differently: at first, one chooses $r_0' < r_0$. *One applies the step a)* ($d_S(j) = \sum_{r=1}^{r_0'} c(r_0'(j-1)+r)\kappa^{r-1}$) *to the* $n_{d'}$ *last c(j) where* $n_{d'}$ *is selected so that the* $n_{d'}$ *last numbers obtained thus (that one notes* $d_S'(J)$) *corresponds to a last line* $f(S,n) \in F^*(m_S)$ *defined in the following subsubsection : cf remark 11.1.2 .*

In this case one can replace m_S^1 *by* m_S *for the constructions defined in the steps b) and c).*

Use of the limit theorems

a) We denote by $e_S^4(t)$, $t = 1, 2, ..., N_2$, $N_2 = NS \leq \lfloor N_3/r_0 \rfloor$, a subsequence of $e_S^3(j)$ obtained by suppressing some sequences $e_S^3(\rho_u), e_S^3(\rho_u + s_{u_1})$, ...,$e_S^3(\rho_u + s_{u_n})$.

Comment One removes possibly some $e_s^3(j)$ in order to ensure independence between the lines defined below. If one does not have independent files, this step is not necessary forcing : $N_3 = NS$ should just be imposed.

b) We set $f_S(i,n) = e_S^4(n + N(i-1))$ for i=1,...,S , n = 1,...,N.

Comment One transforms the matrix line $\{e_S^4(j)\}$ into a SxN matrix.

c) We set $g_S(n) = \sum_{i=1}^{S} f_S(i,n)$ for n = 1,...,N.

Comment One uses the summation of the central limit by summoning the columns.

d) We set $h_S(n) = \overline{g_S(n)} \bmod m_S$ for n = 1,...,N.

Comment This corresponds to use the XORLT.

Remark 11.1.2 *If one applies the changes of construction defined in remark 11.1.1, one chooses the parameters defined in this remark so that the last line f(S,n) is made of numbers not having been transformed by* $T_1^{m_S}$ *described with the step d) of the previous section.*

One uses the changes of construction defined in remark 11.1.1 so that the probabilities associated with these sums can be regarded as selected really randomly. In this case, one can apply the proposition 5.5.1 . One thus avoids for example the difficulties defined in subsubsection "Problem in some case", section 5.5.8.

That is not necessary in all the cases : for example if we use the sequences c(j) resulting from files enough different as in section 11.2.

Checking of S

a) One checks by numerical calculations that the curve of the

$$h \mapsto P\{H_S(n) = h \mid H_S(n+j_2) = h_2, ..., H_S(n+j_p) = h_p\}$$

is enough close to that of the uniformity: it is necessary that the condition of the section 7.2 is satisfied. In general, it is well the case if S= 10.

If it is not the case, one remakes several times the previous operations with various $S > 10$.

One chooses S according to the obtained results by checking that, for the chosen S, the curve of

$$g \mapsto P\{G_S(n) = g \mid G_S(n+j_2) = g_2, ..., G_S(n+j_p) = g_p\}$$

is sufficiently smooth and especially that the curve of the

$$h \mapsto P\{H_S(n) = h \mid H_S(n+j_2) = h_2, ..., H_S(n+j_p) = h_p\}$$

is sufficiently near to that of the uniform distribution.

b) One chooses smallest $S \geq 10$ which is appropriate. It is noted S_0.

c) We set $h(n) = h_{S_0}(n)$ for n = 1,...,N.

Use of the Fibonacci Congruence

For n = 1,...,N, we set $k(n) = \overline{ah(n)}$ modulo m
where $m = m_{S_0}$ and where a is the largest element of the Fibonacci sequence fi_n such as $a = fi_{n_3} < m = fi_{n_3+1}$.
Comment : $k(n) = \overline{T}(h(n))$ where T is the Fibonacci congruence with parameters (a,m).

Getting the random sequences

a) We set $r(n) = k(n)/m$ for n=1,...,N.

b) Let $r(n) = \overline{\overline{0, b_1^n, b_2^n .....}}$, $b_s^n \in \{0, 1\}$, the binary writting of r(n).

c) We set $x(n) = \overline{\overline{0, b_1^n, b_2^n .....b_{q_0}^n}}$ where q_0 was defined previously in 11.1.2.
Comment In fact, in the previous steps, one applied the Fibonacci function T_{q_0} : $x(n) = T_{q_0}\big(h(n)/m\big)$ (cf definition 1.3.5).

d) We set $b'_{q_0 n-r+1} = b_r^n$ for n=1,...,N and $r = 1, ..., q_0$.
Comment One obtains the b'_n by taking the b_t^n's successively : $b'_1 = b^1_{q_0}$, $b'_2 = b^1_{q_0-1}$,....., $b'_{q_0} = b^1_1$, $b'_{q_0+1} = b^2_{q_0}$, $b'_{q_0+2} = b^2_{q_0-1}$,....., $b'_{2q_0} = b^2_1$,

$b'_{2q_0+1} = b^3_{q_0}$,.....

e) The sequence $\{b'_n\}$ is noted $b^0(n')$, $n' = 1, 2, ..., Nq_0$.

11.1.3 Explanation of the conditions about q_0 and r_0

Assumptions

Because the various steps of this construction, one can accept the model of the section 7.2 :

$$P\{\overline{\sum_i F(i,n)} = k \mid \overline{\sum_i F(i,n+j_s)} = h_s, s = 2, 3...\} = \frac{1}{m}\Big[1 + u_k\Big]$$

(cf also section 11.3.1). By property 7.2.2, we admit the following assumption.

Hypothesis 11.1.2 *For all $k/2^{q_0}$, for all finite injective sequence j_s, we assume that*

$$P\big\{X(n) = k/2^{q_0} \;\big|\; H(n+j_2) = h_2, ..., H(n+j_p) = h_p\big\} = 1/2^{q_0} + Ob(1)\epsilon_{I_k} \ ,$$

where $\mid \epsilon_{I_k} \mid \leq \epsilon = \frac{\Gamma^{-1}(4^{-q_0})\sqrt{N_{Iel}}}{m}$.

Some lemmas

First, the following lemma is needed.

Lemma 11.1.3 *For all $k/2^{q_0}$, $k \in \mathbb{N}$, for all finite injective sequence j_s ,*

$$P\big\{X(n) = k/2^{q_0}\big|X(n+j_2) = x_2, X(n+j_3) = x_3, ...\big\} = 1/2^{q_0} + Ob(1)\epsilon_{I_k} \ .$$

Proof Define J_s by $\{H(n+j_s) \in J_s\} = \big\{X(n+j_s) = x_s\big\}$. By proposition 4.2.3,

$$P\big\{X(n) = k/2^{q_0}\big|X(n+j_2) = x_2, X(n+j_3) = x_3, ...\big\}$$

$$= \frac{P\Big\{\big\{X(n) = k/2^{q_0}\big\} \cap \big\{X(n+j_2) = x_2\big\} \cap \big\{X(n+j_3) = x_3\big\} \cap ....\Big\}}{P\Big\{\big\{X(n+j_2) = x_2\big\} \cap \big\{X(n+j_3) = x_3\big\} \cap ....\Big\}}$$

$$= \frac{P\Big\{\big\{X(n) = k/2^{q_0}\big\} \cap \big\{H(n+j_2) \in J_2\big\} \cap \big\{H(n+j_3) \in J_3\big\} \cap ....\Big\}}{P\Big\{\big\{H(n+j_2) \in J_2\big\} \cap \big\{H(n+j_3) \in J_3\big\} \cap ....\Big\}}$$

$$= \sum_{h_2,h_3,...} \frac{P\Big\{\big\{X(n) = k/2^{q_0}\big\} \cap \big\{H(n+j_2) = h_2\big\} \cap \big\{H(n+j_3) = h_3\big\} \cap ....\Big\}}{\sum_{h_2,h_3,...} P\Big\{\big\{H(n+j_2) = h_2\big\} \cap \big\{H(n+j_3 = h_3\big\} \cap ....\Big\}}$$

$$= \sum_{h_2,h_3,...} \eta_{h_2,h_3,...} \frac{P\Big\{\Big\{X(n) = k/2^{q_0}\Big\} \cap \Big\{H(n+j_2) = h_2\Big\} \cap \Big\{H(n+j_3) = h_3\Big\} \cap ....\Big\}}{P\Big\{\Big\{H(n+j_2) = h_2\Big\} \cap \Big\{H(n+j_3 = h_3\Big\} \cap ....\Big\}}$$

where

$$\eta_{h_2,h_3,...} = \frac{P\Big\{\big\{H(n+j_2) = h_2\big\} \cap \big\{H(n+j_3 = h_3\big\} \cap ....\Big\}}{\sum_{h_2,h_3,...} P\Big\{\big\{H(n+j_2) = h_2\big\} \cap \big\{H(n+j_3 = h_3\big\} \cap ....\Big\}} \ .$$

Then, $\sum_{h_2,h_3,...} \eta_{h_2,h_3,...} = 1$.

Therefore,

$$P\big\{X(n) = k/2^{q_0} \big| X(n+j_2) = x_2, X(n+j_3) = x_3, ...\big\}$$

$$= \sum_{h_2,h_3,...} \eta_{h_2,h_3,...} \frac{P\Big\{\Big\{X(n) = k/2^{q_0}\Big\} \cap \Big\{H(n+j_2) = h_2\Big\} \cap \Big\{H(n+j_3) = h_3\Big\} \cap ....\Big\}}{P\Big\{\Big\{H(n+j_2) = h_2\Big\} \cap \Big\{H(n+j_3 = h_3\Big\} \cap ....\Big\}}$$

$$= \sum_{h_2,h_3,...} \eta_{h_2,h_3,...} P\Big\{X(n) = k/2^{q_0} \Big| H(n+j_2) = h_2, H(n+j_3) = h_3, ...\Big\}$$

$$= \sum_{h_2,h_3,...} \eta_{h_2,h_3,...} \Big[1/2^{q_0} + Ob(1)\epsilon_{I_k}\Big]$$

$$= 1/2^{q_0} + Ob(1)\epsilon_{I_k} \ ,$$

because $\sum_{h_2,h_3,...} \eta_{h_2,h_3,...} = 1$ and $0 \leq \eta_{h_2,h_3,...}$. ■

One needs the following proposition.

Lemma 11.1.4 *We set $a_2^{S_0} = a_2$. We suppose that hypothesis 11.1.2 holds. We suppose $m/2^{q_0} \geq 1001$.*

Then, for all Borel set $Bo \subset \{0, 1/2^{q_0}, 2/2^{q_0}, ..., (2^{q_0}-1)/2^{q_0}\}$, $|\epsilon_{Bo}| \leq \frac{2^{q_0/2}\Gamma^{-1}(a_2)}{2\sqrt{m}}$.

Proof We know that $a_2 = \Gamma\Big(\Gamma^{-1}(4^{q_0})\sqrt{\lfloor m/2^{q_0}\rfloor/(m/2^{q_0}) + 2^{q_0}/m}\Big)$.

Then, $\Gamma^{-1}(a_2) = \Gamma^{-1}(4^{q_0})\sqrt{\lfloor m/2^{q_0}\rfloor/(m/2^{q_0}) + 2^{q_0}/m}$.

Then, $\Gamma^{-1}(a_2) = \frac{\Gamma^{-1}(4^{q_0})\sqrt{\lfloor m/2^{q_0}\rfloor+1}}{\sqrt{m/2^{q_0}}}$.

Then, $\Gamma^{-1}(a_2)\sqrt{m/2^{q_0}} = \Gamma^{-1}(4^{q_0})\sqrt{\lfloor m/2^{q_0}\rfloor + 1}$.

Then, $\frac{\Gamma^{-1}(a_2)\sqrt{m/2^{q_0}}}{m} = \frac{\Gamma^{-1}(4^{q_0})\sqrt{\lfloor m/2^{q_0}\rfloor+1}}{m}$.

Suppose that $L(Bo) \leq 1/2$.
There exists $K \leq 2^{q_0}/2$ such that $Bo = \cup_{s=1}^{K} I_k$ where

$$I_k = [k/2^{q_0}, (k+1)/2^{q_0}[\ .$$

Then, by using hypothesis 11.1.2 and lemma 11.1.3
$P\{X(n) \in Bo \mid X(n+j_2) = x_2, ..., X(n+j_p) = x_p\}$
$= \sum_{k=1}^{K} P\{X(n) \in I_k \mid X(n+j_2) = x_2, ..., X(n+j_p) = x_p\}$
$= \sum_{k=1}^{K} [L(I_k) + Ob(1)\epsilon_{I_k}]$
$= L(Bo) + Ob(1) \sum_{k=1}^{K} \frac{\Gamma^{-1}(4^{-q_0})\sqrt{N_{Iel}}}{m}$
$= L(Bo) + Ob(1) \sum_{k=1}^{K} \frac{\Gamma^{-1}(4^{-q_0})\sqrt{\lfloor m/2^{q_0}\rfloor+1}}{m}$
$= L(Bo) + Ob(1) \sum_{k=1}^{K} \frac{\Gamma^{-1}(a_2)\sqrt{m/2^{q_0}}}{m}$
$= L(Bo) + Ob(1) 2^{q_0-1} \frac{\Gamma^{-1}(a_2)}{\sqrt{m*2^{q_0}}}$
$= L(Bo) + Ob(1) \frac{2^{q_0/2}\Gamma^{-1}(a_2)}{2\sqrt{m}}$.

If $L(Bo) \geq 1/2$, we use Bo' his complement set in $\{0, 1/2^{q_0}, ..., (2^{q_0}-1)/2^{q_0}\}$ which satisfies the predicted equality. ■

Remark 11.1.5 *One could maybe to have a finer increase by considering the traditional problem of the samples.*

Study For example, for a sample $x(n) \in \{0/2^{84}, 1/2^{84}, ..., (2^{84}-1)/2^{84}\}$ of size $N = 1.000.000$, let us take the Borel set $Bo_1 = \cup_{n=1}^{N}[x_n - 1/2^{85}, x_n + 1/2^{85}[$. Then, $L(Bo_1) = N/2^{84} = 7.7292/10^{15}$. Though $P_e(Bo_1) = 1$.

It is the matter of the traditional problem of the test a posterori when the sample is known.

Because of the problems of this type, one can wonder whether there would not a way to lower the increase by taking $k \leq N$ instead of $k \leq 2^{q_0-1} = 2^{83}$. ($Bo = \cup_k I_k$). Indeed is it useful, for a sample of size N, to be interested with

intervals length smaller than 1/N, since, in this case, one can always obtain values of $P_e(Bo)$ completely different from L(Bo)?

If one can suppose $k \leq N$, one will improve the size of the final sample. It is the assumption that we had made page 14 of [16].

It is a study which remains to be made. ■

The following proposition is needed.

Lemma 11.1.6 *We suppose $N_3 = Nr_0S$. We suppose that the hypothesis 11.1.2 holds. We assume also that $\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{S}}{\sqrt{N_3}}\sqrt{r_0 m}$. Then, $|\epsilon_{Bo}| \leq \frac{\alpha}{\sqrt{q_0 N}}$*

Proof We have $\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{S}\sqrt{r_0}\sqrt{m}}{\sqrt{N_3}}$.

Then, $\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{m}}{\sqrt{N}}$.

Then, $\frac{2^{q_0/2}\Gamma^{-1}(a_2)}{2\sqrt{m}} \leq \frac{\alpha}{\sqrt{q_0 N}}$.

Then, by using lemma 11.1.4 , $|\epsilon_{Bo}| \leq \frac{\alpha}{\sqrt{q_0 N}}$. ■

These increases are imposed because, in our calculations, (cf section11.1.4), one realizes that it is necessary to impose $|\epsilon_{Bo}| \leq \frac{\alpha}{\sqrt{qN}}$.

Remark 11.1.7 *One chooses q_0 and r_0 such that q_0/r_0 is maximum and checking the condition $\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{S}\sqrt{r_0}\sqrt{m}}{\sqrt{N_3}}$ where m is not too large.*

Study Indeed, $q_0 N \approx q_0 N_3/(r_0 S)$ is the size of the obtained sample of bits $b^0(n')$. It is necessary thus that it is largest possible to have the best possible output.

It is thus necessary to choose q_0 and r_0 such as q_0/r_0 is maximum and checking the condition $\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{S}\sqrt{r_0}\sqrt{m}}{\sqrt{N_3}}$.

However that often forces to choose m very large. Indeed, the previous condition about amounts choosing $\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{S}}{\sqrt{N_3}}\sqrt{r_0\ \kappa^{3r_0/4}}$ considering that m is the largest element of the Fibonacci sequence such as $m \leq \kappa^{3r_0/4}$.

For example, if $\kappa = 32$, that returns to impose about $\frac{\sqrt{q_0}\Gamma^{-1}(4^{q_0})}{\sqrt{r_0}}\sqrt{2^{q_0 - 15r_0/4}} \leq \frac{2\alpha\sqrt{S}}{\sqrt{N_3}}$.

In order that q_0/r_0 is maximum, it is roughly necessary that $q_0 \approx 15r_0/4$. This result is obtained when q_0 and r_0 are big.

But, if one chooses m very large (e.g. about 10^{1000}), that can complicate calculations modulo m much.

It is finally simpler to take m large, but, not very large : cf example of section 11.2. ■

Results about the sequences of bits

Lemma 11.1.8 *The previous assumptions are kept. Then, for all sequences of bits bi_n, for all finite injective sequence j_s,*

$$P\big\{B^0(n') = bi_1 \big| B^0(n'+j_2) = bi_2, B^0(n'+j_3) = bi_3, ...\big\} = 1/2 + Ob(1)\ \frac{\alpha}{\sqrt{q_0 N}}\ .$$

Proof We define the J_s by $\{X(n+j_s) \in J_s\} = \big\{B^0(n'+j_s) = bi_s\big\}$. By proposition 4.2.3,

$$\begin{aligned} &P\big\{B^0(n') = bi_1 \big| B^0(n'+j_2) = bi_2, B^0(n'+j_3) = bi_3, ...\big\} \\ &= P\Big\{X(n) \in J_1 \Big| X(n+j_2) \in J_2, X(n+j_3) \in J_3, ...\Big\} \\ &= 1/2 + Ob(1)\epsilon_{Bo}\ .\blacksquare \end{aligned}$$

11.1.4 Explanation 2 : $\epsilon = \alpha/\sqrt{q_0 N}$

In this section, we have obtained sequences of bits $b^0(n')$ such that

$$P\big\{B^0(n') = bi_1 \big| B^0(n'+j_2) = bi_2, B^0(n'+j_3) = bi_3, ...\big\} = 1/2 + Ob(1)\epsilon\ .$$

By using the method described here, one will impose $\epsilon = \alpha/\sqrt{q_0 N}$ where Nq_0 is the size of sample $\{b^0(n')\}$.

Now apply the theorem 9 and the result of the section 11.2.10 .
Then, by equation 9.3, $\beta_{1,p} \leq \frac{2p\alpha}{A(p)^{1/2}2^{p/2}}$.

Indeed, $\beta_{1,p} \leq \frac{\sqrt{Nq_0}\epsilon_p}{\sqrt{A(p)L(Bo)}} \approx \frac{\sqrt{Nq_0}.2p\epsilon}{A(p)^{1/2}2^{p/2}}$.

Then, in order to use theorem 9, it is necessary that $\beta_{1,p}$ is small, e.g. $\beta_{1,p} \leq 0.1$. For this reason it is necessary that $\epsilon = \alpha/\sqrt{q_0 N}$.

Then, $\beta_{1,p}$ is enough small.

Now, use again the notations of theorem 9 . Then,

$$P\Big\{\sqrt{N}\big|P_e - (1/2)^p\big| \geq \sigma_B x\Big\} \leq \Gamma\big(\theta(x)x\big)\ ,$$

where $\theta(x) = \frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}} \leq 1$ and where

$$P_e = (1/N)\sum_n 1_{bi_1}[B^0(n')]1_{bi_2}[B^0(n'+j_2)]....1_{bi_p}[B^0(n'+j_p)]$$

with $bi_r \in \{0,1\}$ for r=1,2,...,p.

On the other hand, by the section 11.2.10, $\gamma_{1,p}$ is small when p is not too big.

Then, $\beta_{1,p}$ is enough small in order that θ is close to 1, i.e. if $\epsilon = \alpha/\sqrt{q_0 N}$, one will not be able to distinguish P_e in case IID from P_e when the assumptions of this section are checked : the empirical estimate of the probability that $(B^0(n'), B^0(n'+j_2), ..., B^0(n'+j_p)) = (bi_1, ..., bi_p)$ is about also close to $(1/2)^p$ that it would be it in case IID.

One obtains the same type of results for theorem 10.

It is not thus finally possible to distinguish the sequence $b^0(n')$ from an IID sample.

Remark 11.1.9 *The worst approximation takes place for the sample of maximum size : $q_0 N$. More this size is small, more the estimate is close to that of the IID distribution.*

11.1.5 Some other explanations

Practical conditions

When one uses a sequence of datas translated in numbers written base κ with size N_3, one obtains a sample of N' bits where $N' = N_3 q_0/[Sr_0]$. That means that the size of the obtained sample h(n) is smaller than that of the original sample d(j) (e.g. 15 times smaller, if S=10). One thus chooses the parameters in order that N_3/N' is largest possible in order to not to lose too many data : cf also Remark 11.1.7.

Use of other congruences than the Fibonacci congruence

One could choose other congruences than the Fibonacci congruence, for example $sup(h_i) = 3$ with the notation 6.1.2. However the Fibonacci congruence is that which gives the best theoretical results.

Use of limit theorems

Because generally, several files are used to constitute the data, one can use different files for each line: The lines are thus often independent : that facilitates the use of the CLT or the XORLT.

Problem about the sums

In the section 5.5.8, one saw that, with regard to the sums of random variables, to use the proposition 5.5.1 gives absurd results in some cases, whose the case

with continuous density. In order to avoid this problem, one can use the techniques described with the remarks 11.1.1 and 11.1.2.

11.2 Example : building of an IID sequence

In this section we study an example : we obtain a sequence of random bits $b^1(n')$ which can be obtained by asking it to rene.blacher@imag.fr (Laboratory LJK, University Joseph Fourier of Grenoble, France). More precisions on this subject will found in [18]. The data result from texts, mathematical texts and file of programming : cf section 10.3.

11.2.1 Choice of random datas

Then, we use the sequence a(j) defined in section 10.3.

Notations 11.2.1 *We note by a(j), $j = 1, ..., N_3$, the sequence of data which we used to build the sequence $b^1(n')$ where $N_3 = 298.159.056 \geq 28 * 10^7$, and $1 \leq a(j) \leq 256$. Then, a(j) can be regarded as the realization of a sequence of random variables A(j) defined on a probability space (Ω, Δ, P)*

Then, we transform these sequences of letters in numbers by using the appropriate function defined on the computer : a(j) $\in \{1, ..., 256\}$ (cf section 10.3).

Now, there are only 26 letters. But it is necessary to add, the capital letters, the ":" , ";" , etc. There will be many of these 256 numbers which will not appear not or little. Also, we will write these numbers in base 32 so that each number can have a probability reasonable to appear.

Notations 11.2.2 *We set $c(j) = \overline{a(j)}$ modulo 32 for $j = 1, 2, ..., N_3$.*

Therefore, $c(j) \in \{0, 1, ..., 31\}$.

11.2.2 Study of data

It is noted that it will be increasingly difficult to predict the $C(j) = c$ knowing *the past* : $C(j - j'_2) = c_2, ..., C(j - j'_p) = c_p$, $1 < j'_2 < .... < j'_p$ when j'_2 increases, i.e. $P\{C(j) = c \mid C(j - j'_2) = c_2, ..., C(j - j'_p) = c_p\} \to P\{C(j) = c\}$ as $j'_2 \to \infty$.

