



HAL
open science

Extension du concept de convergence en probabilité aux variables ordinaires

Vittorio Castellano

► **To cite this version:**

Vittorio Castellano. Extension du concept de convergence en probabilité aux variables ordinaires. Annales de l'ISUP, 1960, IX (4), pp.359-370. <hal-04092962>

HAL Id: hal-04092962

<https://hal.science/hal-04092962v1>

Submitted on 9 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

EXTENSION DU CONCEPT DE CONVERGENCE EN PROBABILITÉ AUX VARIABLES ORDINAIRES

Vittorio CASTELLANO

RESUME

Ayant défini l'ordre infini d'un point d'accumulation d'une suite et les points principaux d'accumulation ; l'auteur définit la convergence presque exclusive, la convergence d'une suite de nombres réels à un point-limite principal et il étend cette forme de convergence au cas d'une suite de distributions de nombres réels vers une distribution limite principale.

De même que la convergence ordinaire, ou exclusive à une seule limite, ainsi la convergence à une distribution limite principale peut être convergence "partout" ou "presque partout" selon qu'elle se vérifie pour tous les points de la distribution sans exception, ou bien pour tous les points excepté un ensemble de dimension zéro.

De ce mode très général de convergence vers un point-limite principal est la tendance vers une constante de la variable de Bernoulli, lorsque le nombre d'épreuves croît indéfiniment, qui aussi se réduirait sur le terrain de l'Analyse classique.

La "convergence en probabilité" de Fréchet, sous la forme explicite la plus avantageuse, est de même un cas particulier de la convergence "presque exclusive", qui étend aux variables de l'Analyse classique cette particulière "convergence vers une limite" des variables aléatoires, très efficacement étudiée, par Cantelli et d'autres après lui, avec le nom de convergence vers une limite au sens du calcul des probabilités.

I - L'OBJET DE CETTE ETUDE -

La variable de Bernoulli, que nous indiquons par \mathfrak{F}_n = fréquence des résultats favorables en n épreuves faites à probabilité constante P , nous a donné l'idée d'une convergence différente de celle de l'analyse mathématique, que l'on a nommée convergence "au sens du calcul des probabilités", ou "en probabilité", ou "stochastique", selon les différents auteurs.

Cette convergence est représentée par l'inégalité

$$\text{Prob} \{ | \mathcal{F}_n - P | < \varepsilon \} > 1 - \eta \quad (1)$$

valable dès que n est assez grand pour tout couple ε, η arbitrairement petits et positifs, mais si l'on calcule, pour chaque valeur déterminée de n , les résultats possibles \mathcal{F}_n satisfaisant l'inégalité

$$| \mathcal{F}_n - P | < \varepsilon$$

on peut écrire aussi, au lieu de la (1) :

$$\text{Freq} \{ | \mathcal{F}_n - P | < \varepsilon \} > 1 - \eta \quad (2)$$

Dans notre étude nous nous en tiendrons, généralement, à cette forme d'écriture qui, à la probabilité d'un événement, substitue sa fréquence dans l'ensemble des cas possibles, ou "univers des épreuves" (1).

Le but de notre étude est de démontrer que la convergence à une limite de la variable de Bernoulli appartient à un type général de convergence "presque exclusive", pouvant être définie aussi pour la suite des nombres réels ; elle peut être étudiée au moyen de l'Analyse mathématique, indépendamment du concept de variable aléatoire.

II - CONVERGENCE "PRESQUE EXCLUSIVE" D'UNE SUITE DE NOMBRES REELS -

Nous tâchons d'abord de transcrire (2) au cas de suites de nombres réels. Comme le \mathcal{F}_n est une variable, il faut transformer la suite des nombres réels :

$$a_1, a_2, \dots, a_n, \dots \quad (3)$$

en une suite de variables réelles

$$A_1, A_2, \dots, A_n, \dots \quad (3')$$

et nous pouvons faire cela en définissant

$$A_i = A_i(a_1, a_{i+1}, \dots)$$

(1) V. Castellano : "Sull'universo delle prove come fondamento della teoria dei campioni", 31e Session de l'Institut International de Statistique, Bruxelles, 1958.

comme l'ensemble de tous les termes de (3) de rang supérieur à i .

La transcription de (2) par (3) :

$$\text{Freq} \{ |A_n - a| < \varepsilon \} > 1 - \eta \quad (2')$$

signifie que, posant $K(n, n')$ = nombre des termes suivant de (3) :

$$a_{n+1}, a_{n+2}, \dots, a_{n'}$$

n'étant pas dans l'intervalle $a \pm \varepsilon$, pour un certain n assez grand, il y a l'inégalité

$$K(n, n') < (n' - n) \eta,$$

et, par conséquent, que

$$\lim_{n' \rightarrow \infty} \frac{K(n, n')}{n' - n} = 0$$

Nous exprimons ce fait en disant que l'ensemble de ces $K(n, n')$ points est infini d'ordre inférieur relativement à l'ordre infini de l'ensemble total que nous supposons = 1.