In fact, one understands that one can consider that one has a stronger assumption : one can suppose that the following hypothesis holds (cf sections 10.4 10.3.1).

Hypothesis 11.2.1 *The sequence C(j) is Qd-dependent with Qd=22.*

In order to prove this result, we have estimated for various d, various p, p', various injective sequences j_s^1, j_s^2 such that $j_1^1 = j_1^2 = 0$, various Borel sets $Bo = Bo_1 \otimes ..... \otimes Bo_p$, $Bo' = Bo'_1 \otimes ..... \otimes Bo'_{p'}$,

$$\mathbb{E}\big\{1_{Bo}[C(n)]1_{Bo'}[C(n+d)]\big\} - \mathbb{E}\big\{1_{Bo}[C(n)]\big\}\mathbb{E}\big\{1_{Bo'}[C(n+d)]\big\} \ ,$$

where

$$1_{Bo}[C(n)] = 1_{Bo_1}[C(n)]1_{Bo_2}[C(n+j_2^1)].....1_{Bo_p}[C(n+j_p^1)] \ ,$$

$$1_{Bo'}[C(n+d)] = 1_{Bo'_1}[C(n+d)]1_{Bo'_2}[C(n+d+j_2^2)].....1_{Bo'_{p'}}[C(n+d+j_{p'}^2)] \ ,$$

and where $Min_{s,t}\big(|(n+j_s^1) - (n+d+j_t^2)|\big) \geq Qd = 22$.

We have always obtained

$$\mathbb{E}\big\{1_{Bo}[C(n)]1_{Bo'}[C(n+d)]\big\} - \mathbb{E}\big\{1_{Bo}[C(n)]\big\}\mathbb{E}\big\{1_{Bo'}[C(n+d)]\big\} \approx 0 \ . \quad (11.1)$$

We made these estimates for many Borel sets except for those which are too small (in this case this estimate is impossible).

Now for $Min_{s,t}\big(|(n+j_s^1) - (n+d+j_t^2)|\big) \geq 22$, by logical reasons, we know that

$$\big|\mathbb{E}\big\{1_{Bo}[C(n)]1_{Bo'}[C(n+d)]\big\} - \mathbb{E}\big\{1_{Bo}[C(n)]\big\}\mathbb{E}\big\{1_{Bo'}[C(n+d)]\big\}\big| \to 0 \ as \ d \to \infty \ .$$

For example, if the C(j)'s derive from text, it is more difficult to predict, C(j+t) when t increases.

In conclusion, one can admit that equation 11.1 holds always for $Min_{s,t}\big(|(n+j_s^1) - (n+d+j_t^2)|\big) \geq 22$. It is enough to prove that $b^1(n')$ is IID : cf section 11.2.10.

11.2.3 Writing in number with r_0 digits

a) We set $\alpha = 0.02$.

b) We choose S=10.

c) We choose $q_0 = 57$ and $r_0 = 28$, $m^1 = m^F(32^{28})$,
$32^{20} < m = 145448911123277268367830664195 3 < m^F(32^{21}) = m^F(32^{3*28/4})$.
Let $a = 89892370700847998927429085014 5 < 32^{20} \leq m$. Then m and a are the parameters of a Fibonacci congruence.

Then,

$$m/2^{q_0} \geq 1001,$$

$$\sqrt{q_0}2^{q_0/2}\Gamma^{-1}(a_2) \leq \frac{2\alpha\sqrt{S}}{\sqrt{N_3}}\sqrt{r_0 m}\ ,$$

where $a_2 = \Gamma\Big(\Gamma^{-1}(4^{-q_0})\sqrt{\lfloor m/2^{q_0}\rfloor/(m/2^{q_0}) + 2^{q_0}/m}\Big) \approx 4.8148/10^{35}$.

Remark 11.2.1 *We remark that* $a_2 = 4.814824860950729/10^{35}$. *Then, we remark also that* $a_2 \approx 4^{-q_0} = 4.814824860968090/10^{35}$.

To make uniform the marginals distribution

We make uniform the marginal distribution by using the method defined in section 11.1.2.

Now, $\lfloor N_3/28\rfloor = \lfloor 298.159.056/28\rfloor = \lfloor 10.648.537.7142857\rfloor = 10.648.537$.

Notations 11.2.3 *We set* $d(j) = \sum_{r=1}^{28} 32^{28-r} c(28(j-1)+r)$ *for* $j = 1, 2, 3, ...10.648.537$.

Then d(j) can be regarded as extracted of a 2-dependent sequence.

Notations 11.2.4 *For* $j = 1, 2, ..., 10.648.537$, *we set* $e^1(j) = \lfloor d(j)[m^1/32^{28}]\rfloor$.

Notations 11.2.5 *For* $j = 1, 2, ..., 10.648.537$, *we set* $e^2(j) = \overline{e^1(j) + rand_0(j)}$ *where* $rand_0(j)$ *is a pseudo-random generator with values in* $F^*(m^1)$ *and with period* m^1.

Notations 11.2.6 *For* $j = 1, 2, ..., 10.648.537$, *we set* $e^3(j) = m.T_1^m(e^2(j)/m^1)$.

11.2.4 Transformation in table

In order to use the limit theorems, one writes the $e^3(j)$'s in the form of a table with 10 lines.

a) We denote by $e^4(t)$, $t = 1, 2, ..., 10.000.000$ a subsequence of $e^3(j)$ obtained by suppressing some sequences $\{e^3(\rho_u), e^3(\rho_u + 1)....., e^3(\rho_u + n_4)\}$.
Comment We have suppressed some $e^3(j)$ in order that there is independence between the lines defined below.

Thus, the union of several files gave us the beginning of the sequence $e^3(j)$: j = 1,2,....,1.052.007. We have keep only the 1.000.000 first in order to form the first line f(1,n) below. By this way the second line f(2,n) will be formed of $e^3(j)$ coming from other files and will be thus independent of the first one.

Notations 11.2.7 *Let $f(i,n) = e^4(n + N(i-1))$ for i=1,...,10 , n= 1,...,N where $N = 1.000.000$.*

In section 11.2.2, it was understood that

$$P\big\{F(i,n) = f \mid F(i,n-j'_2) = f_2, ..., F(i,n-j'_p) = f_p\big\} \to P\{F(i,n) = f\}$$

as $j'_2 \to \infty$ when $j'_1 = 0 < j'_2 < .... < j'_p$. In order to have the same result for

$$P\{F(i,n) = f | F(i,n+j'_2) = f_2, ..., F(i,n+j'_p) = f_p\} \to P\{F(i,n) = f\},$$

we invert the even lines.

Notations 11.2.8 *If "i" is even, we set $f_1(i,n) = f(i,N-n+1)$ for i=1,...,10, n = 1,...,N. If i is odd, we set $f_1(i,n) = f(i,n)$ for i=1,...,10, n = 1,...,N.*

Therefore, logically, when one will summon the lines $f_1(i,n)$, it is reasonable to think that it will be difficult to predict $\sum_i f_1(i,n)$ knowing elements which are past **or** future. One will thus have, for any injective sequence j_s such that $j_1 = 0$,
$P\Big\{ \sum_i F(i,n) = g \Big| \sum_i F(i,n+j_2) = g_2, ..., \sum_i F(i,n+j_p) = g_p \Big\} \approx \sum_i P\{F(i,n) = g\}$,
if $min_s(|j_s|)$ is large enough.

11.2.5 Use of limit theorems

Then, we can use the CLT and the XORLT.

Notations 11.2.9 *Let $g(n) = \sum_{i=1}^{10} f_1(i,n)$ for n = 1,...,1.000.000.*

Notations 11.2.10 *Let $h(n) = \overline{g(n)}$ modulo m, for n=1,.....,1.000.000.*

11.2.6 Use of the Fibonacci Congruence

Notations 11.2.11 *We set $k(n) = \overline{T}\big(h(n)\big)\big)$ for all n= 1,2,....,N.*

Notations 11.2.12 *For all n = 1,2,....,N, we set $r(n) = k(n)/m$.*

11.2.7 Building of a random sequence $x(n)$

Notations 11.2.13 *Let $r(n) = \overline{\overline{0, b_1^n b_2^n ....}}$, the binary writting of r(n). We set $x(n) = \overline{\overline{0, b_1^n b_2^n ....b_{q_0}^n}}$ ($q_0 = 57$).*

Notations 11.2.14 *We set $b'_{q_0 n-r+1} = b_r^n$ for n=1,...,N and $r = 1, ..., q_0$.*

Notations 11.2.15 *We denote the sequence b'_n by $b^1(n')$, $n' = 1, 2, ..., Nq_0$.*

11.2.8 Some remarks

Study of data

We checked in the chapter 10 the validity of our assumptions on the numerical data of this example. We point out that we found that all the necessary assumptions about the various steps of this construction were checked, and by far!

Problem about sums

In the section 5.5.8, one saw that to take the proposition 5.5.1 gives absurd results for the sums of random variables in some cases. It is not the case for the f(i,n) used in this section. Therefore, one has does not have to use the techniques described with the remarks 11.1.1 and 11.1.2.

The data d(j) are obtained starting from sufficiently different electronic files: texts in various languages, mathematical texts, files of programming. The distribution of the theoretical probability of these data must be regarded as chosen randomly.

It is also true because we have applied the application $T_1^m : e^3(j) = mT_1^m(e^2(j)/m^1)$. According to the section 8.4, that makes uniform and also independent the probabilities. The variations with the uniformity in the case of p dimensions are thus very small. But it remains random considering there is always no connection between texts and the sets A^t described in section 8.4.

That is also true for the conditional probabilities considering that

$$P\{X_n|x_2,....,x_p\} = \frac{P\big\{\{X_{n+j_1} = x_1\} \cap \{X_{n+j_2} = x_2\} \cap ..... \cap \{X_{n+j_p} = x_p\}\big\}}{P\big\{\{X_{n+j_2} = x_2\} \cap ..... \cap \{X_{n+j_p} = x_p\}\big\}} .$$

Now, it is precisely what we want to be able to do : to apply our results to the conditional probabilities (cf hypothesis 11.1.2). The proposition 5.5.2 can thus be applied to results of this section 11.2.

Choice of S

Normally, one would have to make previous calculations for several S in order to improve the results. We did it in study numerical. In fact, in practice, as for much of other files which are on computers, one finds that one can be satisfied with S=10.

11.2.9 Properties of $B^1(n')$

We will study the empirical aspect by using theorems 9 and 10 .

For that let us notice that, according to the previous results, the following assumption will be satisfied.

Hypothesis 11.2.2 *: For all b=0 or 1*

$$P\big\{B^1(n') = b \;\big|\; B^1(n'+j_2) = b_2, ..., B^1(n'+j_p) = b_p\big\} = 1/2 + \frac{Ob(1)\alpha}{\sqrt{Nq_0}} \ ,$$

where $\alpha/(Nq_0)^{1/2} \approx 2.649/10^6$.

Remark that Nq_0 is the size of sample of the $b^1(n')$.

Use of 2-dependence in empirical theorems

We understood in the study of the data which one can consider that sequence C(j) is Qd-dependent with Qd=22. In particular, with the notations of sections 9.2 and 9.3, hypotheses 9.2.1 and 9.3.1, one deduces from equation 11.1 that, for all Borel set, for $Min_{s,t}\big(|(n+j_s^1) - (n+d+j_t^2)|\big) \geq 22$,

$$\mathbb{E}\big\{\big(1_{Bo}[C(n)] - L'_n\big)\big(1_{Bo}[C(n+d)] - L'_{n+d}\big)\big\} = 0 \ ,$$

$$\mathbb{E}\Big\{\big[1_{Bo_1}(C_n) - L(Bo_1)\big]\big[1_{Bo_1}(C_{n+d}) - L(Bo_1)\big]1_J(C_{n+j})1_J(C_{n+d+j})\Big\} = D'_n D'_{n+d} \ ,$$

where $L'_n = \mathbb{E}\big\{1_{Bo}[C(n)]\big\}$, $D'_n = \mathbb{E}\Big\{\big[1_{Bo_1}(C_n) - L(Bo_1)\big]1_J(C_{n+j})\Big\}$.

That means that the hypotheses 9.2.1 and 9.3.1 hold for the sequence C(j).

Now, D(j) is Qd_D dependent with $Qd_D = 2$. It is also the case for the sequences F(i,n). That means that, with the notations of sections 9.2 and 9.3, hypotheses 9.2.1 and 9.3.1, for all Borel set Bo, for all $Min_{s,t}\big(|(n+j_s^1) - (n+d+j_t^2)|\big) \geq Qd_X = Qd_D = 2$,

$$\mathbb{E}\big\{\big(1_{Bo}[X(n)] - L_n\big)\big(1_{Bo}[X(n+d)] - L_{n+d}\big)\big\} = 0 \ ,$$

$$\mathbb{E}\Big\{\big[1_{Bo_1}(X_n) - L(Bo_1)\big]\big[1_{Bo_1}(X_{n+d}) - L(Bo_1)\big]1_J(X_{n+j})1_J(X_{n+d+j})\Big\}$$
$$= D_n D_{n+d} \ .$$

Now, the sequence $b^1(n')$ is Qd_B-dependent with $Qd_B = 57$. Therefore, with the notations of the hypotheses 9.2.1 and 9.3.1, $k_B = 0$ and $K_B = 0$ for $q = Qd_B = 57$.

As a matter of fact we will understand in sections 11.2.10 and 11.2.11 that one has results much more stronger. But this Qd_B-dependence enables us to obtain a sure increase to the probabilities of the samples.

11.2.10 Use of theorem 9

Use of 2-dependence in theorem 9

First, apply the theorem 9 with the 2-dependence.

We study a sample $b^1(\psi(n))$, $n = 1, 2, ..., N_1$, where $N_1 \leq N$ and where $\psi : \{1, 2, ..., N_1\} \to \{1, 2, ..., N\}$ is an injective function.

We choose $p \leq 40$: we choose a such p because there is little interest to study the case where $2^p \geq q_0 N = 57.000.000$, i.e. $p \geq 26$. But in order to be completely sure that all our results are correct, we study this theorem untill p=40 : (cf Remark 11.2.2).

Of course, the simplest technique for using the theorem 9, is to choose $q = Qd_B = 58$.

First, compute $\beta_{1,p}$ and $\gamma_{1,p}$.

By the results of section 9.5.1, we know $\beta_{1,p} \leq \frac{\sqrt{N_1}\epsilon_p}{\sqrt{A(p)L(Bo)}}$.

Then, one can admit
$\beta_{1,p} \leq \frac{\sqrt{N_1}2p\epsilon}{\sqrt{A(p)2^p}} = \frac{\sqrt{N_1}2p\alpha}{\sqrt{q_0 N A(p)2^p}} = \frac{\sqrt{N_1}}{\sqrt{N q_0}} \ \frac{2p\alpha}{A(p)^{1/2}2^{p/2}}$.

Then, for $N_1 = N q_0$,

$\beta_{1,p} \leq$	0.04000000000000	if p=1
$\beta_{1,p} \leq$	0.08000000000000	if p=2
$\beta_{1,p} \leq$	0.12000000000000	if p=3
$\beta_{1,p} \leq$	0.09237604307034	if p=4
$\beta_{1,p} \leq$	0.06030226891555	if p=5
$\beta_{1,p} \leq$	0.01309541208943	if p=10
$\beta_{1,p} \leq$	0.00332528643438	if p=15
$\beta_{1,p} \leq$	0.00078139197225	if p=20
$\beta_{1,p} \leq$	0.00001726334915	if p=25
$\beta_{1,p} \leq$	0.00003662110860	if p=30
$\beta_{1,p} \leq$	0.00000152587891	if p=40

Moreover,

$$\gamma_{1,p} \approx \frac{(p^2 - p + 1)\alpha}{2A(p)\sqrt{q_0 N}} \left[2p + (1 + 4Qd_B)\frac{4p}{2^p} + (1 + 2Qd_B)\frac{4p^2\epsilon}{2^p} \right] .$$

Then,

$\gamma_{1,p} \leq$	0.00121856976850	if p=1
$\gamma_{1,p} \leq$	0.00734320738757	if p=2
$\gamma_{1,p} \leq$	0.02592374729248	if p=3
$\gamma_{1,p} \leq$	0.02176471569129	if p=4
$\gamma_{1,p} \leq$	0.01239039814018	if p=5
$\gamma_{1,p} \leq$	0.00382913360640	if p=10
$\gamma_{1,p} \leq$	0.00855657503967	if p=15
$\gamma_{1,p} \leq$	0.02022398420111	if p=20
$\gamma_{1,p} \leq$	0.03980480582574	if p=25
$\gamma_{1,p} \leq$	0.06922022014741	if p=30
$\gamma_{1,p} \leq$	0.16540760122551	if p=40

As matter of fact, in some cases, $\gamma_{1,p}$ is too large. Then, we can use the approximation of $\gamma'_{1,p}$ of lemma 9.7.3. In this case,

$$\gamma'_{1,p} \approx \frac{Ob(1)\zeta'(p)\alpha}{2A(p)\sqrt{q_0 N}} + \frac{(p^2-p+1)\alpha}{2A(p)\sqrt{q_0 N}}\Big[(1+4Qd_B)\frac{4p}{2^p} + (1+2Qd_B)\frac{4p^2\epsilon}{2^p}\Big] ,$$

where $\zeta'(p) = 2p + 2\sum_{r=1}^{p-1} \frac{2(p+r)}{2^r}$.

Then,

$\gamma'_{1,p} \leq$	0.00121856976850	if p=1
$\gamma'_{1,p} \leq$	0.00021067452526182	if p=2
$\gamma'_{1,p} \leq$	0.00733261112871	if p=3
$\gamma'_{1,p} \leq$	0.02122430648961	if p=4
$\gamma'_{1,p} \leq$	0.01171705405466	if p=5
$\gamma'_{1,p} \leq$	0.00128197696378	if p=10
$\gamma'_{1,p} \leq$	0.000248582465482	if p=15
$\gamma'_{1,p} \leq$	0.000178421399711	if p=20
$\gamma'_{1,p} \leq$	0.00020982313313	if p=25
$\gamma'_{1,p} \leq$	0.00024904181006	if p=30
$\gamma'_{1,p} \leq$	0.0003284840939180	if p=40

Then, by theorem 9, for a sample of size $N_1 \leq Nq$,

$$P\Big\{\sqrt{N_1}\big|P_e^B - (1/2)^p\big| \geq \sigma_B x\Big\} \leq \Gamma[\theta(x)x] ,$$

where $P_e^B = (1/N_1)\sum_n 1_{bi_1}(B_n)1_{bi_2}(B_{n+j_2})....1_{bi_p}(B_{n+j_p})$.

Then, θ is close to 1 with $\theta(x) = \frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}} \leq 1$. One can also choose $\theta(x) = \frac{1-\beta_{1,p}/x}{1+\gamma'_{1,p}} \leq 1$.

Finally, the following table is obtained.

		x=1	x=1.5	x=2	x=2.5	x= 3
p=1	$\theta(x) \geq$	0.9588	0.9732	0.9788	0.9830	0.9854
p=3	$\theta(x) \geq$	0.9114	0.9391	0.9512	0.9580	0.9596
p=5	$\theta(x) \geq$	0.9291	0.94888	0.9587	0.9647	0.9686
p=10	$\theta(x) \geq$	0.9857	0.9900	0.9922	0.9935	0.9944
p=15	$\theta(x) \geq$	0.9964	0.9975	0.9981	0.9984	0.9986
p=20	$\theta(x) \geq$	0.9990	0.9993	0.9994	0.9995	0.9995
p=25	$\theta(x) \geq$	0.9996	0.9997	0.9997	0.9997	0.9997
p=30	$\theta(x) \geq$	0.9997	0.9997	0.9997	0.9997	0.9997
p=40	$\theta(x) \geq$	0.9996	0.9996	0.9996	0.9996	0.9996

In particular, $0.9291 \leq \theta \leq 1$ if $x \geq 1$.

Then, one obtains the following increases for $P\{\sqrt{N_1}\left|P_e^B - (1/2)^p\right| \geq \sigma_B x\}$:

		x=1	x=1.5	x=2	x=2.5	x= 3
Under IID hypothesis	p=1	0.317	0.133	0.045	0.012	0.0027
Under hypothesis 11.2.2	p=1	0.337	0.144	0.050	0.013	0.0031
	p=3	0.321	0.136	0.046	0.012	0.0040
	p=5	0.352	0.154	0.05	0.015	0.0036
	p=10	0.324	0.137	0.047	0.013	0.0028
	p=15	0.319	0.134	0.046	0.012	0.0027
	p=20	0.317	0.133	0.045	0.012	0.0027
	p=25	0.317	0.133	0.045	0.012	0.0027
	p=30	0.317	0.133	0.045	0.012	0.0027
	p=40	0.317	0.133	0.045	0.012	0.0027

One obtains similar results for samples of size $N_1 \leq Nq$.

Then, it is difficult to differentiate the sequence $b^1(n')$ from an IID sample. Indeed, if our data were not IID, that would imply that $\sqrt{Nq_0}\left|P_e^B - (1/2)^p\right|$ would be large. That can certainly occur for some $(bi_1, bi_2, ..., bi_p)$ but, as the previous increases show it, with a probability which is not too different from that of IID case.

Finally, by using P_e^B, it is not possible to differentiate the sequence $b^1(n')$ from an IID sequence.

Moreover, the worst approximation occurs for the sample of maximum size: $N_1 = q_0 N$. The smaller this size is, the more our results approaches those of

IID case : $\beta_{1,p} \leq \frac{\sqrt{N_1}}{\sqrt{Nq_0}} \ \frac{2p\alpha}{A(p)^{1/2}2^{p/2}}$.

One can thus say that, with regard to the behavior of the P_e^B, that it is not possible to differentiate $b^1(n')$ from an IID sequence.

Use of independence in theorem 9

In the previous results, one has increased $\gamma_{1,p}$ by using the 2-dependence. In fact, this 2-dependence is the dependence existing for the sequence D(j). For the sequence $B^1(n')$, the results are much better because we did everything in our building so that it is identical to a sequence IID.

One could thus have finer increases. For that, it is necessary to calculate $\sigma_1^2 = \sigma_B^2[1 + 2\gamma_p^1]$, i.e. it is necessary to calculate γ_p^1 which is defined by this previous equality. We have estimated σ_1^2 and have compared it with the exact value of σ_B^2 for p=1,2,3,4,5.

We thus obtained the following increases of $|\sigma_1^2 - \sigma_B^2|$ for different p and different $bi_1 \otimes ..... \otimes bi_p$.

p	1	2	3	4	5
$\lvert\sigma_1^2 - \sigma_B^2\rvert \leq$	$2 * 10^{-4}$	$1 * 10^{-6}$	$3 * 10^{-7}$	$7 * 10^{-9}$	$2 * 10^{-9}$

That means $\gamma_p^1 \approx 0$. One thus obtains finer increases of

$$P\big\{\sqrt{N_1}\big|P_e^B - (1/2)^p\big| \geq \sigma_B x\big\} \ .$$

Indeed, one obtains the following majorations for

$$P\big\{\sqrt{N_1}\big|P_e^B - (1/2)^p\big| \geq \sigma_B x\big\} \ :$$

		x=1	x=1.5	x=2	x=2.5	x= 3
Under IID hypothesis	p=1	0.317	0.133	0.045	0.012	0.0027
Under hypothesis 11.2.2	p=1	0.317	0.133	0.045	0.012	0.0028
	p=3	0.317	0.133	0.045	0.012	0.0028
	p=5	0.317	0.133	0.045	0.012	0.0028
	p=10	0.317	0.133	0.045	0.012	0.0027

One concludes that the sequence $b^1(n')$ cannot be differentiated from an IID sequence.

11.2.11 Use of theorem 10

Use of the 2-dependence in theorem 10

It is pointed out that $P\Big\{\sqrt{N_1}\Big|\frac{P_e^B}{p_e^B} - L(I)\Big| > \sigma_{cp}\ x\Big\} \leq \Gamma\big(\theta'(x)x\big)$, where $\theta'(x) = \frac{1-\beta_{2,p}/x}{1+\gamma_{2,p}}$, where $P_e^B = (1/N_1)\sum_n 1_{bi_1}(B_{n+j_1})1_{bi_2}(B_{n+j_2})....1_{bi_p}(B_{n+j_p})$ and where $p_e^B = (1/N_1)\sum_n 1_{bi_2}(B_{n+j_2})....1_{bi_p}(B_{n+j_p})$.

By the results of section 9.5.2, the values $\beta_{2,p}$ (lemma 9.5.5) and $\gamma_{2,p}$ are produced by the formulas

$$\beta_{2,p} \approx \frac{\sqrt{N_1}}{\sqrt{Nq_0}}\ \frac{16\alpha}{2^p} = \frac{\sqrt{N_1}}{\sqrt{Nq_0}}\ \frac{0.32}{2^p}$$

and

$$\gamma_{2,p} = \epsilon\Big(2\xi_p + 2[1+\epsilon\xi_p]\sqrt{\theta_p} + \epsilon\theta_p\Big)\Big],$$

where (cf lemma 9.5.7 and 9.5.8)

$$\xi_p \approx \frac{1}{2}\Big[(p-1)[11p+16] + \frac{16(2Qd_B+1)(p^2-p+1)}{2^p}\Big]$$

$$\theta_p \approx \frac{8\sigma_B(J)^2}{L(J)} + 16(p^2-3p+3)(p-1)\epsilon\Big[1 + \frac{2}{2^{p-1}} + \frac{8Qd_B}{2^{p-1}}\Big]\ .$$

Moreover, one can assume that $\sigma_B(J)^2 \leq 4L(J)$ (cf Hypothesis 9.7.1). Lastly, we have always $\epsilon = 2.649/10^6$.

Then, for $N_1 = Nq_0$, the table of $\beta_{2,p}$, p=1,2,... is

0.16	0.008	0.004	0.002	0.001	0.0005.

The table of $\gamma_{2,p}$, p=1,2,...40, is

0.00016	0.0003	0.00048	0.0006	0.0009	0.0011	0.0014	0.0018
0.00222	0.0027	0.00325	0.0038	0.00450	0.0052	0.0059	0.0068
0.00763	0.0085	0.00951	0.0105	0.01159	0.0127	0.0139	0.0151
0.016376	0.0177	0.01908	0.0205	0.02198	0.0235	0.02510	0.0267
0.02842	0.0301	0.03194	0.0338	0.03568	0.0376	0.0396	0.0417

Thus finally, the decreases of $\theta'(x)$ are obtained

		x=1	x=1.5	x=2	x=2.5	x= 3
p=1	$\theta(x) \geq$	0.8399	0.8932	0.9199	0.9359	0.9465
p=3	$\theta(x) \geq$	0.9595	0.9729	0.9795	0.9835	0.9862
p=5	$\theta(x) \geq$	0.9891	0.9925	0.9941	0.9951	0.9958
p=10	$\theta(x) \geq$	0.9970	0.9971	0.9971	0.9972	0.9972
p=15	$\theta(x) \geq$	0.9941	0.9941	0.9941	0.9941	0.9941
p=20	$\theta(x) \geq$	0.9896	0.9896	0.9896	0.9896	0.9896
p=27	$\theta(x) \geq$	0.9813	0.9813	0.9813	0.9813	0.9813
p=28	$\theta(x) \geq$	0.9799	0.9799	0.9799	0.9799	0.9799
p=30	$\theta(x) \geq$	0.9770	0.9770	0.9770	0.9770	0.9770
p=40	$\theta(x) \geq$	0.9600	0.9600	0.9600	0.9600	0.9600

For example

$$P\left\{ \sqrt{N_1} \left| \frac{P_e^B}{p_e^B} - (1/2) \right| > \sigma_{cp}\ x \right\}$$

		x=1	x=1.5	x=2	x=2.5	x= 3
Under IID assumption	p=1	0.317	0.133	0.045	0.012	0.0027
Under assumption 11.2.2	p=1	0.401	0.180	0.065	0.019	0.0045
	p=3	0.337	0.144	0.050	0.014	0.0031
	p=5	0.323	0.137	0.047	0.013	0.0028
	p=10	0.319	0.135	0.046	0.013	0.0028
	p=15	0.320	0.136	0.047	0.013	0.0029
	p=20	0.322	0.138	0.048	0.013	0.0030
	p=27	0.326	0.141	0.045	0.014	0.0032
	p=28	0.327	0.142	0.050	0.014	0.0033
	p=30	0.329	0.143	0.051	0.015	0.0034
	p=40	0.337	0.150	0.055	0.016	0.0040

Remark 11.2.2 *One cannot increase suitably* $P\{\sqrt{N_1}|\frac{P_e^B}{p_e^B} - (1/2)| > \sigma_{cp}\ x\}$ *for any p. But it is not important.*

Study If p increases, θ_p and ξ_p - and therefore $\gamma_{2,p}$ - increase. Then $\theta'(x)$ converges to 0.

The theorem 10 cannot thus be used if p is too large under the assumption of the 2-dependence. In fact, that does not have importance.

Indeed, one will be face to similar problems in case IID : let $N_e = P_e N_1$ be the number of points $\left(b^1(n'), b^1(n'+j_2), ..., b^1(n'+j_p)\right)$ which belongs to $Bo = bi_1 \otimes bi_2 \otimes ..... \otimes bi_p$.

Under the assumption IID, $P\{N_e \geq 1\}$ is almost equal to 0 if L(Bo) is very small. For example for a sample of size 100 and a Borel set Bo of measure L(Bo) =1/10000, there is approximately a probability equal to 1/100 that $N_e \geq 1$. Therefore, probably $P_e^B - (1/2)^p = -(1/2)^p$. For example, if x=1, that means

that, if p is large, $P\{\sqrt{N_1}|P_e - (1/2)^p| > \sigma_B\} \approx 0$ whereas for a sample of sufficient size, $P\{\sqrt{N_1}|P_e - (1/2)^p| > \sigma_B\} \approx 0.3173$.

Moreover, $\frac{P_e^B}{p_e^B} - (1/2) = \frac{0}{0} - (1/2)$ it is not computable.

There are the same results for our data which are not IID. But they are Qd-dependent. That is enough. If p is large, $N_e = 0$ in the majority of the cases : $P\{\sqrt{N_1}|P_e - (1/2)^p| > \sigma_B\} \approx 0$.

Thus for a fixed $(bi_1, ..., bi_p)$ it is always possible that there is a " n' " such that $1_{bi_1}(b^1(n'))1^1_{bi_2}(b^1(n' + j_2))..... 1_{bi_p}(b^1(n' + j_p)) = 1$. But, it is what occurs in the case IID

Of course, it is the same thing for $N_e = p_e N_1$.

Then, it doesn't make sense to study the behavior of P_e^B/p_e^B when $p > 40$. That means that our results include all the cases which had to be studied. ■

Use of independence in theorem 10

As for $\gamma_{1,p}$, one increases $\gamma_{2,p}$ by estimate. One calculates $\sigma_2^2 = \sigma_{cp}^2[1 + \gamma_p^2]$.

We thus obtained the following increases of $|\sigma_2^2 - \sigma_{cp}^2|$ for various p and various $bi_1 \otimes ..... \otimes bi_p$.

p	1	2	3	4	5
$\lvert\sigma_2^2 - \sigma_{cp}^2\rvert \leq$	$1.1 * 10^{-4}$	$1.2 * 10^{-4}$	$1 * 10^{-4}$	$0.9 * 10^{-4}$	$1 * 10^{-4}$

For example

$$P\left\{\sqrt{N_1}\left|\frac{P_e^B}{p_e^B} - (1/2)\right| > \sigma_{cp}\ x\right\} =$$

		x=1	x=1.5	x=2	x=2.5	x= 3
Under IID hypothesis	p=1	0.317	0.134	0.045	0.012	0.0027
Under hypothesis 11.2.2	p=1	0.317	0.134	0.046	0.012	0.0040
	p=3	0.317	0.134	0.046	0.012	0.0027
	p=5	0.317	0.134	0.046	0.013	0.0028
	p=10	0.317	0.134	0.045	0.012	0.0027

One obtains the same conclusions that previously. The results are not changed so much. But, one does not have to worry for the case where p is large.

11.2.12 Conclusion

The inequalities above show that it could be that $b^1(n')$ does not check certain tests of an IID sequence, but that will occur with hardly more probabilities that

for a sample of a really IID sequence. It is thus not possible to differentiate the sequence $b^1(n')$ from an IID sample by using these tests.

Thus by an empirical opinion, it is not possible to differentiate the sequence $b^1(n')$ from an IID sample by using P_e^B and P_e^B/p_e^B. Now, we understood in section 2.1 that these are these conditions that one admits for the definition of the randomness

Let us notice that these results remain true for any subsequence $b^1(t(n'))$ that we could choose : it is a very strong result

The sequence $b^1(n')$ satisfy all the conditions which we have indicates in section 2.1. Then, one can already consider that $b^1(n')$ is an IID sample.

Morever by an not-empirical opinion, we know that

$$P\big\{B^1(n') = b\big|B^1(n'+j_2) = b_2, ..., B^1(n'+j_p) = b_p\big\} = 1/2 + Ob(1)\epsilon \ .$$

It is an additional property. However it is very important. If there is only this only property, that proves already that one can admit that $b^1(n')$ is an IID sample if ϵ is small enough.

For these two reasons, one concludes that $B^1(n')$ can be regarded as a sample of an IID sequence of random variables.

11.2.13 Results about $x(n)$

One obtains the same type of equations with X(n) as with the $B^1(n')$, i.e. x(n) could be regarded as sample of IID variables X(n) : for all Borel set Bo,

$$P\big\{X(n) \in Bo\big|X(n+j_2) = x_2, ..., X(n+j_p) = x_p\big\} = L(Bo) + Ob(1)2.649/10^6 \ .$$

Now study the theorem 9 :

$$P\Big\{\sqrt{N_1}\big|P_e^X - P(Bo_1)....P(Bo_p)\big| \geq \sigma_B x\Big\} \leq \Gamma(\theta(x).x) \ ,$$

where

$$P_e^X = (1/N_1)\sum_n 1_{Bo_1}(X(n+j_1))....1_{Bo_p}(X(n+j_p)) \ .$$

Then, one has to study the p such that $2^p \leq N = 1.000.000$, i.e. $p \leq 20$. In fact one chooses to study $p \leq 30$.

Then, for $N_1 = N$,

$\beta_{1,p} \leq$	0.0044	if p=1
$\beta_{1,p} \leq$	0.0131	if p=3
$\beta_{1,p} \leq$	0.0065	if p=5
$\beta_{1,p} \leq$	0.0014	if p=10
$\beta_{1,p} \leq$	0.000014	if p=20
$\beta_{1,p} \leq$	0.000004	if p=30

Moreover,

$\gamma'_{1,p} \leq$	0.000560	if p=1
$\gamma'_{1,p} \leq$	0.00357	if p=3
$\gamma'_{1,p} \leq$	0.0007	if p=5
$\gamma'_{1,p} \leq$	0.00004	if p=10
$\gamma'_{1,p} \leq$	0.00000029	if p=20

Thus finally, one obtains the following decreases of $\theta(x)$:

		x=1	x=1.5	x=2	x=2.5	x= 3
p=1	$\theta(x) \geq$	0.9595	0.9728	0.9795	0.9834	0.9861
p=3	$\theta(x) \geq$	0.8769	0.9167	0.9367	0.9486	0.9566
p=5	$\theta(x) \geq$	0.9390	0.9591	0.9692	0.9752	0.9792
p=10	$\theta(x) \geq$	0.9869	0.9912	0.9934	0.9947	0.9956
p=20	$\theta(x) \geq$	0.9992	0.9995	0.9996	0.9997	0.9997

For example

$$P\Big\{\sqrt{N_1}\big|P_e^X - L(Bo)\big| \geq \sigma_B x\Big\} =$$

		x=1	x=1.5	x=2	x=2.5	x= 3
Under IID hypothesis	p=1	0.317	0.134	0.045	0.012	0.0027
Under hypothesis 11.2.2	p=1	0.337	0.144	0.050	0.014	0.0031
	p=3	0.381	0.170	0.061	0.018	0.0041
	p=5	0.348	0.150	0.053	0.015	0.0033
	p=10	0.324	0.137	0.047	0.013	0.0028
	p=20	0.318	0.134	0.046	0.012	0.0027

One obtains similar results with the second empirical theorem.

These results shows that x(n) behave like an IID sample. It is an additional proof that one can regards $b^1(n)$ as an IID sample.

11.2.14 Tests

We are going to verify the previous conclusions by making tests of randomness. We use the classical Diehard tests cf [2], [1]. We tested the sequences $b^1(n')$ or x(n).

Results are in accordance with what we waited: for sequences $b^1(n')$ and x(n) the hypothesis "randomness" can be accepted by all the Diehard tests.

We denote by N_1 the size of used samples. We denote by α_5 the selected percentage points with probability 0,95.

In the following results, we test several samples of size N_1. We take the maximum of the statistics associated in these tests. It is not thus surprising that they are sometimes larger than α_5. As a matter of fact, We notice that these maximums are close to α_5 and that this result is even almost too good

Equidistribution Test We use the chi-squared tests. First, we test $b^1(n')$: of course the associated partition is $\big\{\{0\},\{1\}\big\}$.

We use sample with size N_1. Let $\chi^2_{N_1}$ be the associated chi squared statistics. We take the maximum of $\chi^2_{N_1}$. The number of samples increase if N_1 decrease (not linearly) . Thus it is normal that maximums have tendency to increase.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each N_1. Moreover α_5 is the selected percentage points with probability 0,95 of chi squared statistics with one degree of freedom.

N_1	$57*10^6$	$57*10^5$	$57*10^4$	$57*10^3$	5700	570	100
$Max(\chi^2_{N_1})$	0.321	0.789	1.874	3.456	1.987	1.54	4.012
α_5	3.84	3.84	3.84	3.84	3.84	3.84	3.84

Now we test the sequence x(n). We use partitions in $D(N_1)+1$ equal intervals : $D(N_1)$ is a function of N_1.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each N_1 (then, $D(N_1)$ is constant). Moreover α_5 is the selected percentage points with probability 0,95 of chi squared statistics.

N_1	5700	570	100
$Max(\chi^2_{N_1})$	40.851	18.21	17.32
α_5	43.77	16.92	16.92

if N_1 is larger, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \xrightarrow{D} X^*_G$ where $X^*_G \sim N(0,1)$: cf proposition A.1.2. Moreover α_5 is the selected percentage points with probability 0,95 of G.

N_1	$57*10^6$	$57*10^5$	$57*10^4$	$57*10^3$
$Max(\lvert N_G\rvert)$	0.754	1.506	2.033	2.114
α_N	1.960	1.960	1.960	1.960

Serial Test We test the sequence $b^1(n')$. We use the chi squared test with the partitions of $\{0,1\}^p$: there are 2^p partitions. Here we denote by $\chi^2_{N_1}$ the chi-squared statistics with $2^p - 1$ degrees of freedom.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each p. Then, for all p, we use several values of N_1.

p	2	3	4	5
$max(\chi^2_{N_1})$	3.431	8.801	19.823	38.456
α_5	7.815	14.07	25.00	36.41

If p is larger, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \sim N(0,1)$.

p	10	15	20
$max(\lvert N_G \rvert)$	1.952	1.874	2.074
α_N	1.960	1.960	1.960

Now, we test the sequence $x(n)$. We use partitions in $D(N_1) + 1$ equal hypercubes where $D(N_1)$ is function of N_1. Because $D(N_1)$ is big, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each p.

p	2	3	4	5	10	15	20
$max(\lvert N_G \rvert)$	2.221	1.947	2.004	2.121	1.745	2.2144	2.002
α_N	1.960	1.960	1.960	1.960	1.960	1.960	1.960

Gap Test One test the sequence x(n). We keep the notations of [1] page 62. We choose $[\alpha, \beta[= [0, 1/2[$. Then, one uses the chi squared statistics with t degrees of freedom (with the notations de [1]) : we denote them by $\chi^2_{N_1}$.

One chooses $t = 5, 6, 7, 8, 9, 10$. We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each t.

t	5	6	7	8	9	10
$Max(\chi^2_{N_1})$	12.45	15.04	19.02	20.31	20.99	19.11
α_5	11.07	12.59	14.07	15.51	16.92	18.31

For this test, we took many different samples. It is not surprising that maximum are close to α_5.

Poker Test One tests the sequence x(n). We keep the notations of [1] page 63. Then, one uses also chi squared statistics : we denote them by $\chi^2_{N_1}$. Then, $\chi^2_{N_1}$. is used for testing the number of k-tuples. We choose $k = 5, 6, 7, 8, 9, 10, 11, 12$. We lump a few categories of low probability together.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each k.

k	5	6	7	8	9	10	11	12
$Max(\chi^2_{N_1})$	7.89	10.84	13.25	12.15	13.45	20.75	13.64	16.82
α_5	7.81	9.49	11.07	12.59	12.59	14.07	15.51	16.92

Coupon collector's Test One tests the sequence x(n). We keep the notations of [1] page 64. We choose d=3,4,5,6,7,8 (with the notations of [1]). Then, one uses also chi squared statistics : we denote them by $\chi^2_{N_1}$. We use various t (with the notations of [1]). We choose t as a function of d. We lump a few categories of low probability together.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each d.

d	3	4	5	6	7	8
$max(\chi^2_{N_1})$	11.02	13.59	18.08	16.73	20.08	22.84
α_5	11.07	12.59	14.07	15.51	16.92	18.31

Permutation Test One tests the sequence x(n). We keep the notations of [1] page 65. We choose t=3,4,5,6,7 (with the notations of [1]). Then, one uses also chi squared statistics with t!-1 degrees of freedom i.e. 5, 23, 119, 719, 5039 degrees of freedom. We denote them by $\chi^2_{N_1}$.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for t=3,4.

t	3	4
$Max(\chi^2_{N_1})$	12.67	40.8
α_N	11.07	35.17

If t is bigger, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $|N_G|$ obtained for each t.

t	5	6	7
$max(\lvert N_G \rvert)$	2.191	2.432	1.891
α_N	1.960	1.960	1.960

Run Test One tests the sequence x(n). We keep the notations of [1] page 66. We uses the chi squared statistics with 6 degrees of freedom. We denote it again by $\chi^2_{N_1}$.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each N_1.

N_1	$57*10^6$	$57*10^5$	$57*10^4$	$57*10^3$	$57*10^2$
$max(\chi^2_{N_1})$	5.710	8.848	10.640	12.312	12.584
α_N	12.59	12.59	12.59	12.59	12.59

Maximum-of-t Test One tests the sequence x(n). We keep the notations of [1] page 70. We test the distribution function of maximum V_n. Then, one uses V^t_N : it is enough to apply equidistribution test with partitions of size D+ 1. Then, one has chi squared statistics with D degrees of freedom.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$ when the degree of freedom depends on N_1 : $D(N_1)$.

We use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $|N_G|$ obtained for each N_1.

N_1	$57*10^6$	$57*10^5$	$57*10^4$	$57*10^3$	$57*10^2$
$Max(\|N_G\|)$	0.431	1.554	2.025	1.905	2.102
α_N	1.960	1.960	1.960	1.960	1.960

Collision Test One tests the sequence $b^1(n')$. We keep the notations of [1] page 70. We do the test with samples which have a size of 2^{14} in 2^{20} urns as in [1] . Then, one can use the tables of probabilities of [1] page 70.

Let co(t) be the numbers of collision of t-th test. Then, we have the following inequalities for t=1,2,...,1000

$$109 \leq co(t) \leq 141 \ .$$

We recall that $P\{co(t) \leq 108\} = 0.043$ and $P\{co(t) \geq 145\} = 0.946$: cf [1] page 70.

Birthday spacings Test One tests the sequence x(n). We keep the notations of [1] page 71. We use the test exacly as in [1]. Then, we use the chi squared statistics with 3 degrees of freedom. We denote it by χ^2_3.

The following table is that of χ^2_3.

χ^2_3	3.66	4.02	1.94	5.67	3.32	6.77	7.01	2.78	4.25	6.11
α_5	7.81	7.81	7.81	7.81	7.81	7.81	7.81	7.81	7.81	7.81

As a matter of fact, we made this test more than hundred times. For these other tests, similar results have been obtained.

Serial Correlation Test One tests the sequence x(n). We keep the notations of [1] page 72. One chooses $N_1 \geq 100$. Let C be the serial correlation coefficient. Then, $\alpha_5 \approx 2$ when one tests $|\sqrt{N_1}C|$.

The following table is that of the maxima of $|\sqrt{N_1}C|$obtained for each N_1.

.

N_1	$57 * 10^6$	$57 * 10^5$	$57 * 10^4$	$57 * 10^3$	$57 * 10^2$
$max\|\sqrt{N_1}C\|$	0.31	1.11	1.51	1.86	1.55
α_5	2	2	2	2	2

Higher order correlation coefficient Test One tests the sequence x(n). We keep the notations of [10] page 72. In this test, we test not only the linear correlation coefficient : we test also the polynomial correlation coefficients.

For this test one can use samples $(x(2n), x(2n+1))$. But - as for the serial correlation coefficient - results are similar if one uses samples $(x(n), x(n+1))$.

Therefore one uses the statistics $||S_{N_1}||^2$ defined in theorem 6-11 of [10] : $||S_{N_1}||^2$ has asymptotically a chi squared distribution with h^2 degrees of freedom (here h=k with the notations of [10]). One chooses h according to N_1.

Because h^2 is big, we use $N_G = \sqrt{2\chi_2^d} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $|N_G|$ obtained for each N_1.

$\|N_1\|$	$57 * 10^6$	$57 * 10^5$	$57 * 10^4$	$57 * 10^3$	$57 * 10^2$
$Max(\|N_G\|)$	0.541	1.121	1.584	1.659	1.832
α_N	1.960	1.960	1.960	1.960	1.960

Conclusion Every test made conclude that the x(n) and $b^1(n')$ are IID sequence. It is only confirming our previous study. It brings a supplementary proof to the correctness of our reasoning.

11.3 Certainty of the randomness of $\{b^1(n')\}$

Now, one wants to show that there is no doubt about the randomness of the sequence $b^1(n')$: all was proven.

11.3.1 Detail of the method

Our method is based on different steps.

1) We use data a_j which can be regarded like the realization $a_j = A_j(\omega)$ of a random sequence A_j having a certain asymptotic independence.

2) We transform them into a sequence $e^2(j)$ having marginal laws rather well distributed thanks to the addition of a pseudo-random generator.

3) Thanks to the Fibonacci function we transform the $e^2(j)$ into a sequence $e^3(j)$ almost independent and with marginal distributions extremely near to the uniform distribution.

The advantage, it is that **almost all** the logical models $E^3(j)$ of the sequence $e^3(j)$ are almost independent and have marginal distribution extremely near to the uniform distribution.

4) Then, one uses the XORLT. One makes the sums modulo m of the S=10 lines of the matrix of the f(i,n) : $h(n) = \overline{\sum_i f(i,n)}$.

By the sections 5.7 and 8.5.2, the advantage of using the XORLT, it is that the conditional probabilities satisfy

$$P\Big\{\overline{\sum_i F(i,n)} \in I \;\Big|\; F(i,n+j_s) = f_{i,s}, s = 2,3,..\Big\} \approx P\Big\{\overline{\sum_i F(i,n)} \in I\Big\} \; .$$

It is even truer for

$$P\Big\{\overline{\sum_i F(i,n)} \in I \;\Big|\; \overline{\sum_i F(i,n+j_s)} = h_{i,s}, s = 2,3,...\Big\} \; .$$

However, the law of each F(i,n) is extremely close to the uniform distribution.

Therefore, the conditional distribution of $\overline{\sum_i F(i,n)}$ given $\overline{\sum_i F(i,n+j_s)} = h_{i,s}, s = 2,3,...$ is also extremely close to the uniform distribution.

One can thus accept the model of

$$P\{\overline{\sum_i F(i,n)} = y \mid \overline{\sum_i F(i,n+j_s)} = h_{i,s}, s = 2,3...\} = \frac{1}{m}\Big[1 + u_y\Big] \; ,$$

where u_k is a sample of an IID sequence U_k.

5) One uses once again the function of Fibonacci T_q. That allows us to conclude that $P\{X(n) \in Bo \mid X(n+j_2) = x_2, ..., X(n+j_p) = x_p\} = L(Bo) + Ob(1)\epsilon$ for any Borel set Bo.

6) Then, one transforms the sequence X(n) into a sequence of bits $B^1(n')$ satisfying $P\{B^1(n') = b \mid B^1(n'+j_2) = bi_2, ..., B^1(n'+j_p) = bi_p\} = 1/2 + Ob(1)\epsilon$ for any bits b.

7) The condition imposed on the parameters, m and q_0 imply that $\epsilon = \frac{\alpha}{\sqrt{q_0 N}}$, where $q_0 N$ is the size of the sequence $B^1(n')$. This equality means that one will not be able to differentiate the sequence $B^1(n')$ from an IID sequence of bits.

11.3.2 Details of the certainty

The results on the various steps which we have just pointed out are logically certain. Indeed,

1) One is logically sure that the sequence $A(j)$ has some asymptotic independence. In particular, if one takes a dictionary or an encyclopaedia, the various definite words represent Qd-dependent events.

2) One showed by reasoning that the sequence $e^2(j)$ has marginal distributions enough well distributed : $e^2(j) = \overline{e^1(j) + rand_0(j)}$ where $rand_0$ is independent with respect to the behavior of $e^1(j)$.

The reasoning could even be useful in order to show that $E^2(j)$ has not only uniform marginal distributions, but is even IID. It is besides the reasoning of Marsaglia in the construction of its CD-ROM : cf chapter 3.

3) Thanks to the functions of Fibonacci we transform the $e^2(j)$ into a sequence $e^3(j)$ having marginal distributions extremely near to the uniform distribution for all the logical models $E^3(j)$. Moreover, these $E^3(j)$ are almost independent.

This result is proved mathematically. As a matter of fact the sequence $e^3(j)$ is maybe already an IID sequence : cf section 12.2.

4) Then the XORLT shows that

$$P\{\overline{\sum_i F(i,n)} \in I \mid F(i,n+j_s) = f_{i,s}, s = 2,3,..\} \approx L(I).$$

5) One can thus accept the model

$$P\{\overline{\sum_i F(i,n)} = y \mid \overline{\sum_i F(i,n+j_s)} = h_{i,s}, s = 2,3...\} = \frac{1}{m}\Big[1 + u_y\Big] \ ,$$

where u_k is a sample of an IID sequence of random variables U_k with mean 0.

Now according to the CLT, one knows that the curve of the probabilities of the sums has well the shape of a bell of the type

$$P\Big\{ \sum_i F(i,n) = y \ \Big| \ \sum_i F(i,n+j_s) = g_s, s = 2,3...\Big\} = \frac{Se^{-\frac{S(y-0.5)^2}{2\sigma^2}}}{m\sqrt{2\pi\sigma^2}}\Big[1 + v_y\Big] \ ,$$

where $m\sigma_V^2 = O(1)$ (σ_V^2 is the variance of V_y).

Therefore, it will be the same for σ_U^2. The supposed condition $\sigma_U^2 \leq 1$ is thus too weak. One is thus sure that it will be satisfied.

6) The results on the Fibonacci function show that

$$P\{B^1(n') = b \mid B^1(n'+j_2) = bi_2, ..., B^1(n'+j_p) = bi_p\} = 1/2 + Ob(1)\epsilon \ ,$$

where $\epsilon = \frac{\alpha}{\sqrt{q_0 N}}$.

7) Our study, as well as the traditional results on the samples, thus shows that one will not be able to differentiate the sequence $B^1(n')$ from a sequence of IID random bits.

8) As a matter of fact, we took precautions so much that ϵ is probably much smaller than $\frac{\alpha}{\sqrt{q_0 N}}$.

There is thus no uncertainty on the randomness of $b^1(n')$. All was proved.

Remark 11.3.1 *So that the sequence $b^1(n')$ is not representative of an IID sequence, there are only two solutions*
1) The computer has a defect. But in this case, any calculation also risks to be false.
2) By an extraordinary chance, it is the same case where a really IID sequence does not satisfy the necessary tests. But that can occur only with the same probability, i.e. negligible.

11.3.3 Presentation of the result

In the beginning, one has a sequence of data a(j). But the sequence $e^3(j)$ can also be regarded as a sequence of data. One can thus expound the result in the following way.

One considers the sequence of data $e^3(j)$.

One can thus say that, for all the logical models $E^3(j)$, associated to the sequence of data $e^3(j)$, the sequence $b^1(n')$ transformed of $e^3(j)$ according to the transformations described above, cannot be differentiated from an IID sequence.

It is always possible that a test of randomness is not checked. But that will happen only with a probability which has the same order as that of an IID sequence.

We thus have turns the difficulty of a mathematical definition of an IID sequence because our result is true for *all* the models $E^3(j)$, and thus for all the models $B^1(n)$ which are deduced from it.

Chapter 12

Other methods of building of IID sequences

12.1 Second method

In this section one uses the rate of convergence of the XORTL (cf 5.5). One does not apply it to a sequence of numbers of the same type that f(i,n), n=1,...,N, but to random numbers of size N, i.e. very large, for example with a sequence of bits of size 100.000.000 regarded as a number with values in $\{0, 1, ..., 2^{100.000.000} - 1\}$.

This technique makes it possible to avoid using conditional probabilities.

12.1.1 Method of construction of the sequence

One keeps a part of the steps of the method used in the chapter 11.

1) One supposes that one has a sequence of data which one writes in the forms of bits $a_j \in \{0, 1, ..., \kappa - 1\}$, j=1,.....,N. One supposes that one can write $a_j = A_j(\omega)$. It is supposed made up of several independent files.

2) One rewrites it in the form of a table c^i_j, $j = 1, ..., J_i$, i=1,2,..,S. The J_i's are chosen in order that the lines are made up with independent files.

3) For each line i, one chooses a congruence of Fibonnacci $T_i(x) = a_i x$ modulo m_i so that N/m_i is very small. We assume also that there exists $\gamma_i \in \mathbb{N}$ such that $m_i = m^F(\kappa^{\gamma_i})$.

One groups the data togheter by set of γ_i elements :
$d_i(j) = \overline{c^i_{J_i-(s-1)\gamma_i}, c^i_{J_i-(s-1)\gamma_i-1}....c^i_{J_i-(s-1)\gamma_i-(\gamma_i-1)}}$, $s = 1, 2, ..., J'_i$ ou $J'_i = \lfloor J_i/\gamma_i \rfloor$.

4) For all $j = 1, 2, ..., J'_i$, one transforms each $d_i(j)$ by Fibonacci functions T_{q_i} where $q_i \leq 3\gamma_i/4$.

4-a) One defines $e^1_i(j) \in \{0, 1, ..., m_i - 1\} : e^1_i(j) = \lfloor d_i(j) m_i / 2^{\gamma_i} \rfloor$.

4-b) One sets $e_i^2(j) = \overline{e_i^1(j) + rand_i(j)}$ modulo m_i, where $rand_i(j)$ is a pseudo-random generator of period m_i with values in $F^*(m_i)$.

4-c) One defines $e_i^3(j) \in \{0, 1, ..., 2^{q_i} - 1\}$ by : $e_i^3(j) = 2^{q_i} T_{q_i}(e_i^2(j)/m_i)$.

5) One rewrites the $e_i^3(j)$ in the form of a sequence of bits.

5-a) For $j = 1, ..., J_i'$, let $e_i^3(j) = \overline{\overline{bi_1^i(j) bi_2^i(j) .... bi_{q_i}^i(j)}}$ be the writing of the $e_i^3(j)$ base 2.

5-b) One defines the sequence of bits $bt_i(n)$, $n = 1, 2, ..., q_i J_i'$, by setting $bt_i(q_i s - u + 1) = bi_u^i(s)$ for $s = 1, ..., J_i'$, and $u = 1, ..., q_i$.

6) One modifies the lines $bt_i(n')$, $n' = 1, 2, ..., q_i J_i'$, thanks to transformations having a behavior close to that of the permutations. In this aim, one uses other sequences of datas $c_i^1(n') \in \{1, 2, ..., q_i J_i'\}$, $n' = 1, 2, ..., q_i J_i'$, where $i = 1, 2, ..., 3S$. Because we use transformations similar to permutations we set $c_i^1(n') = Perm_i(n')$ in order that the notations are clearer.

6-a) One groups them togheter by package of three successive sequences $Perm_i^t(n')$ for t=1,2,3, i=1,2,...,S, $n' = 1, 2, ..., q_i J_i'$.

6-b) For each line i, for $n' = 1, 2, ..., q_i J_i'$, one sets, $r_0^i(n') = bt_i(n')$ and, for each t=1,2,3, $r_t^i(n') = bt_i(Perm_t^i(n'))$ for $n' = 1, 2, ..., q_i J_i'$.

6-c) For each line i, we set $r_i(n') = \overline{r_0^i(n') + r_1^i(n') + r_2^i(n') + r_3^i(n')}$ modulo 2 for $n = 1, 2, ..., q_i J_i'$.

Remark 12.1.1 *Less large sequences $c_i^1(s)$ can also be used in order to define these transformations. For example, one can suppose i=1,2,...,S : then, we choose three lines i_1, i_2, i_3 which we associate at each line i in the step 6-a).*

7) One definite g_i as the number with J bits whose writing bases 2 is $g_i = \overline{\overline{r_i(1) r_i(2) .... r_i(J)}}$, ou $J = min_i(q_i J_i')$.

The numbers g_i are thus number of very large size.

8) We set $h = \sum_{i=1}^{S} g_i$. For calculation one uses the method described in section 12.1.2.

9) We set $k = \overline{h}$, mod $M_2 + 1$ where $M_2 = 2^J - 1 : 0 \leq k \leq M_2$.

10) Let $k = \overline{\overline{b_1, b_2, ...}}$, the writing of k base 2. Then, one regards the sequence $b_1, b_2, ..., b_J$.

It is the sequence of bits b_s which will check the properties that we want for our IID sequence.

12.1.2 Calculation algorithms

As we make calculations on extremely large numbers, one cannot directly use the various mathematical programming systems: e.g. if $g_i \in \{0, 1, ..., 10^{1.000.000}\}$. We thus clarify now the method of calculating $h = \sum_{i=1}^{S} g_i$ when S=15.

It is known that g_i represents a number with J digits in base 2 : $g_i = \overline{r_i(1)r_i(2)....r_i(J)}$. Therefore, $g_i = \sum_{j=1}^{J} r_i(j)2^{J-j}$ where $r_i(j) \in \{0, 1\}$.

Then, $g_i = \sum_{n=0}^{J-1} r_i(J-n)2^n$. Therefore $h = \sum_{n\geq 0} \sum_i r_i(J-n)2^n = \sum_{n\geq 0} z_n^1 2^n$ where $z_n^1 = \sum_{i=1}^{15} r_i(J-n)$. Therefore, $z_n^1 \leq 15$.

We set $z_n^1 = 0$ if $n < 0$. Two algorithms A1 and A2 thus will transform a sequence z_n^1 corresponding to $h = \sum_{n\in\mathbb{Z}} z_n^1 2^n$, $z_n^1 \leq 15$ into a sequence z_n^k corresponding to $h = \sum_{n\in\mathbb{Z}} z_n^k 2^n$ where $z_n^k \in \{0, 1\}$.

Algorithm A1

Algorithm A1 is composed of several successive steps.

Transformation of z_n^1 into z_n^2 For any n, one writes z_n^1 in base 2. Because, $z_n^1 \leq 15$, for all $n \in \mathbb{Z}$, $z_n^1 = \rho_{1,n}^4 2^3 + \rho_{1,n}^3 2^2 + \rho_{1,n}^2 2 + \rho_{1,n}^1$ ou $\rho_{1,n}^j \in \{0, 1\}$ for j=1,2,3,4.

Therefore,

$$h = \sum_{n\in\mathbb{Z}} z_n^1 2^n = \sum_{n\in\mathbb{Z}} (\rho_{1,n}^4 2^3 + \rho_{1,n}^3 2^2 + \rho_{1,n}^2 2 + \rho_{1,n}^1) 2^n$$

$$= \sum_{n\in\mathbb{Z}} \rho_{1,n}^4 2^{n+3} + \sum_{n\in\mathbb{Z}} \rho_{1,n}^3 2^{n+2} + \sum_{n\in\mathbb{Z}} \rho_{1,n}^2 2^{n+1} + \sum_{n\in\mathbb{Z}} \rho_{1,n}^1 2^n$$

$$= \sum_{n\in\mathbb{Z}} z_n^2 2^n,$$

where $z_n^2 = \rho_{1,n}^1 + \rho_{1,n-1}^2 + \rho_{1,n-2}^3 + \rho_{1,n-3}^4$. Therefore, $z_n^2 \leq 4$.

Transformation of z_n^2 into z_n^3 Because $z_n^2 \leq 4$, we set, for $n \in \mathbb{Z}$, $z_n^2 = \rho_{2,n}^3 2^2 + \rho_{2,n}^2 2 + \rho_{2,n}^1$.

Therefore,

$$h = \sum_{n\in\mathbb{Z}} z_n^2 2^n = \sum_{n\in\mathbb{Z}} (\rho_{2,n}^3 2^2 + \rho_{2,n}^2 2 + \rho_{2,n}^1) 2^n$$

$$= \sum_{n\in\mathbb{Z}} \rho^3_{2,n} 2^{n+2} + \sum_{n\in\mathbb{Z}} \rho^2_{2,n} 2^{n+1} + \sum_{n\in\mathbb{Z}} \rho^1_{2,n} 2^n = \sum_{n\in\mathbb{Z}} z^3_n 2^n,$$

where $z^3_n = \rho^1_{2,n} + \rho^2_{2,n-1} + \rho^3_{2,n-2}$. Therefore, $z^3_n \leq 3$.

Transformation of z^3_n into z^4_n Because $z^3_n \leq 3$, we set, for $n \in \mathbb{Z}$, $z^3_n = \rho^2_{3,n} 2 + \rho^1_{3,n}$ with $\rho^2_{3,n} \leq 1$.

Then,

$$h = \sum_{n\in\mathbb{Z}} z^3_n 2^n = \sum_{n\in\mathbb{Z}} (\rho^2_{3,n} 2 + \rho^1_{3,n}) 2^n = \sum_{n\in\mathbb{Z}} \rho^2_{3,n} 2^{n+1} + \sum_{n\in\mathbb{Z}} \rho^1_{3,n} 2^n$$

$$= \sum_{n\in\mathbb{Z}} z^4_n 2^n,$$

where $z^4_n = \rho^1_{3,n} + \rho^2_{3,n-1}$.

Because $z^4_n \leq 2$, we set for $n \in \mathbb{Z}$, $z^4_n = \rho^2_{4,n} 2 + \rho^1_{4,n}$ with $\rho^2_{4,n} \leq 1$.

The first algorithm is finished in this way. Indeed, by continuing the same method, the following step transforms z^4_n into a sequence z^5_n which checks the *same* conditions : $z^5_n \leq 2$.

Algorithm A2

Then, A2 algorithm transforms a sequence z^s_n into a sequence z^{s+1}_n which checks the same conditions.

Then, one supposes $z^s_n \leq 2$.

Now, if $h = \sum_{n\in\mathbb{Z}} z^s_n 2^n$, with $z^s_n = \rho^2_{s,n} 2 + \rho^1_{s,n}$, $\rho^2_{s,n} \leq 1$, $z^s_n \in \{0, 1, 2\}$, $z^s_n = 0$ if $n < 0$, then

$$h = \sum_{n\in\mathbb{Z}} z^s_n 2^n = \sum_{n\in\mathbb{Z}} (\rho^2_{s,n} 2 + \rho^1_{s,n}) 2^n = \sum_{n\in\mathbb{Z}} \rho^2_{s,n} 2^{n+1} + \sum_{n\in\mathbb{Z}} \rho^1_{s,n} 2^n$$

$$= \sum_{n\in\mathbb{Z}} z^{s+1}_n 2^n,$$

where $z^{s+1}_n = \rho^2_{s,n-1} + \rho^1_{s,n}$

Therefore, $z^{s+1}_n = \rho^2_{s+1,n} 2 + \rho^1_{s+1,n}$ with $\rho^2_{s+1,n} \leq 1$, $z^{s+1}_n \in \{0, 1, 2\}$, $z^{s+1}_n = 0$ if $n < 0$.

Therefore one did not change conditions on z^s_n when one transforms "s" into "s+1" by A2 algorithm

Finally, A2 algorithm transforms any sequence $z_n^s \in \{0,1,2\}$, $z_n^s = \rho_{s,n}^2 2 + \rho_{s,n}^1$, $\rho_{s,n}^2 \leq 1$, into a sequence $\{z_n^{s+1}\} = A2(\{z_n^s\}) \in \{0,1,2\}$, $z_n^{s+1} = \rho_{s,n-1}^2 + \rho_{s,n}^1$.

However, the number of 2 tends to decrease. One then will repeat a certain number of times the A2 algorithm in order to make disappear the 2 from the sequence z_n^s.

Indeed, in order that $z_n^{s+1} = 2$, it is necessary and sufficient that $\rho_{s,n}^1 = \rho_{s,n-1}^2 = 1$.

Finally, after each step which transforms z_n^s into z_n^{s+1} the new sequence z_n^{s+1} has less from 2 (or the same number) that the sequence z_n^s. One can understand that on the following examples.

Example 1 We suppose z = 0 1 1 1 2 0 1

Then,

z = 0 1 1 1 2 0 1

A2(z)
= 0 1 1 1 0 0 1
+ 0 0 0 1 0 0 0

= 0 1 1 2 0 0 1

$A2^2(z)$
= 0 1 1 0 0 0 1
+ 0 0 1 0 0 0 0

= 0 1 2 0 0 0 1

$A2^3(z)$
= 0 1 0 0 0 0 1
+ 0 1 0 0 0 0 0

= 0 2 0 0 0 0 1

$A2^4(z)$

= 0 1 0 0 0 0 1
+ 1 0 0 0 0 0 0

= 1 1 0 0 0 0 1

Example 2 We suppose z= 0 1 1 0 2 0 1

Then,

z=0 1 1 0 2 0 1

A2(z)
= 0 1 1 0 0 0 1
+ 0 0 0 1 0 0 0

= 0 1 1 1 0 0 1

Exemple 3 We suppose z= 0 1 1 2 2 0 1

Then,

z=0 1 1 2 2 0 1

$A2(z)$
= 0 1 1 0 0 0 1
+ 0 0 1 1 0 0 0

= 0 1 2 1 0 0 1

$A2^2(z)$
= 0 1 0 1 0 0 1
+ 0 1 0 0 0 0 0

= 0 2 0 1 0 0 1

$A3^4(z)$
= 0 0 0 1 0 0 1
+ 1 0 0 0 0 0 0

$= 1\ 0\ 0\ 1\ 0\ 0\ 1$

As a matter of fact, the existence of a "2" after the step "s" is still checked in the step "s+1" if this "2" is preceded by a "1".

If it is preceded by a "0" or by a "2" it is removed.

Therefore, if there is still "2" in the step "s" that means that there was several "1" which are consecutive in step 0 (followed by a "2").

Because the probability that n "1" are consecutive decreases if n increases, there will be normally no "2" in sequence $z_n^{s_0}$ after s_0 steps.

Method of calculating of b_i

The calculation of $b_1, ..., b_q$, is very simple. Let $h = \overline{h_{J+r}h_{J+r-1}.....h_2h_1}$ (where $r \geq 0$) the binary writing of h . The bits b_s are the bits h_s except the r first bits $h_{J+r}, h_{J+r-1}, .., h_{J+1}$ which one removes : $b_1, b_2, ..., b_J = h_J h_{J-1}.....h_1$.

12.1.3 Properties

One uses the properties on the rate of convergence of the XORLT. For example, one reminds the following proposition (cf section 5.5.1).

Proposition 12.1.1 *One uses the set of the probabilities on the* $(x_1, ..., x_S) \in \{0, 1, ..., 2^J - 1\}^S$ *defined as in section 5.5, hypothesis 5.5.1 :* $P_{x_1,....,x_S} = \frac{P'_{x_1,....,x_S}}{\sum P'_{x_1,....,x_S}}$ *provided with the uniform probability on the* $P'_{x_1,.,x_S}$.

One supposes that with these notations, and our previous notations (section 12.1.1, step 7), $X_i = G_i$.

Then, for all $y \in \{0, 1, ..., 2^J - 1\}$, *with a probability equal to* $1 - 2\Gamma(b)$,

$$P\{K = y\} \approx (1/2^J)\Big[1 + \frac{bOb(1)}{\sqrt{3}\sqrt{2^{J(S-1)}}}\Big] \ .$$

Proof Under our assumptions, the N of the assumptions 5.5.1 is equal to the 2^J defined in section 12.1.1 : $N = 2^J$.

Moreover $E_{P'} = 1/2$ and $\sigma^2_{P'} = 1/12$.

Finally, $\overline{x_1 + ... + x_S} = \overline{g_1 + ... + g_S} = k$. ■

One can thus obtain probabilities very close to the uniform distribution with a probability very close to 1. Indeed,

$$P\{K = y\} = \sum_{\overline{x_1+...+x_S}=y} p_{x_1,...,x_S} \ ,$$

$$P\{K=y\}=P\Big\{\{B_1=b_1\}\cap\{B_2=b_2\}\cap...\cap\{B_{2^J}=b_{2^J}\}\Big\}\ .$$

Then, if b^0 is large, with a probability infinitely close to 1,

$$P\Big\{\{B_1=b_1\}\cap\{B_2=b_2\}\cap...\cap\{B_{2^J}=b_{2^J}\}\Big\}=(1/2^J)\Big[1+\frac{b^0Ob(1)}{\sqrt{3}\sqrt{2^{J(S-1)}}}\Big]\ .$$

12.1.4 Study of datas

Then, we have a sample of S lines $d_i(1),...,d_i(J_i')$ where

$$d_i(j)=\overline{\overline{c^i_{J_i-(s-1)\gamma_i},c^i_{J_i-(s-1)\gamma_i-1}....c^i_{J_i-(s-1)\gamma_i-(\gamma_i-1)}}}\ ,$$

which is transformed in

$$g_i=\overline{\overline{r_i(1)r_i(2)....r_i(J)}}.$$

Concentration of the G_i's

The problem most difficult to solve logically, it is that the G_i's should not have a probability which is cancelled in many points : cf section 5.5.10.

First, to have an idea of the effects of the transformations of the data, we suppose that we did not make a transformation on the d_i and that we summon them directly, i.e. $g_i=\overline{\overline{r_i(1)r_i(2)....r_i(J)}}$, J=10.000.000, means a sum of text.

Then, it is important that these lines does not come from the same type of files, for example from English books. Indeed, if a line comes from an English text, that is not of consequence about the CLT: this text is only a sample of size 1 chosen randomly.

But if all the lines are obtained from English texts, that will involve logically that the probability associated to each line is very particular. Thus any line which does not correspond to an English text will have a null probability theoretically and there will be many of such sequences. Thus, the measure associated to the English texts is a priori concentrated in a small number of points on the set of the possible points $\overline{\overline{r_i(1)r_i(2)....r_i(J)}}$. There is $2^{10.000.000}$ possible points, but there are much less possible English texts than one can write with 10.000.000 of bits.

Therefore, there are many points γ_s which have a null or negligible probability: $P\{G_i=\gamma_s\}=0$.

Now, it was shown that the curve of the probabilities obtained by the CLT has the shape of bell, except in certain cases where the measure is concentrated nearly a small number of points.

In the case which interests us here, one can regard as example, the case where, for any i, $\overline{\overline{r_i(1)r_i(2)....r_i(J)}} = \overline{\overline{r_i(1)r_i(2)....r_i(P)0.0....0.0}}$ where P=1.000.000 and J=2.000.000. It is understood indeed that $h = \overline{\overline{h_1h_2.....h_P0.0.....0}}$. Therefore, in fact $b_{P+1}b_{P+2}.....b_q = 0, 0, ..., 0$.

If one directly applies the method above without modifying the $d_i(j)$, one would not be sure to have a curve of the probabilities in the shape of bell. The quality of the result would be thus dubious.

Of course, $g_i = \overline{\overline{r_i(1)r_i(2)....r_i(P)0.0....0.0}}$ is an extreme case. But it provides an example and shows that it is necessary to avoid being in a case of this kind.

Because of that, we made uniform the $d_i(J)$ and then we employed transformations of a type similar to the permutations.

Remark 12.1.2 *One cannot directly add pseudo-random generators as in the chapter 11. In the chapter 11, one used the numbers with 10 or 20 digits. Each provided number could be regarded as random because one can easily take a germ with 20 digits. Here, it would be necessary to use a generator with values in the numbers which have a order of* $2^{10.000.000}$ *if* $J = 10.000.000$ *for example. In no case, it could not to be random.*

12.1.5 Study of sums of text

First, let us consider a sequence resulting from texts or mathematical texts or programs. Let us suppose that for any line $\ell^i = \overline{\overline{\ell^i_1, ..., \ell^i_J}}$, $J = 400.000$ where ℓ means a letter of the Latin alphabet : $\ell \in \{0, 1, .., 25\}$.

If one uses all the possible sequences of letters ℓ^i , there is $26^{400.000}$ possible combinations.

Of course, there are much less texts possible in English. How much is there of them? It is rather difficult to evaluate correctly.

In the tests and the numerical results which we calculated, we obtained a Qd-dependence with Qd=22 : cf section 10.4. To avoid any error, let us increase Qd: let us consider that there is Qd-dependence with Qd=40.

Study of a particular case

To illustrate the ideas, let us group the letters by group of 40. There is 400.000/40= 10.000 of it. Let us take 1 group out of 2. There remain about 5000 of them.

Because of the 40-dependence, there would be thus more than $(40terms)^{5000}$ possible combinations.

Let us suppose that there are 20.000 words in the English language (it is an average: one can admit that there are 500.000 words).

But how much there are sequences of 40 terms in the English language?

To try to estimate it, let us suppose that one has words of 10 letters and that each word is independent. Then, there would be for each group of 4 words of 10 letters, $(10000)^4$ possible combinations.

But, one is not in this case. In fact, it seems that there would be at least $(20000)^3$ possible combinations for each group of 40 terms. But it is extremely difficult to be sure for it without making a long study.

Altogether, there would be thus $((20000)^3)^{5000} \approx 26^{56538}$ possible combinations.

That is much smaller than 26^{400000}, the number of all the possible combinations with 26 letters. There will be thus many l_s checking $Proba\{\ell^1 = l_s\} = 0$.

But these points of probability equal to 0 can disappear if several lines are summoned.

Thus let us suppose that we summon 15 lines by group of 40 terms.

Let $\ell_1^i(j)$ be the obtained lines with 5000 groups of 40 terms.

Let $h_1(j) = \sum_{i=1}^{15} \ell_1^i(j)$. One thus will group 15 groups of 40 independent terms togheter.

Then for each $h_1(j)$ associated with a group of 40 terms, there will be $(40terms)^{15} = (20000^3)^{15} = 20000^{45}$ possible combinations.

Thus altogether, there will be $((20000)^{45})^{5000} \approx (26^{45595})^{15} \approx 26^{683925}$ possible combinations: that would be more than the total number of possible combinations : $26^{400.000}$.

On the other hand, there will be even more possible combinations with the sum $h_1(j) = \overline{\sum_{i=1}^{15} \ell_1^i(j)}$.

It will be the same thing if one takes all the groups of 40 terms instead of removing 1 of them out of 2.

That thus means that by summoning 15 lines, one can remove the risk that $Proba\{h(j) = x\} = 0$.

Study of the assumptions of this particular case

Is what the assumptions which we have made are sufficiently close to reality?

The assumptions that we have made are approximate. For example: there is really 22-dependence? It is what numerical results show. Here we suppose that there are 40-dependence. But is this sufficient?

On the other hand, one can improve this type of results if one adds files of the different type: English texts, music, mathematical texts , computer programs.

The problem, it is that, in order to use this type of reasoning directly, it would have to be shown that for each line, there are about no points of probability equal to 0. Here, it is shown that the sum of the lines has this property.

But is this true? For example if all last the bits of the lines are equal to 0, that is not true.

General case

In fact, it is to avoid this kind of risk which we have initially transforms the data by the techniques of standardization and by transformations having similar properties to those of the permutations. But, all these results are to be confirmed.

In fact, these reasonings show that it would be maybe possible to summon the data directly. We improve the chances that it is possible by the various transformations carried out.

12.1.6 Permutations and associated transformations

Permutations

To avoid having points of concentrations of the probability, the simplest method, it is to make permutations on a certain number of sequence of c_j^i. Indeed, the permutations can transform into independent sequences about any sequence of real numbers (including the pseudo-random sequences) : cf section 8.1.2.

It is quite clear that, if permutations are made, that changes immediately the type of probability associated to each line d_i: for example, if one permutes $d_i \in \{\overline{c_1^i c_2^i .....c_J^i 0.....0}\}$.

Therefore, if the various lines d_i were summoned, there is not thus a priori reasons of considering that they are a sample of a random variable concentrated nearly a small number of points.

Choice of permutations

Like permutation, one can use for example the various Matlab permutations.

But, it poses a problem: they are not permutations taken randomly. In fact, one is in the case envisaged by Knuth ([1] : cf also definition 2.1.5) and which it is necessary to avoid. One would need permutations taken randomly to be sure that one can considers the probabilities as not concentrated nearly a small number of points.

For that, one can use nondeterministic sequences of data to define the permutations.

For example let us suppose that one wants to permute a sequence x(j) of size N and that one has data $d(j) \in \{0, 1, ..., m-1\}$, $m > N$. One would like to be able to define the permutation $P(j) = \lfloor N * d(j)/(m-1) \rfloor$. Unfortunately, there is no reason that P is injective.

One can try to remove the j, j' such that P(j)=P(j'), $j \neq j'$. But if N is large, that can be long.

It is easier to use these functions differently

Transformation having some characteristics of permutations

To define these transformations, nondeterministic sequences of data are used. In our construction defined in section 12.1.1, one used other sequences of datas $c_i^1(s) \in \{0, 1\}$.

First, we must define the numbers of which we have need to define these functions close to the permutations: we build them starting from the $c_i^1(s)$. We define these numbers in section 12.1.1, 6-a, 6-b, 6-c, 6-d. We obtain random sequences (not IID) $P_i^t(n)$ for t=1,2,3, i=1,2,...,S, $n = 1, 2, ..., q_i J_i'$

Then, one applies these transformations on each line i: we set $r_t^i(j) = bt_i(P_t^i(j))$.

Lastly, they are summoned modulo 2^{q_i} : $r_i(j) = \overline{r_0^i(j) + r_1^i(j) + r_2^i(j) + r_3^i(j)}$.

It is easy to note that this technique allows a mixture of the lines (and data) which is as random as it would be the choice of a permutation taken randomly. That is thus appropriate perfectly so that we can suppose that the probabilities of each line cannot be regarded as concentrated nearly a small number of points.

12.1.7 Question about sums

In section 5.5.8, one understood that to use proposition 5.5.1 involves in some cases absurd results for the sums of random variables. In the section 11.1.5, one understood that that does not pose a problem for the example studied in section 11.2. Is also the case for the sum of g_i used in section 12.1.1?

As a matter of fact, in the construction defined in section 12.1.1 there is obviously no problem.

Indeed, by using the method described in section 12.1.1, one makes directly an operation corresponding to a permutation. That thus amounts well taking probabilities randomly.

One can thus consider really that the probabilities that one uses in the sum behave as being taken randomly.

12.1.8 Example

By using, the technique defined in section 12.1.1, we have created a real sequence $b^2(n')$ which can be obtained by asking it to rene.blacher@imag.fr (Laboratory

LJK, University Joseph Fourier of Grenoble, France). More precisions on this subject will found in [18]

Detail of the construction

The following steps are carried out.

1) One has a sequence of digits in base 32 obtained as in the chapter 11. $a_j \in \{0, 1, ..., 31\}$, j=1,..., 34785384 : one thus has a sequence of 173926920 bits.

The data of the various lines were extracted from various texts in various languages: the Bible, dictionary, mathematical program, mathematical texts, encyclopaedia.

2) One rewrites it in the form of a table c_j^i, $j = 1, ..., J_i$, i=1,2,..,5 with $J_1 = 7.222.542$ $J_2 = 6.980.250$ $J_3 = 6.987.224$ $J_4 = 6.774.012$, $J_5 = 6.821.356$.

Each line consists of independent file

3) One chooses the congruence of Fibonnacci $T(x) \equiv ax$ modulo $m = m^F(32^{28})$ (it is the same one which in the chapter11).

One groups the data togeteher by unit of $\gamma = 28$ elements : $d_i(j) = \overline{\overline{c^i_{J_i-(s-1)\gamma}....c^i_{J_i-(s-1)\gamma-(\gamma-1)}}}$, $j = 1, 2, ..., J_i'$ where $J_i' = \lfloor J_i/\gamma \rfloor \geq \lfloor 6.774.012/28 \rfloor$ = 241.929. We set $J' = min_i(J_i') = 241.929$.

Then, we use only the matrix $d_i(j)$, j=1,2,...,J', i=1,2,3,4,5.

4) For all $j = 1, 2, ., J'$, one transforms each $d_i(j)$ by the function of Fibonnacci.

4-a) One defines $e_i^1(j) \in \{0, 1, ..., m-1\} : e_i^1(j) = \lfloor d_i(j)m/2^\gamma \rfloor$.

4-b) One sets $e_i^2(j) = \overline{e_i^1(j) + rand_i(j)}$ modulo m, where $rand_i(j)$ is a pseudo-random generator MATLAB of period m with values in $F^*(m)$.

4-c) One defines $e_i^3(j) \in \{0, ..., 2^q-1\}$, j=1,2,...,J', by : $e_i^3(j) = 2^q T_q(e_i^2(j)/m)$ where q=105 ($q = 21 * 5$, $21 = 3\gamma/4$) .

5) One rewrites the $e_i^3(j)$ in the form of a sequence of bits.

5-a) For $n = 1, ..., J'$, let $e_i^3(n) = \overline{\overline{bi_1^i(n)bi_2^i(n)....bi_q^i(n)}}$ be the writing of the $e_i^3(j)$ base 2.

6-b) One defines the sequence of bits $bt_i(n')$, $n' = 1, 2, ..., J$, J=qJ'=25402545, by setting $bt_i(qs - u + 1) = bi_u^i(s)$ for $s = 1, ..., J'$, and $u = 1, ..., q$.

6) One modifies the lines $bt_i(n')$, $n' = 1, 2, ..., J$, thanks to transformations having a behavior close to that of the permutations. In this aim, one uses other sequences of datas $c^1(n') \in \{1, 2, ..., J\}$, $n' = 1, 2, ..., J$ where $i = 1, 2, ..., 3S$. Because we use transformations similar to permutations we set $c^1(n') = Perm_i(n')$ in order that the notations are clearer.

6-a) One groups them togheter by package of three successive sequences $Perm_i^t(n')$ for t=1,2,3, i=1,2,...,S, $n' = 1, 2, ..., J$.

6-b) For each line i, for $n' = 1, 2, ..., J$, one sets, $r_0^i(n') = bt_i(n')$ and, for each t=1,2,3, $r_t^i(n') = bt_i(Perm_t^i(n'))$ for $n' = 1, 2, ..., J$.

6-c) For each line i, we set $r_i(n') = \overline{r_0^i(n') + r_1^i(n') + r_2^i(n') + r_3^i(n')}$ modulo 2 for $n = 1, 2, ..., J$.

7) For i=1,2,3,4,5, one defines g_i like being the number with 25402545 bits whose writing bases 2 of it is $g_i = \overline{\overline{r_i(1)r_i(2)....r_i(J)}}$.

8) We set $k = \overline{\sum_{i=1}^{5} g_i}$ modulo 2^J. Let $k = \overline{\overline{b^2(1)b^2(2)....b^2(J)}}$ the writing bases 2 of k . The sequence $b^2(n)$ is a sequence of random bits.

Programming

In some computation's systems, it can be long to execute the program of the transformations similar to permutations, $Perm_i(n) \in \{1, 2, ..., J\}$ when J is large (step 6). For example, if $J = 25402545$, the used memory is large. In Matlab 2008, the execution can be long.

In order to use not too memory, one can proceed step by step : one uses the $Perm_i(n')$ by packages : for example $Perm_i(n')$, n'= 508050(s-1)+1 : 508050*s. Now we show how processing with the first line.

1) We have a sequence of bits $bt(n') = bt_1(n')$, n'= 1,....,J, where $J = 25402545$.

2) We write it in a form of a sequence $Bt(v) \in \{0, 1, ..., 2^{50} - 1\}$, $v = 1, 2, ..., 508051$ where $Bt(v) = \overline{\overline{bi_{50}(v)....bi_2(v)bi_1(v)}}$ and $508051 = \lfloor J/50 \rfloor + 1$: if u-1=50v+s is the Euclidean division of u-1 by 50, $bt(u) = bi_{s+1}(v + 1)$.

3) We denote by $res(n') \in \{0, 1\}$, n'=1,2,...,J, the sequence which we want to obtain : at the end of computation $res(n') = bt(Perm(n'))$ for n'= 1,....,J. We write it in a form of a sequence $Res(v) \in \{0, 1, ..., 2^{50} - 1\}$, $v = 1, 2, ..., 508051$: $Res(v) = \overline{\overline{bit_{50}(v)....bit_2(v)bit_1(v)}}$ and $res(u) = bit_{s+1}(v+1)$, where u-1 =50v+s is the Euclidean division of u-1 by 50.

We execute the following steps :

A) At step "1", we set res = zeros(1,J) and Res=zeros(1,508051).

B) At step "s" :
for n= 508050(s-1)+1 : 508050*s,

1) Let $Perm(n) - 1 = 50j_0 + r_0$ be the Euclidean division of $Perm(n) - 1$ by 50.

2) We set $b = bi_{r_0+1}(j_0 + 1)$. Then $b = bt(Perm(n))$.
Indeed, for $j_0 = 0, 1, ..., 508050$, we use $Bt(j_0 + 1) = \overline{\overline{bi_{50}(j_0 + 1).....bi_1(j_0 + 1)}}$. We extract from it $b = bi_{r_0+1}(j_0 + 1)$ which is thus equal to $bt(Perm(n))$.

3) Let $n - 1 = 50j_2 + r_2$ be the Euclidean division of n-1 by 50.

4) We set $bit_{r_2+1}(j_2+1) = b$.
That means that one regards $Res(j_2+1) = \overline{\overline{bit_{50}(j_2+1).....bit_1(j_2+1)}}$. By replacing the old $bit_{r_2+1}(j_2+1) = 0$ by b, one obtains thus a new $Res(j_2+1) = \overline{\overline{bit_{50}(j_2+1).....bit_2(j_2+1)bit_1(j_2+1)}}$.

Then, we obtain at the end of step "s" a transformed sequence res(n') representing the sequence $bt(Perm(n'))$ obtained for n'= 1,2,...,508050.s. After, we will execute the following step.

In order to make computations at the step "s", one needs memory for
1) $Perm(n) \in \{1, 2, ..., J\}$, $n = 508050(s-1)+1 : 508050.s$,
2)$Bt(v) \in \{0, 1, ..., 2^{50}-1\}$, $v = 1, 2, ..., 508051$,
3) $Res(n') \in \{0, 1, ..., 2^{50}-1\}$, $n' = 1, 2, ..., 508051$.
Then, we do not need too memory if we use Matlab 2008.

Remark 12.1.3 *It is also possible to write the sequence Perm(n), n=1,....,J, in a form* $PErm(v) \in \{0, 1, ..., 2^{50}-1\}$, *v=1,..., 508051 as for bt(n') and Bt(v).*

Remark 12.1.4 *Here, we write the method to execute computations. We do not write the program. For example, in order to compute the step 4), the most simple way is to use* $Res(j_2+1) = Res(j_2+1) + b.2^{r_2}$.

An other method

Now the previous way of programming is a little long. Then, one can use a fast way : one can apply the transformations similar to permutations to $e_i^3(j)$. With this aim, one changes the method defined above in "Detail of the construction".

We replace steps 5) 6) 7) by the following way.

5) One modifies the lines $e_i^3(j) \in \{0, 1, ..., 2^q-1\}$, $j = 1, 2, ..., J'$, thanks to transformations having a behavior close to that of the permutations. In this aim, one uses other sequences of datas $Perm_i(n') \in \{1, 2, ..., J'\}$, $n = 1, 2, ..., J'$ where $i = 1, 2, ..., 3S$.
5-a) One groups them togheter by package of three successive sequences $Perm_i^t(n)$ for t=1,2,3, i=1,2,...,S, $n = 1, 2, ..., J'$.
5-b) For each line i, for $n = 1, 2, ..., J'$, one sets, $r_0^i(n) = e_i^3(n)$ and, for each t=1,2,3, $r_t^i(n) = e_i^3(Perm_t^i(n))$ for $n = 1, 2, ..., J'$.
5-c) For each line i, we set $r_i(n) = \overline{r_0^i(n) + r_1^i(n) + r_2^i(n) + r_3^i(n)}$ modulo 2^q for $n = 1, 2, ..., J'$.

6) One rewrites the $r_i(n)$ in the form of a sequence of bits.

6-a) For $n = 1, ..., J'$, let $r_i(n) = \overline{\overline{bi_1^i(n)bi_2^i(n)....bi_q^i(n)}}$ be the writing of the $r_i(n)$ base 2.

6-b) One defines the sequence of bits $bt_i(n')$, $n' = 1, 2, ..., J$, J=qJ'=25402545, by setting $bt_i(qs - u + 1) = bi_u^i(s)$ for $s = 1, ..., J'$, and $u = 1, ..., q$.

7) For i=1,2,3,4,5, one defines g_i like being the number with 25402545 bits whose writing bases 2 of it is $g_i = \overline{\overline{bt_i(1)bt_i(2)....bt_i(J)}}$.

Now, it will be necessary to know if this way is sufficient in order to have random bits $b^2(n)$ whose the randomness is sure.

12.1.9 Properties

First, let us notice the following remark.

Remark 12.1.5 *One can admit that the sequences $d_i(j)$ and $e_i^3(j)$ are 2-dependent : cf section 10.4. The $d_i(2j+1)$ thus form an independent sequence. As each $e_i^3(j)$ has a distribution close to the uniform distribution (cf chapter 8), that means that, at least, the sequence $d_i(2j+1)$ is IID.*

On the other hand, by using proposition 12.1.1 one finds that in the set of probabilities provided with the uniform distribution, with a probability infinitely close to 1,

$$P\Big\{\{B_1 = b_1\} \cap \{B_2 = b_2\} \cap ... \cap \{B_{2^J} = b_{2^J}\}\Big\} = (1/2^J)\Big[1 + \frac{Ob(1)}{2^{50.000.000}}\Big] .$$

It is a very strong approximation.

Now,

$$P\Big\{\{B_2 = b_2\} \cap ... \cap \{B_{2^J} = b_{2^J}\}\Big\}$$

$$= P\Big\{\{B_1 = 0\} \cap \{B_2 = b_2\} \cap ... \cap \{B_{2^J} = b_{2^J}\}\Big\}$$
$$+P\Big\{\{B_1 = 1\} \cap \{B_2 = b_2\} \cap ... \cap \{B_{2^J} = b_{2^J}\}\Big\}$$

$$= (1/2^{J-1})\Big[1 + \frac{Ob(1)}{2^{50.000.000}}\Big] .$$

More generally

$$P\Big\{\{B_n = b_1\} \cap \{B_{n+j_2} = b_2\} \cap ... \cap \{B_{n+j_p} = b_p\}\Big\} = (1/2^p)\Big[1 + \frac{Ob(1)}{2^{50.000.000}}\Big] \ .$$

Therefore,

$$P\Big\{B_n = b_1 \Big| B_{n+j_2} = b_2, ..., B_{n+j_p} = b_p\Big\}$$

$$= \frac{P\Big\{\{B_n = b_1\} \cap \{B_{n+j_2} = b_2\} \cap ... \cap \{B_{n+j_p} = b_p\}\Big\}}{P\Big\{\{B_{n+j_2} = b_2\} \cap ... \cap \{B_{n+j_p} = b_p\}\Big\}}$$

$$= \frac{(1/2^p)\Big[1 + \frac{Ob(1)}{2^{50.000.000}}\Big]}{(1/2^{p-1})\Big[1 + \frac{Ob(1)}{2^{50.000.000}}\Big]}$$

$$\approx (1/2)\Big[1 + \frac{2Ob(1)}{2^{50.000.000}}\Big] = (1/2) + \frac{Ob(1)}{2^{50.000.000}} \ .$$

Therefore,

$$P\Big\{B_n = b_1 \Big| B_{n+j_2} = b_2, ..., B_{n+j_p} = b_p\Big\} = (1/2^J) + Ob(1)\epsilon \ ,$$

where $\epsilon = \frac{Ob(1)}{2^{50.000.000}}$.

Now, we apply the theorem 9 with the Qd-dependence where $Qd \leq 22 * 5 = 110$ considering that $a_j \in \{0, 1, ..., 31\}$ and that $32 = 2^5$: cf section 11.2.9.

By the results of section 9.5.1, and by lemma 9.2.5,
$\beta_{1,p} \leq \frac{\sqrt{N_1}2p\epsilon}{\sqrt{A(p)2^p}}$ where $N_1 \leq 25402545$, $\epsilon \leq \frac{1}{2^{25.000.000}}$, $A(p) \geq (1/8)$. Then,

$\beta_{1,p} \leq \frac{\sqrt{8*25402545}}{2^{50.000.000}} \ \frac{2p}{2^{p/2}} = \frac{4*14255.538}{2^{50.000.000}} \leq \frac{2^{17}}{2^{50.000.000}} \leq \frac{1}{2^{49.999.983}}$.

Moreover, $\gamma_{1,p} \approx \frac{(p^2-p+1)\epsilon}{2A(p)}\Big[2p + (1 + 4Qd)\frac{4p}{2^p} + (1 + 2Qd)\frac{4p^2\epsilon}{2^p}\Big]$.

Now $2J^3 \approx 3.278398058033158 * 10^{22}$.
Then, $2J^3 < 2^{76} \approx 7.555786372591432 * 10^{22}$.
Then, $\gamma_{1,p} \leq J^3\epsilon < \frac{2^{75}}{2^{50.000.000}} = \frac{1}{2^{49.999.925}}$.
Finally,

$$\gamma_{1,p} \leq \frac{1}{2^{49.999.925}} \ .$$

Therefore, with the same notations as in section 11.2.10,

$$P\Big\{\sqrt{N_1}\big|P_e^B-(1/2)^p\big|\geq\sigma_B x\Big\}\leq\Gamma\Big[\Big(\frac{1-\beta_{1,p}/x}{1+\gamma_{1,p}}\Big)x\Big]\leq$$

$$\leq\Gamma\Big[\Big(\frac{1-\frac{1}{2^{49.998.983}}}{1+\frac{1}{2^{49.999.925}}}\Big)x\Big]\leq\Gamma\Big[\Big(1-\frac{1}{2^{49.998.982}}\Big)x\Big]\ ,$$

for $x\geq\frac{1}{2^{1000}}$.

With the notations of section 11.2.11, the same results are obtained for $P\Big\{\sqrt{N_1}\Big|\frac{P_e^B}{p_e^B}-(1/2)\Big|>\sigma_{cp}\ x\Big\}$.

It is quite clear that with such an approximation, nothing could differentiate such a sequence from an IID sequence if one has sample with size 26.000.000.

The question which is asked is: <<How can one obtain such a fine approximation?>>. The answer is that this is because the property of the XORLT and its speed of convergence. We understood that this asks a question in section 5.5.8, et 12.1.7.

It is necessary however not to forget that these results are true with a probability infinitely close to 1 in the set of the probabilities provided with the uniform distribution (as defined in proposition 12.1.1). Then, it asks the question to know if this measure on the set of the probabilities is quite selected. To arrive at *absoluetely sure* conclusions , that will require a still long study which we will make later on.

12.1.10 Tests

We are going to verify the previous conclusions by making tests of randomness. We use the classical Diehard tests cf [2], [1]. We tested the sequences $\{b^2(n')\}$ defined in the step 8 of section 12.1.8 and $\xi_n=\overline{0,b_1^n.....b_{50}^n}$. .

Results are in accordance with what we waited: for sequences $b^2(n)$ and ξ_n the hypothesis "randomness" can be accepted by all the Diehard tests.

We denote by N_1 the size of used samples. We denote by α_5 the selected percentage points with probability 0,95.

Equidistribution Test We use the chi-squared tests. First, we test $b^2(n)$ with the partition $\big\{\{0\},\{1\}\big\}$.

We use sample with size N_1. We use the chi squared statistics $\chi^2_{N_1}$. We take the maximum of $\chi^2_{N_1}$. The number of samples increases if N_1 decreases (not linearly) . Thus it is normal that maximums have tendency to increase.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each N_1. Moreover α_5 is the selected percentage points with probability 0,95 of chi squared statitics with one degree of freedom.

N_1	25402545	2540200	254020	25402	2500	250	100
$Max(\chi^2_{N_1})$	0.875	2.974	2.091	3.115	3.028	3.985	3.770
α_5	3.84	3.84	3.84	3.84	3.84	3.84	3.84

Now we test the sequence $\xi(n)$. We use partitions in $D(N_1)+1$ egal intervals : $D(N_1)$ is a function of N_1.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each N_1 (then, $D(N_1)$ is constant). Moreover α_5 is the selected percentage points with probability 0,95 of chi squared statistics.

N_1	2500	250	100
$Max(\chi^2_{N_1})$	41.828	15.01	15.25
α_5	43.77	16.92	16.92

if N_1 is bigger, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \xrightarrow{D} X^*_G$ where $X^*_G \sim N(0,1)$: cf proposition A.1.2. Moreover α_5 is the selected percentage points with probability 0,95 of G.

N_1	25402545	2540200	254020	25402
$Max(\lvert N_G\rvert)$	1.627	1.489	1.778	1.990
α_N	1.960	1.960	1.960	1.960

Serial Test We test the sequence $b^2(n)$. We use the chi squared test with the 2^p partitions of $\{0,1\}^p$. Here we denote by $\chi^2_{N_1}$ the chi-squared statistics with $2^p - 1$ degree of freedom.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each p. Then, for all p, we use several values of N_1.

p	2	3	4	5
$max(\chi^2_{N_1})$	5.287	9.984	25.765	36.704
α_5	7.815	14.07	25.00	36.41

If p is bigger, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \sim N(0,1)$.

p	10	15	20
$max(\lvert N_G\rvert)$	1.847	1.789	1.954
α_N	1.960	1.960	1.960

Now, we test the sequence $\xi(n)$. We use partitions in $D(N_1)+1$ equal hypercubes where $D(N_1)$ is function of N_1. Because $D(N_1)$ is big, we use $N_G = \sqrt{2\chi_2^d} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each p.

p	2	3	4	5	10	15	20
$max(\|N_G\|)$	1.874	1.801	1.647	2.210	1.521	2.010	1.984
α_N	1.960	1.960	1.960	1.960	1.960	1.960	1.960

Gap Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 62. We choose $[\alpha, \beta[= [0, 1/2[$. Then, one uses the chi squared statistics with t degrees of freedom (with the notations de [1]) : we denote them by $\chi^2_{N_1}$.

One chooses $t = 5, 6, 7, 8, 9, 10$. We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each t.

t	5	6	7	8	9	10
$Max(\chi^2_{N_1})$	10.21	13.27	14.02	17.84	17.73	20.02
α_5	11.07	12.59	14.07	15.51	16.92	18.31

For this test, we took many different samples. It is not surprising that maximum are close to percentages in 95 percent or in 99 percent.

Poker Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 63. Then, one uses also chi squared statistics : we denote them by $\chi^2_{N_1}$. Then, $\chi^2_{N_1}$. is used for testing the number of k-tuples. We choose $k = 5, 6, 7, 8, 9, 10, 11, 12$. We lump a few categories of low probability together.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each k.

k	5	6	7	8	9	10	11	12
$Max(\chi^2_{N_1})$	7.24	7.3	13.25	13.80	14.54	14.04	12.22	18.00
α_5	7.81	9.49	11.07	12.59	12.59	14.07	15.51	16.92

Coupon collector's Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 64. We choose d=3,4,5,6,7,8 (with the notations of [1]). Then, one uses also chi squared statistics : we denote them by $\chi^2_{N_1}$. We use various t (with the notations of [1]). We choose t as a function of d. We lump a few categories of low probability together.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each d.

d	3	4	5	6	7	8
$max(\chi^2_{N_1})$	12.74	12.27	14.88	19.90	19.72	20.92
α_5	11.07	12.59	14.07	15.51	16.92	18.31

Permutation Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 65. We choose t=3,4,5,6,7 (with the notations of [1]). Then, one uses chi squared statistics with t!-1 degrees of freedom i.e. 5, 23, 119, 719, 5039 degrees of freedom. We denote them by $\chi^2_{N_1}$.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each t=3,4.

t	3	4
$Max(\chi^2_{N_1})$	14.22	33.17
α_N	11.07	35.17

If t is bigger, we use $N_G = \sqrt{2\chi^d_2} - \sqrt{2d-1} \sim N(0,1)$. The following table is that of the maxima of $|N_G|$ obtained for each t.

t	5	6	7
$max(\|N_G\|)$	1.777	2.029	2.223
α_N	1.960	1.960	1.960

Run Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 66. We uses the chi squared statistics with 6 degrees of freedom. We denote it again by $\chi^2_{N_1}$.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$.

The following table is that of the maxima of $\chi^2_{N_1}$ obtained for each N_1.

N_1	25402545	2540200	254020	25402	2500
$max(\chi^2_{N_1})$	2.254	10.087	11.774	11.454	13.821
α_N	12.59	12.59	12.59	12.59	12.59

Maximum-of-t Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 70. We test the distribution function of maximum V_n. Then, one uses V_N^t : it is enough to apply equidistribution test with partitions of size D+ 1. Then, one has chi squared statistics with D degrees of freedom.

We use samples with various sizes N_1. We are interested in the maximum of these various $\chi^2_{N_1}$ when the degree of freedom depends on N_1 : $D(N_1)$.

We use $N_G = \sqrt{2\chi_2^d} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $|N_G|$ obtained for each N_1.

N_1	25402545	2540200	254020	25402	2500
$Max(\|N_G\|)$	1.115	1.602	1.754	1.994	1.813
α_N	1.960	1.960	1.960	1.960	1.960

Collision Test One tests the sequence $b^2(n)$. We keep the notations of [1] page 70. We do the test with samples which have a size of 2^{14} in 2^{20} urns as in [1] . Then, one can use the tables of probabilities of [1] page 70.

Let co(t) be the numbers of collision of t-th test. Then, we have the following inequalities for t=1,2,...,1000

$$110 \leq co(t) \leq 137 \ .$$

We remind that $P\{co(t) \leq 108\} = 0.043$ and $P\{co(t) \geq 145\} = 0.946$.

Birthday spacings Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 71. We use the test exacly as in [1]. Then, we uses the chi squared statistics with 3 degrees of freedom. We denote it by χ^2_3.

The following table is that of χ^2_3.

χ^2_3	5.31	6.77	5.87	1.07	7.11	3.01	2.25	4.21
α_5	7.81	7.81	7.81	7.81	7.81	7.81	7.81	7.81

As a matter of fact , we made this test more than hundred times. For these other tests, similar results have been obtained.

Serial Correlation Test One tests the sequence $\xi(n)$. We keep the notations of [1] page 72. One chooses $N_1 \geq 100$. Let C be the serial correlation coefficient. Then, $\alpha_5 \approx 2$ when one tests $|\sqrt{N_1}C|$.

The following table is that of the maxima of $|\sqrt{N_1}C|$obtained for each N_1.

N_1	25402545	2540200	254020	25402	2500
$max\|\sqrt{N_1}C\|$	1.27	1.13	1.62	1.94	1.77
α_5	2	2	2	2	2

Higher order correlation coefficient Test One tests the sequence $\xi(n)$. We keep the notations of [10] page 72. In this test, we test not only the linear correlation coefficient. We test also the polynomial correlation coefficients.

For this test one can use samples $(\xi(2n), \xi(2n+1))$. But - as for the serial correlation coefficient - results are similar if one uses samples $(\xi(n), \xi(n+1))$.

Therefore one uses the statistics $||S_{N_1}||^2$ defined in theorem 6-11 of [10] : $||S_{N_1}||^2$ has asymptotically a chi squared distribution with h^2 degrees of freedom (i.e. h=k with the notations of [10]). One chooses h according to N_1.

Because h^2 is big, we use $N_G = \sqrt{2\chi_2^d} - \sqrt{2d-1} \sim N(0,1)$.

The following table is that of the maxima of $|N_G|$ obtained for each N_1.

$\|N_1\|$	25402545	2540200	254020	25402	2500
$Max(\|N_G\|)$	0.541	1.121	1.584	1.659	1.832
α_N	1.960	1.960	1.960	1.960	1.960

Conclusion All the tests conclude that $\xi(n)$ and $b^2(n)$ form IID sequences. It is only confirming our previous study. It brings a supplementary proof to the correctness of our reasoning.

12.1.11 The methods of sections 11.1 and 12.1

Comparisons of the two methods

The method defined in this section 12.1.1 has theoretical results much better than those defined in the chapter 11.

But, such a quality of the approximation seems useless since one reasons on samples. In our method defined in chapter 11, we obtained an approximation theoretically less fine and yet, we saw that one can regard it as sufficient.

The improvement made in this section by the method defined in the chapter 11 seems not to mean much. For example, there exists always a probability close to 0.045 such as $\frac{|P_e - (1/2^p)|}{\sigma_B \sqrt{qN}} \geq 2$.

The approximation provided by the method defined in this section 12.1.1 can thus be only one additional guarantee which one can take when one builds a sequence of random bits b_n. It could however to be useful if one wanted to build functions of the b_n with certain mathematical properties

However considering the quality of the theoretical results, why then not employ only the method defined in this section 12.1.1?

To answer to this question, it should initially be recalled that it there many other possible methods that those defined in this article to transform sequences of data into IID sequences : cf section 12.2.

After, we studied much more the method defined in the chapter 11 than that defined in this section 12.1.1. Following this study, we do not understand any possible fault for the method defined in the chapter 11.

With regard to the second method, the rate of convergence of the XORLT should be the subject of a systematic study for better specifying the result. Indeed, we studied it only in the case of sequence $p'_{x_1,...,x_S}$ which are IID and chosen randomly: cf section 5.5. Therefore, this study about the XORLT would deserve to be developed and the results to be better specified.

Joint use of the two methods

A means of reducing the uncertainty of the model would be also to employ as sequence a(j) - defined in construction in section 12.1.1 - sequences of bits $b^0(n')$ created by the method of Fibonacci in the chapter 11.

One could thus be ensured that the lines are taken randomly and do not have probabilities concentrated nearly a small number of points. Moreover, one could apply the reasoning of the section 12.1.5 for each line.

12.1.12 Use of an additional congruence

One could want to use the method described hereinbefore in section 12.1.1, and after to finish by the use of theorem 1 as one does it in section 11.1. In fact it would be useless. One however will describe this method because it makes possible to better understand the utility of congruences in the construction of the sequences of random bits.

Theoretical method of construction of the sequence b_n

One keeps the same general method that in section 12.1.1 until step 9, the definition of k with S=15. One suppresses the step 6. One thus starts again the definition of construction from this step by supposing J=2J', $J' \in \mathbb{N}^*$.

9) We set $k = \overline{h} \bmod M_2 + 1$ where $M_2 = 2^{2J'} - 1$ is very big.

10) We set $\ell = T^*(k)$ where T^* is the function defined in proposition 4.1.1 with $m = M_2$.

11) Let $q_2 \in \mathbb{N}^*$. Let $Q = J' - q_2$. Then, one considers the sequence $b_1, b_2, ..., b_Q$ where $\ell = \overline{b_1 b_2 ....}$, is the writing of ℓ base 2.

Under these hypotheses, it is the sequence of bits b_s which will check the properties that we want for our IID sequence : one thus selected q_2 according to these conditions : cf proposition 12.1.3 below.

By acting thus, one wants to apply the same method as in the chapter 11 but one avoides the introduction of the conditional probabilities which require strong assumptions.

On the other hand, it is obvious that as soon as J' is big, one cannot use congruences of Fibonnacci, their calculation being too much long. One thus replaces by congruences easy to calculate : the binary congruences.

Method of calculation of the b_i

First, et us remind that to apply binary congruences, it is enough to invert the J' first bits with the J' last ones: cf section 4.1.2. Therefore, to obtain the b_i, the following operations are made.

1) If $h = \overline{\overline{h_{2J'+r}.....h_{2J'+1}h_{2J'}h_{2J'-1}.....h_1}}$, $k = \overline{\overline{h_{2J'}h_{2J'-1}.....h_1}}$

2) Then one calculates $\ell = T^*(k)$ modulo $M_2 = 2^{2J'} - 1$.

For that, we set, $k = \overline{\overline{k_{2J'}k_{2J'-1}.....k_1}}$, the writing of k base 2. Then, by section 4.1.2, $T^*(k) = \overline{\overline{k_{J'}k_{J'-1}.....k_1k_{2J'}k_{2J'-1}.....k_{J'+1}}}$.

Let us notice that the chance that $k = 2^{2J'} - 1 = \overline{\overline{1.1.1.....1}}$ is negligible. One excludes this case which it would not be possible to consider as representative of a IID sequence.

3) In order to apply the property 12.1.2, one limits oneself to the Q first bits : $b_1, b_2, ..., b_Q = k_{J'}k_{J'-1}.....k_{J'-(J'-q_2)+1}$.

Properties of the sequence B_i

One will understand now that there is no improvement by using binary congruences. To simplify, it is supposed that the following assumption is satisfied.

Hypothesis 12.1.1 *One supposes that the following inequality holds*

$$\left| P\{K = k\} - P\{K = k'\} \right| \leq K_K |k - k'| / 4^{2J'} \ ,$$

where $K_K \leq 1$.

Let us notice that $K_K = K_0$ the coefficient of Lipischitz of the theorem 1 increasing the slope of the density. Indeed, that is a consequence of lemma 5.6.3.

Now, one wants to apply the theorem 1.

Lemma 12.1.6 *We keep the notations of proposition 4.1.1 with d=2. Let I be an interval of* $F(M_2)$. *Then* $P\{T^*(K)/M_2 \in I\} = L(I) + Ob(1)\epsilon_I$ *where* $\epsilon_I \leq 1/d^{J'-4}$.

Proof By proposition 4.1.1 with d=2, p=J',
$P\{T^*(K)/M_2 \in I\} = L(I) + e(K_K, d, p)$, where
$e(K, d, p) = 2(3K+4)/d^p + O(K/M_2)$ and
$O(K/M_2) = \frac{2(K+1)}{M_2}\big(2 + \frac{d^p+1}{d^p(d^p-1)}\big) + \frac{3(K+1)}{d^p(d^p-1)}\big(1 + 1/d^p\big)$.

Therefore, $\epsilon_I = 2(3K_K+4)/d^{J'} + O(K_K/M_2) < 16/d^{J'} = 1/d^{J'-4}$ because d=2. ■

Proposition 12.1.2 *For all injective sequence n_s, s=1,2,...,p, $p \le Q$, $1 \le n_s \le Q$, the following equality hold*

$$P\Big\{\{B_{n_1} = b_1\} \cap \{B_{n_2} = b_2\} \cap ... \cap \{B_{n_p} = b_p\}\Big\} = \frac{1}{2^p}\big[1 + Ob(1)2^{3-q_2}\big].$$

Proof Let $T^*(K) = \overline{\overline{b_1, b_2, ..., b_{2J'}}}$. Then

$$T^*(K)/M_2 = \frac{T^*(K)}{2^{2J'}-1} = \frac{T^*(K)}{2^{2J'}} + T^*(K)\Big[\frac{1}{2^{2J'}-1} - \frac{1}{2^{2J'}}\Big]$$

$$= \frac{T^*(K)}{2^{2J'}} + T^*(K)\Big[\frac{2^{2J'} - (2^{2J'}-1)}{(2^{2J'}-1)2^{2J'}}\Big]$$

$$= \frac{T^*(K)}{2^{2J'}} + \frac{T^*(K)}{2^{2J'}-1}\ \frac{1}{2^{2J'}} = \frac{T^*(K)}{2^{2J'}} + \frac{e}{2^{2J'}}\ ,$$

where $0 \le e < 1$.

Then,

$$\frac{T^*(K)}{2^{2J'}} = \overline{\overline{0, b_1, b_2, ..., b_{2J'}}}\ .$$

Moreover,

$$\frac{e}{2^{2J'}} = \overline{\overline{0, 000....0b'_{2J'+1}b'_{2J'+2}.....}}\ ,$$

where $b'_s \in \{0, 1\}$.

Then

$$T^*(K)/M_2 = \overline{\overline{0b_1b_2, ...b_{2J'}b'_{2J'+1}b'_{2J'+2}.....}}\ .$$

Then, by lemma 12.1.6,

$$P\Big\{\{B_1 = b_1\} \cap \{B_2 = b_2\} \cap ... \cap \{B_Q = b_Q\}\Big\}$$

$$= P\Big\{T^*(K)/M_2 \in [\overline{\overline{0, b_1, ..., b_Q, 0, ..., 0}}, \overline{\overline{0, b_1, ..., b_Q, 0..., 0}} + 2^{-Q}\Big\}$$

$$= \frac{1}{2^Q} + Ob(1)2^{-(J'-4)} \ .$$

Therefore, by lemma 4.3.2,

$$P\Big\{\{B_{n_1} = b_1\} \cap \{B_{n_2} = b_2\} \cap ... \cap \{B_{n_p} = b_p\}\Big\}$$

$$= \frac{1}{2^p} + Ob(1)2^{Q-p}2^{-(J'-4)}$$

$$= \frac{1}{2^p} + Ob(1)2^{-q_2+4-p} = \frac{1}{2^p}\big[1 + Ob(1)2^{3-q_2}\big] \ . \ \blacksquare$$

Proposition 12.1.3 *If $q_2 \geq 14$, For all injective sequence j_s, s=1,2,...,p, the following equality holds*

$$P\Big\{\{B_{n+j_1} = b_1 \ \Big| \ B_{n+j_2} = b_2, ..., B_{n+j_p} = b_p\}\Big\} = \frac{1}{2}\ \big[1 + 2.1Ob(1)2^{3-q_2}\big] \ .$$

Proof We have

$$P\{B_{n+j_1} = b_1 \ \Big| \ B_{n+j_2} = b_2, ..., B_{n+j_p} = b_p\}$$

$$= \frac{P\Big\{\{B_{n+j_1} = b_1\} \cap \{B_{n+j_2} = b_2\} \cap ... \cap \{B_{n+j_p} = b_p\}\Big\}}{P\Big\{B_{n+j_2} = b_2\} \cap ... \cap \{B_{n+j_p} = b_p\}\Big\}}$$

$$= \frac{(1/2^p)\big[1 + Ob(1)2^{3-q_2}\big]}{(1/2^{p-1})\big[1 + Ob(1)2^{3-q_2}\big]}$$

$$= \frac{1}{2}\ \big[1 + 2.1Ob(1)2^{3-q_2}\big] \ . \ \blacksquare$$

This approximation is much less good than that of section 5.5.1 :

$$P\{K = y\} = (1/N)\Big[1 + Ob(1)\frac{C_1}{2^{2J'(S-1)/2}}\Big] \ .$$

That means that

$$P\Big\{\{B_{n_1} = b_1\} \cap \{B_{n_2} = b_2\} \cap ... \cap \{B_{n_p} = b_p\}\Big\} = \frac{1}{2^p}\Big[1 + Ob(1)\frac{C_1}{2^{2J'(S-1)/2}}\Big] \ .$$

Therefore, in this case, the use of binary congruence does not improve anything. This fact is confirmed by the use of virtual sequences.

Use of sequences of virtual numbers

Let us suppose that instead of having lines g_i length 2J' bits, one has lines of 4J' bits. One applies the technique defined in the section 12.1.12. It is then clear that the algorithms will be use primarily on the last 2J' (bits) of g_i.

Now, one can define this sequence of 4J' bits, by taking the sequence of 2J' bits g_i and by joining on the left a virtual sequence of 2J' bits g'_i.

Finally, one applies the algorithm A1 and A2 to all the sequence g_i and not only with his right half. The properties remain the same ones. The advantage, it is that one obtains a sequence of size (about) twice larger.

Therefore, in this case, the use of the theorem 1 with binary congruence nothing brings. It is normal considering it is limited to permute the J' first bits with the J' last ones.

Utility of the transformations with congruence

One wants to understand why there's not advantage in using here congruences and why it is useful with congruences of Fibonacci in the chapter 11.

At first, that would not be the same thing if one employed results on H which is close to the normal distribution whereas K is close to the uniform distribution.

On the other hand, the conguence of Fibonacci mixes the bits. It is not the case for binary congruences which permutes the J' first bits with the J' last ones.

In fact, the events $T_q^{-1}(k/d^q)$ provided by the Fibonacci sequence in section 11.1.2 are independent of the data which we use, for example text. It is clear that it is not the case of binary congruences: it transforms a text into text with two parts: it is always text. It is the origin of the problem.

12.2 Other methods of construction of an IID sequence

It there certainly several methods other than the ones introduced into this report to transform sequences of data into IID sequences. In this report, we studied two methods. But these methods are maybe too strong and one could have IID sequences by a simpler way.

12.2.1 Methode of Marsaglia

In this construction, one stops at the construction of the $e_S^2(j)$ in section 11.1.2: $e_S^2(j) = \overline{e_S^1(j) + rand_0(j)}$ where $e_S^1(j)$ are the sum of two sequences of data.

One studied the randomness of this type of sequences in chapter 3. By choosing the different parameters well, one can have good reasons so that the

obtained sequence have properties close to the randomness.

12.2.2 Method of transformations T_q

In this case, one stops at the following step of the construction defined in section 11.1.2: $e_S^3(j) = m_S T_1^{m_S}(e_S^2(j)/m_S^1)$.

It is known that the use of T_q implies that the $e_S^3(j)$'s are made uniform : cf chapter 8. Normally, it implies also their independence : cf section 8.4.

It is confirmed by a simple reasoning: let us try to know what means to apply the function T_q to a nondeterministic sequence. If a sequence X_n is nondeterminist, for any conditional distribution, one knows that the probability of X_n given $X_{n-1} = x_1, X_{n-2} = x_2, X_{n-3} = x_3, ...$ is not concentrated in one point. Let us suppose that a sequence is sufficiently not determinist, i.e. that X_n given $X_{n-1} = x_1, X_{n-2} = x_2, X_{n-3} = x_3, ...$ have a distribution distributed sufficiently well in an interval. Then if one applies T_q with this distribution, considering the properties of the functions of Fibonacci, one will make uniform the distribution : i.e. $P\{X_n \in Bo | X_{n-1} = x_1, X_{n-2} = x_2, X_{n-3} = x_3, ...\} \approx L(Bo)$, for all Borel set , and for all n. That implies that there is independence.

Therefore, the use of the function T_q does not make only uniform, but also seems to make independent. That is particularly understandable in the case where $(X_n, X_{n+1}, ..., X_{n+p})$ is close to a vector with a continuous density. Thus, the $e^3(j)$ are maybe IID.

It is all the more possible as, in the previous step (technique of Marsaglia), one wondered already if there were not randomness. But certainties are wanted. To have them, one would thus need a more thorough study of the functions of Fibonacci.

Counterexample.

However, there is a simple theoretical counterexample.

Example 12.2.1 *Let $(Y, Y') \in \{0, 1, ..., m-1\}^2$ be a random vector having a density f with respect to $\mu_m \otimes \mu_m$: $f(x, x) = m\beta$ and $f(x, y) = \alpha$ if $x \neq y$ where $\beta > \alpha$.*

Study First, $m^* m\beta/m^2 + \alpha(m^2 - m)/m^2 = 1$, i.e. $\beta + \alpha(1 - 1/m) = 1$ i.e. $\alpha = \frac{1-\beta}{(1-1/m)}$. For example, if $\beta = 1/2$, $\alpha = \frac{1}{2(1-1/m)}$.

Let $\Delta = \{(n, n) | n = 0, 1, .., m-1\}$ be the diagonal of $\{0, 1, .., m-1\}^2$. Let $\Delta' = \{(k/d^q, k/d^q) | k = 0, 1, ..., d^q - 1\}$ be the diagonal of $\{0/d^q, 1/d^q, ..., (d^q-1)/d^q\}^2$.

Then, $P\{(T_q(Y), T_q(Y')) \in \Delta'\} \geq P\{(Y, Y') \in \Delta\} = \beta$.

There will be a breakdown of independence as soon as β is larger than α : in this case, the absolute value of the linear coefficient of correlation $|\rho(T_q(X), T_q(Y))|$ will be much larger than 0.

As a matter of fact this results is true for $|\rho\big(\gamma[T_q(X)], \gamma[T_q(Y)]\big)|$ for all function γ. In particular, it is true if $\gamma = P_n$ the orthonormal polynomial of degree n, i.e. it is true for the polynomial correlation coefficient of order (n,n) $|\rho_{n,n}\big(T_q(X), T_q(Y)\big)| = |\rho\big(P_n[T_q(X)], P_n[T_q(Y)]\big)|$ (cf [10]). These coefficients $|\rho_{n,n}\big(T_q(X), T_q(Y)\big)|$ show an important breakdown of independence. ■

This example could be apply for text y_n if m is rather small. It is possible that, for example, words of 10 letters (or word groups) are repeat several times. But one is in the case of marginal probability concentrated nearly a small number of points : we can eliminate this case by choosing m large and $my'_n = \overline{my_n + g_n}$ where y_n means the sequence of data and g_n means a pseudo-random generator : cf section 8.2.2.

Another reasons so that there is not independence

Still let us regard the $my'_n = \overline{my_n + g_n}$.

Contrary to the unidimensional case, the $(g_n, g_{n+1}, ..., g_{n+p})$, n=1,2,.....,m, generated by the pseudo-random generators g_n does not generate all the set $\{0/m, 1/m, ..., (m-1)/m\}^p$, but a subset of the type of a subspace with one dimension.

In the same way, all the points of $\{0/m, 1/m, ..., (m-1)/m\}^p$ cannot be image of a $(y_n, y_{n+1}, ..., y_{n+p})$, $n = 1, 2, ..., n_0$. Indeed, they means texts with p elements.

One will be thus, theoretically at least, in the case where probability associated with the $(g_n, g_{n+1}, ..., g_{n+p})$ is concentrated in a small number of points. There will be thus the possible disadvantages described in section 8.2.2: one will not be able to apply the same reasoning that those used to obtain the uniformity.

Reminders

It is infinitely more difficult to prove than there is independence than to prove than there is uniformity. For a sequence of size 10^6, there is a quasi infinite number of possible dependence whereas there is only one uniformity : cf section 2.1.5.

However, it should be remined that certain data seem of a type simpler: it behaved as Qd-dependent sequences (cf section 10.4), sometimes even 2-dependent (cf section 11.2.9).

Conclusion

It is thus possible that there is independence in some cases there. But, that would require to be confirmed by a more thorough study.

12.2.3 Other congruences than Fibonacci congruences

One has used the Fibonacci congruences in the chapter 11 and in section 12.1 because they have the best properties. But, one could use other congruences having very close properties and satisfying for example $sup(h_i) = 3$ or 4 (cf notations 6.1.2).

12.2.4 Use of random permutation

In the method of the section 12.1, one understood how to use random transformations of type of permutations. Then, one could thus want to stop with $r_i(j) = \overline{r_0^i(j) + r_1^i(j) + r_2^i(j) + r_3^i(j)}$ modulo 2.

Indeed, the use of permutations involves the randomness. But one has not exactly permutations. As, it would have to be proven that one can stop with this step.

On the other hand, the use of all the method defined in section 12.1 allows to obtain mathematical properties such as that defined in proposition 12.1.1 which can be useful for some research. It can thus be useful in some cases to use all the method defined in section 12.1.

Chapter 13

Study of models

In this chapter we study the sequence of random variables $D_S(j)$, $E_S^1(j)$ and $E_S^2(j)$ associated to sequences of reals $d_S(j)$, $e_S^1(j)$, $e_S^2(j)$ used in section 11.1.2 or 12.1.1. More generally, we study the sequence of random variables $X_n = T_q(Y_n)$, n=1,2,...,N, where T_q is the Fibonacci function.

13.1 Continuous densities

In this section we study the case where $Y_n \in F(m)$ have a continuous density with respect to μ_m with a Lipschitz corfficient K_0 which is not too large.

13.1.1 General case

Let us suppose that one has a sample resulting from texts, y_n, $n = 1, 2, ..., n_0$, $y_n \in F(m)$ where $n_0 << m$. We suppose $y_n \neq y_{n'}$ if $n \neq n'$. Generally, that occurs always if m is great enough with respect to n_0. One can always obtain this assumption for normal texts. In order to be sure that this assumption holds, one can compute $Min(|y_n - y_{n'}|)$.

One can suppose that $Min(|y_n - y_{n'}|)$ is not too small with respect to the size of sample.

This assumption involves that, for all subsequence $y_{t(n)}$ and for all p,

$$(y_{t(n)}, y_{t(n+1)}, ..., y_{t(n+p-1)}) \neq (y_{t(n')}, y_{t(n'+1)}, ..., y_{t(n'+p-1)})$$

if $n \neq n'$.

Then, one can assume that $Y_n \in [0, 1]$ and that one is boiled down to case $Y \in F(m)$ according to the traditional method.

Then, one can regard y_n as the realization of a sequence of random variables $Y_n : y_n = Y_n(\omega)$ such that Y_n has a a differentiable density. One can also assume that this density have a Lipschitz coefficient K_0 which is not too large.

Now, if $Y \in F(m)$, there exist always a such density with respect to $\mu_m \otimes .... \otimes \mu_m$. In this case the only problem is to estimate K_0 in order to check that K_0 is not too large

It is a logical assumption: there is nothing which shows that it cannot be checked. Admittedly, it is not a sufficient reason. But, it is felt well that if one produces a sample of such a model, one will can obtain very likely the sample y_n : one wants to mean that for a such model, there exists a reasonable chance to obtain the fixed sample y_n (of course all that requires to be specified).

As a matter of fact it is an assumption which most researchers admit: that is especially clear when they estimate the densities (which they suppose to exist).

Therefore, when one has a sample y_n, n=1,...,N where $N << m$, it is reasonable to admit that it is the realization of a sequence Y_n such as the associated Lipschitz coefficient is not too large

Now, if one uses texts y_n, the model is not exactly known: it is only known that there exists a fixed model Y_n with a differentiable density such that K_0 is not too large, there exists a reasonable chance to obtain the fixed sample y_n : $y_n = Y_n(\omega)$

It is a rather intuitive conjecture. But it is thus admited implicitly by most mathematicians.

We have studied numerous examples which corroborate this hypothesis. But it would be too long to detail. As a matter of fact all this needs to be explicated.

For example, one can use a process of estimation. But there is a problem.

Indeed, let us clarify at first what we shall want as model: let N be the size of the sample. We want that $(Y_1, ..., Y_N)$ have a density with a Lipschitz coefficient which is not too big . The problem is that this density is theoretically impossible to estimate with samples y_n of size N. Indeed, in theory, it is necessary that there is more than N "y_n" to use a process of estimation of the density of $(Y_1, ..., Y_N)$ (cf section 10.2.3).

If one wants to use a process of estimation, we could only estimate the marginal distributions and, for example, a few higher order correlation coefficients. But we do not want an estimation but a right model. Then, one can, also estimate more coefficients in order to have an idea of the model. After, we vary by a continuous way the obtained model. By this way, we can have an idea about model $(Y_1, ..., Y_N)$. One estimates thus the Lipschitz coefficient K_0. By using this method, we understand that it is not too big (maybe except if a higher order correlation coefficient is big). That means that our model is right.

It is also possible to estimate $(Y_1, ..., Y_N)$ where $N << N'$ with sample of size N' in order to have an idea of K_0. One can use other texts in order to increase the size of the sample. All these methods corroborate that K_0 is not too large. As a matter of fact, it is not surprising because it is the assumption which is used often in estimation (except if one uses higher order correlation coefficients).

It is understood that this problem is complicated. The fact remains that this assumption is intuitively logical and that it is very often admited.

Let us notice that there is another difficulty to solve : when one wants to study the couples $(y_{t(2n)}, y_{t(2n+1)})$ associated with $y_{t(n)}$, there is a problem. One can often for example choose the permutation t such as it exists n_1 and n_2 such that $|y_{t(2n_1)} - y_{t(2n_2)}|$ and $|y_{t(2n_2+1)} - y_{t(2n_2+1)}|$ are both minimum elements in the set of the $|y_n - y_{n'}|$. Therefore, the distance

$$\left|\left|(y_{t(2n_1)}, y_{t(2n_2+1)}), (y_{t(2n_2)}, y_{t(2n_2+1)})\right|\right|$$

is too small.

It will be similar in 3,4,5 dimensions etc. One will thus obtain Lipschitz coefficients which are a little too large.

But it is not an important problem because, for samples, there will be always a priori empirical higher order correlation coefficients larger than others ones. That thus boils down to the traditional problem to choose tests associated with a sample when this one is known.

In any case, this problem would require to be studied in detail and risk to imply a long study if one wants to solve it completely.

13.1.2 Case of text

We have just understood that, if $N << m$, one can admit that $(X_1, ..., X_N)$ has a density with a Lipschitz coefficient K_0 not too large. Then, this model is logical.

Now, by an other way one can admit that it is not logical. Indeed, the set of texts is s much smaller than $F(m)$. Then there exists many points p_s such that $P\{Y_n = p_s\} = 0$.

It is a contradiction. The two assumptions are logical but contradictory.

Then, what is the good model?

In the theory of the models, the important thing, it is what one wants: here, it is that one can manage to have numbers which are unpredictable . The model with K_0 not too large seems right for this sample. Then, the conclusion are also right.

Now, because one knows that one uses texts, one can choose some ways in order to improve the result. Thus, we add a pseudo random generator g'_n : $\overline{y_n + g'_n}$ (cf section 8.2.2, page 200).

If one use many texts, one knows that one will have some points p_s such that $P\{Y_n = p_s\} = 0$ and other points p'_s such that $P\{Y_n = p'_s\} > 1/m$. But is is in the future (with respect to N). We do not need to use this property.

Anyway nothing proves that one will continue to choose English texts like sources of numbers. Thus certain conclusions can be erroneous.

For choosing the model well, it is necessary to take account of that which one needs. Admittedly one could choose the ideal model only concentrated on all the possible English texts. A priori it would be logical. But one understands well that an infinity of model can be associated with a sample. The model with differentiable density is as logical as the English texts.

Finally, all that one can say, it is that there is an infinity of possible models which are well reasonable. The fact of knowing that they are texts can bring to us some additional conclusions like asymptotic independence or like the addition of a pseudo random genrator $g'_n : \overline{y_n + g'_n}$.

13.1.3 Case of conditional probabilities

Now, in the case where densities are continuous, the conditional densities are also continuous.

That means that the conditional probabilities $P\{Y_n | y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\}$ has a continuous density $f_{y_2,...,y_p}$ with a coefficient Lipschitz K_0 which is not too great.

Let $\overline{T}^{-1}(I) = \{a_1, a_2, ..., a_{c'-c}\}$ where $I = [c/m, c'/m[$ is an interval. Then, by classical integration techniques,

$$\begin{aligned} & P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} \\ &= \sum_{j=1}^{c'-c} P\{Y_n = a_j | Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} \\ &= \sum_{j=1}^{c'-c} (1/m) f_{y_2,...,y_p}(a_j) \approx L(I) \end{aligned}$$

It is exactly the technique which we developed for properties 7.1.22 and 7.1.18. We recall that we obtained the following equalities :

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{c_2}{\sqrt{c'-c}}\Big]$$

or

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{O(1)K_0^{cp}}{c'-c}\Big] .$$

Of course, K_0^{cp} is the Lipschitz coefficient associated with conditional densities. To be sure our results for a fixed model, it is necessary to estimate it.

In order to increase the slope of conditional probabilities, one can increase the slopes of

$$\frac{f_1(y^0, y_2, y_3, ...)}{f_2(y_2, y_3, ..., y_p)}$$

where $f_2(y_2, y_3, ..., y_p)$ is the density of $(Y_{n+j_2}, Y_{n+j_3}, ..., Y_{n+J_p})$ and where $f_1(y^0, y_2, y_3, ...)$ is the density of $(Y_{n+j_1}, Y_{n+j_2}, Y_{n+j_3}, ..., Y_{n+J_p})$ with $j_1 = 0$ (it is the density with respect to the uniform measure).

In order to obtain K_0^{cp}, it is thus necessary at first to increase the slope of $f_1(y^0, y_2, y_3, ...)$ for the variable y^0. That can be done by ordering the points y_n^0 of the sample. After, one takes those of minimal distances, which gives us an idea of the value of K_0^{cp}. It is also necessary to underestimate $f_2(y_2, y_3, ..., y_p)$. That can be done by estimate.

A simpler mean is to estimate the density of dependence $f^{dep}(y_1; y_2, y_3, ..., y_p)$ by using higher order correlation coefficients considering than one has $f_{y_2,...,y_p}(y_1)$ $= f^{dep}(y_1; y_2, y_3, ..., y_p) f_2(y_2, y_3, ..., y_p)$: cf proposition A.3.2.

One can also estimate the $P\{Y_n \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\}$. In this case one chooses $j_s = s - 1$. Indeed, in the case of text, the dependence is maximum for $j_s = s - 1$.

Finally, one can estimate K_0^{cp}. We deduce from this estimation an increase which will have to be certain. That will enable us to have the following property (for the model where K_0 is not too large) :

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{O(1)K_0^{cp}}{c' - c}\Big] .$$

13.2 Another group of models

The model that we have just studied is that where the coefficient of Liptchitz K_0 is not too large. We know that, under this assumption, one has always by properties 7.1.18 or 7.1.22 :

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} = L(I)\Big[1 + \frac{O(1)K_0}{c' - c}\Big] .$$

However, in order to prove these results, we use K_0. But it is enough to read the proofs of these properties 7.1.18 or 7.1.22 in order to understand that it would be possible to use the coefficients of Lipschitz K^r associated with each interval $[r/N(I), (r + 1)/N(I)[$ to obtain the same result. In this case, there would be weaker hypotheses: it would be enough that $\sum_r K^r$ is not too large.

It is felt well intuitively that this kind of conditions is satisfied by our models.

Admittedly, it is easier to understand for the classical densities of $(Y_1, ..., Y_N)$. But what interests us, are the conditional probabilities. Then to understand that the property "$\sum_r K^r$ not too large" is checked for the conditional densities, simplest way is to remember that the conditional density $f_{y_2,,y_p}(y_1)$ is equal to the product of the marginal densities $f_2(y_2, y_3, ., y_p)$ and of the density

of dependence (cf proposition A.3.2) :

$$f_{y_2,...,y_p}(y_1) = f^{dep}(y_1; y_2, y_3, ..., y_p) f_2(y_2, y_3, ..., y_p) \ .$$

Therefore, this kind of conditions on $\sum_r K^r$ seems checked by our models.

As these assumptions are more general, that makes our end result even surer. Our results are thus true for all the models checking this new assumption "$\sum_r K^r$ not too large". And in these models, there is of them inevitably one (and even an infinity) which is models correct for the sample y_n.

But in our opinion, it is likely useless to study these new models, the model supposing only "K_0 not too large" is sufficient as soon as $N << m$ is chosen.

13.3 General case

The use of the previous models is interesting because it shows that y_n behaves indeed like a sample of one of these possible models Y_n. We thus do not make any error while putting to us under these assumptions. Our calculations are thus right. That implies that the bits $b^0(n')$ obtained by our construction (cf section 11.1 and 12.1.1) behave indeed like IID sequences.

But it is probable that there do not need even to suppose to be under the assumptions of one of these models: it is probable that, for *any logical model*, one will still obtain

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} \approx L(I) \ .$$

It is almost the results of sections 8.2 and 8.4 .

13.3.1 A very strong result

Indeed, we understood that if one provides the set of possible probabilities with the measure defined in section 8.2, our results are checked for almost all the possible probabilities.

There is thus a slight restriction: some models are not appropriate. This restriction is normal: one knows very well that in the set of ALL the models (thus without no a priori) there will be an infinity of them which will not be appropriate.

However, it is already extraordinary that the result is true for almost all the possible models.

In order to understand it, let us take for example a sample really IID y_n, i.e. obtained starting from a sequence of random variables Y_n^0 when we do not know that Y_n^0 is IID.

One wants to associate with y_n a model Y_n. Under assumption IID, all the possible models without a priori (except one: Y_n^0) are bad: a priori there are a correct model and a noncountable infinity of bad one.

Now, if one knows nothing a priori about y_n, there will be an infinity of good models and an infinity of bad ones. One can think that, in a certain way (for example if it is supposed that the probabilities are written with a limited number of decimals), there will be much more bad models than goods.

If one uses the results of sections 8.2 and 8.4, it is the opposite: in the set of all the possible models, almost all will be good. It is a very strong result.

However, a priori there remain bad models associated with our sample: we want to mean thus that they will be such as

$$P\{T_q(Y_n) \in I\} \neq L(I)$$

or

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} \neq L(I) \ .$$

13.3.2 A result checked by all the logical models

As a matter of fact, one can remove these bad models. Indeed, one can admit that $P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} \approx L(I)$ will be checked for all the *logical* possible models.

Indeed, for the moment, one has used no assumption on the y_n. However, there is no connection between the $T^{-1}(I_k) = \{a_1, a_2, ..., a_{c'-c}\}$ and text (in any case, it is an assumption extremely extremely reasonable. Moreover we also checked it by numerical computations).

Therefore, if a model was bad, that would mean that there is a logical connection between the $T_q^{-1}(I_k)$ and text. A priori, it seems that can occur only with one similar probability (even smaller or infinitely smaller or even null) to that which a really IID sample has to check the tests that it has to check.

One can thus a priori exclude a such model.

Thus our result holds with all the possible logical models, those where there is no connection between text and the $T_q^{-1}(I)$.

Of course, nothing proves completely that this assumption is true. But if it was wrong, it would be with a negligible probability due to the text and thus due to the sample : one would find the fact that a sample can always be bad for a fixed model.

In all way, a priori it seems that it has there no connection between the $T_q^{-1}(I_k)$ and text. That means that, empirically there is no connection between the y_n and the $T_q^{-1}(I_k)$.

Let us suppose now that they are false: let us suppose that the French texts, for a strange reason of structure, have a connection with the $T_q^{-1}(I)$. In this

case, it is enough to add modulos 2 another sequence obtained starting from other data of another type (texts in another language, another type of data).

Indeed, if one adds two unspecified sequences modulos 2, it is enough that only one is IID so that the sum is IID : cf section 2.6.7, page 44.

However, it appears incredible which the program data processing, for example, have also a connection with the $T^{-1}(I)$ if it is also the case for the French texts, so much is large the differences in structure between these programs and the French texts.

Of course, one can always consider that there is a negligible probability in order that that carried out for both. But it is so negligible that one can admit that it is impossible.

In any case, it is not here the matter of statistical study, but about a reasonable and logical connection. It is very probable that this connection does not exist. Therefore, one can use the functions of Fibonacci and find

$$P\{T_q(Y_n) \in I\} \approx L(I) \ .$$

That means that, very probably, **our result is true for all the possible logical models**.

It is a very strong result.

13.3.3 What logical models?

Then, it is obtained that $P\big\{Y_1 \in \widehat{T}^{-1}(I)\big\} \approx \frac{(c'-c)}{m} \ \Big[1 + \frac{Ob(1).b}{\sqrt{3N(I)}}\Big]$ for all the logical models because there is no logical connection between text and $T_q^{-1}(I)$. Then the question is put: which value to choose for b?

If the previous result is not true, that means that b is too small. In order to be sure that for the logical models one has an equation of this type, one can increase b. For example, one could choose

$$P\big\{Y_1 \in \widehat{T}^{-1}(I)\big\} = \frac{(c'-c)}{m} \ \Big[1 + \frac{Ob(1)}{\sqrt{3}N(I)^{1/4}}\Big] \ .$$

As a matter of fact, the question is: How to know of which is order must be b in order to admit that the previous equations are logically true?

In order to know that, it is necessary to go back to the themself texts them: i.e. it is necessary to study the associated empirical probabilities.

We thus estimated b for various texts and for various $T_q^{-1}(I)$.

If p=1, all the numerical studies that we have made show that, for intervals I of the same length, the sets $T_q^{-1}(I)$ contains about the same number of possible texts. More precisely, one has carries out the chi squared tests of uniformity with $D = d^q - 1$ degrees of liberty for various texts. At first, it was supposed that one has a partition of F(m) in d^q subsets $T_q^{-1}(I)$. Because D is large, we use $N_G = \sqrt{2\chi_2^D} - \sqrt{2D-1} \xrightarrow{D} X_G^*$ where $X_G^* \sim N(0,1)$: cf proposition A.1.2.

Moreover we denote by α_5 the selected percentage points with probability 0,95 of G. Then, we have obtained the following results.

D	1000	2500	1000	3500	200	400	1000	5000	10000
$\|N_G\|$	0.640	0.854	0.302	1.154	1.241	0.541	0.921	1.084	1.321
α_N	1.960	1.960	1.960	1.960	1.960	1.960	1.960	1.960	1.960

One has made the chi squared tests of uniformity with $D' > 30$ degrees of freedom without supposing that one has a partition : various subsets $T_q^{-1}(I)$'s were used where $\widehat{T}^{-1}(I)$ is smaller. The following results were obtained.

$\|N_G\|$	0.721	0.801	1.254	0.345	0.905	0.677	1.784	1.112	0.549
α_N	1.960	1.960	1.960	1.960	1.960	1.960	1.960	1.960	1.960

Many of other tests were made. Finally, it is found that one can admit - and by far - in all the cases

$$P\big\{Y_1 \in \widehat{T}^{-1}(I)\big\} = \frac{(c'-c)}{m}\ \Big[1 + \frac{Ob(1).20}{\sqrt{3N(I)}}\Big]\ . \tag{13.1}$$

This increase (b=20) is not astonishing. Indeed, according to proposition 8.4.1, it occurs with a probability larger than $1 - \Gamma_1(b)$. However, if b=20, $\Gamma_1(b) \approx 1.12/10^{88}$. Then, a priori, in order to find a case where that is not true, it would be necessary to use a such large number of texts that it is impossible to realize.

In any case it is even not sure that one can find cases where this equality is not checked empirically. Indeed, one does not use texts representing samples which have a fixed law. What one uses, it is, on the one hand sequences which have the logic of the English language, and on the other hand sets which have simple mathematical properties. Anyway, we never have encounter such a case.

Thus, if one writes the sets $\widehat{T}^{-1}(I)$ with letters one understands that those are unions of the elements of the type

[whgkudf ly cuqhjg]
[aamxgusdggbxckmp]
[x;cbkutcc ze xycyc x]
[qtdxucdzlcxy yx vyxy]
[uezuxcuazvxaoaqzq]
[,hqdsgcize cqy bxq]
[a picykhgkkl hfqfqqq]
[ory of Relativity in 190]
[xwtex pez! i yi qy yqhfg]

.....

It is thus possible logically that it has no text not checking the equation 13.1.

In two dimensions, one will have the same type of results. But, in this case, it is better to use the results of [18]. In particular,

$$P\Big\{\big\{Y_1 \in \widehat{T}^{-1}(I_1)\big\} \cap \big\{Y_2 \in \widehat{T}^{-1}(I_2)\big\}\Big\} \approx \frac{\prod_s (c'_s - c_s)}{m^2} \left[1 + \frac{Ob(1)2.20}{\sqrt{3.Inf\{N(I_s)\}}}\right].$$

As a matter of fact, if $\widehat{T}^{-1}(I_1) \neq \widehat{T}^{-1}(I_2)$, sets $\widehat{T}^{-1}(I_1)$ and $\widehat{T}^{-1}(I_2)$ behave like randomly selected compared to the text.

If $\widehat{T}^{-1}(I_1) = \widehat{T}^{-1}(I_2)$, sets $\widehat{T}^{-1}(I_1)$ behave also like randomly selected compared to the text. It will be due to chance if two texts which can be consecutive belong to this same set (one chooses two texts which are consecutive because it is there that there is the strongest dependence). Now, it will be impossible that two identical texts belongs to the set $\widehat{T}^{-1}(I_1) \otimes \widehat{T}^{-1}(I_2)$. But this case is also excluded if $I_1 \neq I_2$.

If $p > 2$, we have obtained results equivalent for $p \leq \frac{log(n_0)}{log(10)}$ [1]:

$$P\Big\{\big\{Y_1 \in \widehat{T}^{-1}(I_1)\big\} \cap ..... \cap \big\{Y_p \in \widehat{T}^{-1}(I_p)\big\}\Big\} \approx \frac{\prod_s (c'_s - c_s)}{m^p} \left[1 + \frac{Ob(1).20.p/\sqrt{3}}{\sqrt{Inf\{N(I_s)\}}}\right].$$

Let us notice that *a priori* it is possible that the $\{a_1, a_2, ...\}$ have a connection with the empirical probability , i.e. if p=1,

$$P\big\{Y_1 \in \widehat{T}^{-1}(I)\big\} \neq \frac{(c' - c)}{m} \left[1 + \frac{Ob(1).20}{\sqrt{3N(I)}}\right].$$

But that is likely to occur with a negligible probability as it is the case when one tests if an IID sample is well IID.

Thus, *a priori* it is always possible that the sequence x_n is not IID, but it would be with a negligible probability.

13.3.4 Conditional probabilities

The fact that there is no connection between text and the sets $T_q^{-1}(k/d^q) = \{a_1, ..., a_{c'-c}\}$ applies to the conditional probabilities. Indeed, there is always no logical connection between text and the sets $T_q^{-1}(I) = \{a_1, ..., a_{c'-c}\}$ in the conditional probabilities:

$$P\{X_n \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\}$$

[1] We recall that n_0 is the size of sample : if one use $p \geq log(n_0)/log(10)$, the results do not mean anything any more and are increasingly random

$$= \frac{P\Big\{ \ \{Y_n \in T_q^{-1}(I)\} \cap \{Y_{n+j_2} = y_2\} \cap .... \cap \{Y_{n+j_p} = y_p\} \ \Big\}}{P\Big\{ \ \cap \{Y_{n+j_2} = y_2\} \cap .... \cap \{Y_{n+j_p} = y_p\} \ \Big\}} \ .$$

Therefore, the conditional probabilities behave well as sums on sets taken randomly, i.e.

$$P\{T_q(Y_n) \in I \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\}$$
$$= \sum_s P\{Y_n = a_S \mid Y_{n+j_2} = y_2, ..., Y_{n+j_p} = y_p\} \approx L(I) \ .$$

13.4 Consequences 1

By applying our results to the techniques of section 11.1, for all the reasonable models D(j), or $E^1(j)$ or $E^2(j)$ or $H(n)$, one understand that, by properties 7.1.22 or 7.1.18,

$$P\big\{T_{q_0}(H(n)/m) \in I \ \big| \ H(n+j_2) = h_2, ..., H(n+j_p) = h_p\big\} = L(I)\Big[1 + \frac{c_2}{\sqrt{N(I)}}\Big]$$

or

$$P\big\{T_{q_0}(H(n)/m) \in I \ \big| \ H(n+j_2) = h_2, ..., H(n+j_p) = h_p\big\} = L(I)\Big[1 + \frac{O(1)K_0}{N(I)}\Big] \ .$$

This result is true for the all models D(j) (or H(n)) where K_0 is not too large, or for all the models D(j) where $\sum_r K^r$ is not too large, or for all the logical models D(j) associated to the sample d(j).

Thus one is sure that according to our assumptions, for the sequence of bits $B^0(n')$,

$$P\big\{B_n^0 = b \ \big| \ B_{n+j_2}^0 = b_2, ..., B_{n+j_p}^0 = b_p\big\} = \frac{1}{2}\Big[1 + \epsilon\Big] \ ,$$

where ϵ is small enough with respect to Nq_0 the size of sample.

Now, one knows that when ϵ is small enough with respect to Nq_0, it is not possible to distinguish $b^0(n')$ from a sample IID. It is particularly obvious for $\epsilon = \frac{1}{2^{50.000.000}}$ as in section 12.1.9.

In this way, one has well proved that the sequence $b^1(n')$ behaves like a sample of an IID sequence of random variable.

Remark 13.4.1 *If $\epsilon = \frac{1}{2^{50.000.000}}$, it likely possible to obtain an ultimate sequence which is completely IID (and not only which has the same behavior that a sequence IID), i.e. a sequence checking*

$$P\big\{B_n^0 = b \ \big| \ B_{n+j_2}^0 = b_2, ..., B_{n+j_p}^0 = b_p\big\} = \frac{1}{2} \ .$$

It is likely enough to use functions $(X'_1, ..., X'_N) = f(X_1, ..., X_N) \in [0,1]^N$ *and to vary slowly f and continuously in order to obtain that* $X'_1, ., X'_N$ *is IID.*

Then it is very probable that $b^1(n')$ *will be also a reasonable sample of the new model, i.e. of the IID model.*

It is only one remark. This method seems complicated to prove and useless since one has already proved that $b^1(n')$ *behaves like an IID sequence.*

13.5 True for all models

One thus has proves that $b^1(n')$ behaves like a sample of an IID sequence of random variables :
1) for the all models D(j) where K_0 is not too large
2) for all models D(j) where $\sum_r K^r$ is not too large
3) for all the logical models D(j) associated to the sample d(j).

Our result is thus well proved and it is proved for all the models that one can associate reasonably with the sequence d(j).

By this way, we obtain thus also a solution to the problem of the definition of a random sequence : thanks to the use of the Fibonacci functions, our results hold for all the reasonable models possible: that resolves the problem of definition.

13.6 Consequences 2

With the previous results and other chapter of this report, on can prove that, for all $p \leq Log(N)/Log(2)$, for all injective sequence j_s, for all logical models B_n,

$$P\left\{\sqrt{N}\big|P_e - \frac{1}{2^p}\big| \geq \sigma_B x\right\} = K_1\Big(\big[1-\eta\big]x\Big) \ ,$$

$$P\left\{\sqrt{N}\left|\frac{P_e}{p_e} - (1/2)\right| > \sigma_{cp}\ x\right\} = K_2\Big(\big[1-\eta'\big]x\Big) \ ,$$

where η and η' are small enough and σ_B^2 and σ_{cp}^2 are the variances associated to P_e and P_e/p_e when $P_e = (1/n_1)\sum_{n=1}^{n_1} 1_b(B_{t_n}) 1_{b_2}(B_{t_{n+j_2}}).....1_{b_p}(B_{t_{n+j_p}})$ and $p_e = (1/n_1)\sum_{n=1}^{n_1} 1_{b_2}(B_{t_{n+j_2}}).....1_{b_p}(B_{t_{n+j_p}})$ where t is a permutation and $n_1 \leq N$.

These results correspond well to definitions 2.1.7 and 2.1.8 for all logical model B_n. That means well that one cannot differentiate B_n from an IID sequence.

For example, in section 12.1.9. One has $\eta = O(\epsilon)$ and $\eta' = O(\epsilon)$ where $\epsilon = \frac{1}{2^{50.000.000}}$.

Appendix A

Summary of some mathematical properties

In this chapter, we remind some mathematical properties used in this report.

A.1 Chi squared independence test

When the marginal distributions are known, the chi-squared independence test is obtained by using the following property : cf [17].

Proposition A.1.1 *Let $(X,Y) \in \mathbb{R}^2$ be a random vector. Let I_s s=1,..,d, be a partition of $\mathbb{R}$. Let $\chi^2_{X,Y}$, χ^2_X, χ^2_Y the chi-squared statistics of, respectively, (X,Y) , X, Y, associated with the partitions $I_s \otimes I_t$, I_s and I_t, s,t=1,2,...,d.*

Then, if X and Y are independent, $\chi^2_{X,Y} - \chi^2_X - \chi^2_Y$ has asymptotically a chi squared distribution with (d-1)(d-1) degrees of freedom.

This test is more powerfull than that associated with $\chi^2_{X,Y}$: cf [17] .

On the other hand, if the degrees of freedom are large, one can approximate the chi-squared statistics thanks to the normal distribution : cf [1] page 44.

Proposition A.1.2 *Let χ^d_2 d=1,2,..., be a sequence of random variables which have the chi-squared statistics with d degrees of freedom.*

Then, $\sqrt{2\chi^d_2} - \sqrt{2d-1} \xrightarrow{D} X^*_G$ as $d \to \infty$, where $X^*_G \sim N(0,1)$.

A.2 Stochastic "O" and "o"

Notations A.2.1 *A sequence of random variable X_n is bounded in probability, if, for every $\epsilon > 0$, there exists M_ϵ and N_ϵ such that $P\{|X_n| \leq M_\epsilon\} \geq 1 - \epsilon$ for all $n \geq N_\epsilon$. Then, one writes $X_n = O_P(1)$. Moreover, we write $X_n = o_P(1)$ for sequence of random variable X_n if X_n converges in probability to 0 : (cf [42], page 8)*

A.3 Higher order correlation coefficients

First, we remind the definition of these coefficients : cf [10] and [9].

Definition A.3.1 *Let $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ be two random vectors defined on a probability space $(\Omega, \mathcal{A}, P)$. Let μ_X and μ_Y, be the laws of X and Y, respectively. Let $\{A_i\}$, i=0,1,2,.. and $\{B_j\}$, j=0,1,2,.. be two families of orthonormal functions associated. One assumes that $A_0 \equiv B_0 \equiv 1$. Then, one defines $\rho_{i,j}$, the correlation coefficient of order (i,j) by $\rho_{i,j} = E\{A_i(X)B_j(Y)\}$ for i=1,2,.., j=1,2,...*

Now the following propositions hold.

Proposition A.3.1 *Let $F_{X,Y}$,F_X,F_Y be the distribution function of (X,Y), X and Y, respectively. Assume that $\{A_i\}$, i=0,1,2,.. and $\{B_j\}$, j=0,1,2,.. are two bases of $L^2(\mathbb{R}^p, \mu_X)$ and $L^2(\mathbb{R}^q, \mu_Y)$, respectively. Then, for all (x,y),*

$$F_{X,Y}(x,y)$$

$$= F_X(x)F_Y(y) + \sum_{i\geq 1, j\geq 1} \rho_{i,j} \Big(\int_{u=-\infty}^{x} A_i(u).\mu_X(du) \Big) \Big(\int_{v=-\infty}^{y} B_j(v).\mu_Y(dv) \Big) \ .$$

Proposition A.3.2 *Assume that (X,Y) has a probability density function f with respect to $\mu_X \otimes \mu_Y$. Then, f is the dependence density.*

Let f^x be the function defined by $f^x(y) = f(x,y)$. Let μ_Y^x be the conditional distribution of Y given $X = x$. Then, f^x is μ_X-almost surely the probability density function of μ_Y^x with respect to μ_Y.

Proposition A.3.3 *Assume that p=q=1 and that $A_i = P_i$ and $B_i = Q_i$, i=0,1,2,.. are the families of orthonormal polynomials associated with μ_X and μ_Y, respectively. Then, $\rho_{i,j}$ is the polynomial correlation coefficient of order (i,j) and $\rho_{1,1} = \rho$, the classical linear correlation coefficient.*

Then, ρ is the first coefficient of a sequence which measures dependences more and more fine : $\rho_{1,1}$ measures linear dependence, $\rho_{1,2}$, $\rho_{2,1}$, $\rho_{2,2}$ measure quadratic dependences, etc.

Proposition A.3.4 *Assume that $\{P_i\}$, i=0,1,2,.. and $\{Q_j\}$, j=0,1,2,.. are two bases of $L^2(\mathbb{R}, \mu_X)$ and $L^2(\mathbb{R}, \mu_Y)$, respectively. Then, in $L^2(\mathbb{R}^2, \mu_X \otimes \mu_Y)$,*

$$f(x,y) = 1 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho_{i,j} P_i(x) Q_j(y) \ ,$$

with $\rho_{1,1} = \rho$.

One can generalize these results by the following way.

Definition A.3.2 *Let $(X_1, ..., X_p) \in \mathbb{R}^p$ be a random vector defined on a probability space $(\Omega, \mathcal{A}, P)$. Let $\{P_i^s\}$, i=0,1,2, be the families of orthonormal polynomials associated to μ_{X_s} the distribution of X_s. Then, one defines $\rho_{i_1,....,i_p}$, the polynomial correlation coefficient of order $(i_1, ..., i_p)$ of $(X_1, ..., X_p)$ by $\rho_{i_1,....,i_p} = E\{P_{i_1}^1(X_1)....P_{i_p}^p(X_p)\}$.*

Remark A.3.1 *One can estimate $\rho_{i_1,....,i_p}$ by using the empirical orthonormal polynomials associated with the empirical measure.*

Bibliography

[1] KNUTH D.E. (1998) the Art of Computer Programming; Vol 2. Third Edition Addison-Wesley, Reading, Massachusetts.

[2] GENTLE J. (1984) Random Number Generation and Monte Carlo Method, Springer 13, 61-81.

[3] MENEZES A., VAN OORSCHOT P. , VANSTONE S. (1996) Handbook of Applied Cryptography, CRC Press, 1996.

[4] VON NEUMANN J . (1951) Various techniques used in connection with random digits. Monte Carlo method, Applied Mathematics series N12, US National Bureau of Standards, Washington DC, 36-38.

[5] SCHNEIER B (1996) Applied Cryptography 2nd Edition, John Wiley and sons, Inc

[6] ROSENBLATT M. (1972) Uniform ergodicity and strong mixing. Z. Wahrsch. Werw. Gebiete. 24, 79-84.

[7] PINSKER M.S. (1964) Information and information stability of random variables and processes. Holden Day, San Francisco.

[8] ELIAS P. (1972) The efficient construction of an unbiased random sequence. Annals Math Stat, vol 43, n3, 865-870.

[9] LANCASTER H. O. (1960) Orthogonal models for contingency tables. Developments in statistics. Academic Press, New York.

[10] BLACHER R. (1993) Higher Order Correlation Coefficients. Statistics 25, 1-15.

[11] BLACHER R. (1995) Central limit theorem by polynomial dependence coefficients. Journal of computational and applied mathematics 57, 45-56.

[12] BLACHER R. (1990) Theoreme de la limite centrale par les moments. Compte rendus de l'Academie des Sciences de Paris. t-311 serie I, p 465-468.

[13] BLACHER R. (1983) Quelques propriétés des congruences linéaires considérées comme générateur de nombres pseudo-aléatoires. Rapport de recherche n 345 IMAG, Université Joseph Fourier de Grenoble.

[14] BLACHER R. (2007) Central Limit Theorem by moments. Statistics and Probability Letters, 2007; 77 (17) 1647-1651

[15] BLACHER R. (2007) Une nouvelle condition d'independance pour le theoreme de la limite centrale http://hal.archives-ouvertes.fr/hal-00144878/en/ HAL: hal-00144878, version 1

[16] BLACHER R. (2002) Transformation d'une suite aléatoire q-dépendante. Rapport de Recherche LMC-IMAG RR 1054-M, Universite de Grenoble.

[17] BLACHER R. (1988) A new form for the chi-squared test of independence. Statistics 19,4, 519-536

[18] BLACHER R. (2009-2010) A Perfect Random Number Generator II. Rapport de Recherche Laboratoire LJK Universite de Grenoble.

[19] GNEDENKO B.V., KOLMOGOROV A.N. (1968) Limit distributions for sums of independent random variables, Addison Wesley Publishing Co, Reading, Massachussets, London.

[20] MARSAGLIA G (1995) CD ROM. Florida State University, site internet http://stat.fsu.edu/pub/diehard/

[21] IBRAGIMOV I.A. LINNIK Yu. V.(1971) Independent and stationary sequences of random variables. Wolters-Noordhoff, Groningen.

[22] BRADLEY R.C. (1984) On a very weak Bernouilli condition. Stochastics, 13, 61-81.

[23] DEHLING H. DENKER M. PHILLIPPS W. (1984) Versik Processes and very weak Bernouilli processes with summable rates are independent. Proc. Amer. Math Soc. 91, 618-624.

[24] WITHERS C. S. (1981) Central limit theorems for dependent random variables. I. Z. Wahrsch. verw. Gebiete, 54, 509-534.

[25] COGBURN R. (1960) Asymptotic properties of stationary sequences. Univ. Calif. Publ. Statist. 3, 99-146.

[26] ROSENBLATT M. (1972) Uniform ergodicity and strong mixing. Z. Wahrsch. Werw. Gebiete. 24, 79-84.

[27] BERNSTEIN (1939) Quelques remarques sur le theoreme limite Liapounoff. Dokl. Akad, Nauk SSSR, 24, 3-8.

[28] BROWN (1970) Characteristics functions, moments and the Central Limit Theorem. Ann. Math. Statist. 41 658-664.

[29] ESSEEN C.G. JANSON S. (1985) On moments conditions for normed sums of independent random variables and martingales differences. Stochastics Processes and their Applications, 19, 173-182.

[30] HERRNDORF N. (1984) A functional Central Limit Theorem for ρ-mixing sequences. Journal of multivariate analysis. 15, 141-146.

[31] BIRKEL T. (1988) Moment bounds for associated sequences. Ann. Prob. 16-3, 1184-1193.

[32] KRUGOV V.M. (1988) The convergence of moments of random sums. Theory of Probability and its applications. 33-2, 339-342.

[33] MAIROBODA R. E. (1989) The Central limit Theorem for empirical moment generating functions. Theory of Probability and its applications, vol 34-2, 332-335.

[34] YOKOHAMA R. (1980) Moment bounds for stationary mixing sequences. Z Wahrscheinlichkeistheorie, ver. Gebiete, 52, 45-57.

[35] YOKOHAMA R. (1983) The convergence of moments in the Central limit Theorem for stationary ϕ-mixing processes. Analysis mathematica, 9, 79-84.

[36] COX D., KIM T.Y. (1995) Moment bounds for mixing random variables useful in non-parametric function estimation. Stochastics Processes and their Applications. 56, 151-158.

[37] IBRAGIMOV I. LIFSHITS M. (1998) On the convergence of generalized moments in almost sure central limit theorem. Stat and Prob Letters 343-351.

[38] SOULIER P. (2001) Moment bounds and the central limit theorem for functions of Gaussian vectors. Stat. and Prob. Letters 193-203.

[39] ROZOVSKY L.V. . (2002) An estimate of the remainder in the central limit theorem for a sum of independent random variables with infinite moments of a higher order. Theor. Proba. and its appl, 174-183.

[40] DOUKHAN P, LOUHICHI S. (1999) A new weak dependence condition and application to moments inequalities. Stochastics Processes and their Applications. (84) 313-342.

[41] HALL P. , HEYDE C.C. (1980) Martingale Limit Theory and Its Application, Academic Press, London

[42] SERFLING R.J. (1980) Approximation theorems of mathematical statistics. Wiley, New York

[43] JOHNSON N.L. KOTZ S. (1969) Discrete distributions. Wiley, New York

[44] JOHNSON N.L. KOTZ S. (1970) Continuous univariate distributions. Wiley, New York

[45] SANTHA M. , VAZIRANI U. V. (1984). Generating quasi-random sequences from slightly-random sources. Proceedings of the 25th IEEE Symposium on Foundations of Computer Science: pages 434440, University of California. ISBN 0-8186-0591-X.

[46] SANTHA M. , VAZIRANI U. V. (1986). Generating quasi-random sequences from semi-random sources, Journal of Computer and System Sciences, v.33 n.1, p.75-87.