Comme $K(0, n') < n + K(n, n')$, on a aussi $\lim_{n' \rightarrow \infty} \frac{K(0, n')}{n'} = 0$ et par conséquent tous les termes de (3), en dehors de l'intervalle $a \pm \varepsilon$, constituent un sous-ensemble de (3) d'ordre infini < 1 en comparaison de l'ensemble (3).

Nous indiquons plus simplement par $K(n)$ le nombre des termes de cet ensemble parmi les premiers n termes de la (3).

Etant donné cela, nous disons que la suite (3) converge vers la limite "principale" a et nous écrivons :

$$\lim_{n \rightarrow \infty} p. a_n = a \quad (4)$$

s'il arrive que :

$$\text{Freq} \{ |a_n - a| < \varepsilon \} > 1 - \eta \quad (5)$$

c'est-à-dire si toute la suite appartient à l'intervalle $a \pm \varepsilon$, sauf un sous-ensemble infini d'ordre < 1 .

Si de $n > n_0$ assez grand

$$K(n, n') = 0$$

pour tout n' , on a

$$\text{Freq} \{ | a_n - a | < \varepsilon \} = 1 \quad (6)$$

en toutes les

$$A_{n_0, n'} = A_n - A_{n_0}$$

et a_n tend vers a dans le sens ordinaire de l'Analyse.

D'après les définitions que nous avons données, (5) définit un type de convergence qui diffère de la convergence considérée ordinairement par l'Analyse, à cause de l'existence de plusieurs points d'accumulation au lieu d'un seul point. Mais un seul de ces points a une valeur numérique du même ordre infini que la suite totale, et nous l'appellerons "point d'accumulation principale".

Si, par exemple, la suite des valeurs réciproques des nombres naturels est :

$$1, 1/2, 1/3, \dots, 1/n \quad (7)$$

celle-ci converge à la limite zéro.

La suite

$$1, 2, 1/3, 2^2, 1/5, 1/6, 1/7, 2^3, 1/9, \dots \quad (8)$$

a deux points d'accumulation : zéro et ∞ .

Dans les premiers n termes d'elle, il y a dehors de l'intervalle $(0, 1)$ un nombre $K(n)$ de termes égaux au maximum entier contenu en $\log_2 n$ et comme

$$\lim_{n \rightarrow \infty} \frac{\log_2 n}{n} = 0 \quad (9)$$

zéro est le point d'accumulation principale et (8) a zéro comme limite principale.

Les considérations précédentes sont valables pour tout ensemble d'un nombre fini de valeurs et, en général, pour toute variable

dont la distribution des fréquences est connue. Si nous supposons, en général, comme définition de convergence d'une variable X_n à une limite finie a , la (2).

$$\text{Freq } \{ | X_n - a | < \varepsilon \} > 1 - \eta \quad (10)$$

il s'ensuit immédiatement que : si une suite de variables réelles converge vers une limite déterminée et finie, cette limite est unique ou elle est principale, selon que

$$\lim_{n \rightarrow \infty} w_n \stackrel{?}{=} 0 \quad (11)$$

où w_n est le champ de variation de X_n .

III - OBSERVATIONS PRELIMINAIRES SUR LA REPRESENTATION D'UNE VARIABLE -

Avant d'appliquer les définitions, que nous venons de donner, à une suite de variables réelles, tendant vers une variable-limite, nous faisons quelques observations sur les représentations ordinaires des variables réelles.

Les variables réelles ne sont pas des ensembles ordonnés en soi, mais elles peuvent certainement être ordonnées selon l'intensité des valeurs qu'elles acquièrent. Cette distribution des valeurs est représentée, sur le plan carthésien, indifféremment par la courbe de graduation ou par la fonction de répartition, selon que sur l'axe des coordonnées on met les valeurs de la variable ou les fréquences des valeurs ne surpassant pas une certaine limite, seulement si l'on considère, évidemment, les fonctions de répartition en leur propre champ d'existence, au lieu de les considérer dans le domaine le plus vaste possible de $-\infty$ à $+\infty$, comme le font plusieurs auteurs.

Si la variable est donnée sous une forme continue, il faut toujours supposer l'ensemble comme infini, car la masse est distribuée en petits intervalles infinitésimaux dx .

Ayant convenu cela, la constante $X=a$ a comme fonction de répartition

$$F(X) = \begin{cases} 0 & \text{pour } x < a \\ 1 & \text{pour } x = a \\ 0 & \text{pour } x > a \end{cases} \quad (12)$$

tandis que la "presque constante" zéro représentée par la (7), à la fonction de répartition

$$F(X) = \begin{cases} \emptyset & x < 0 \\ 0 & x = 0 \\ 1 & 0 < x \leq 1 \\ \emptyset & 1 < x \end{cases} \quad \text{pour} \quad (13)$$

où le symbole \emptyset signifie : inexistant.

La suite (8) aussi, dont la fonction de répartition existe en tout le champ $x > 0$, représente zéro comme une "presque constante", mais sa fonction de répartition est :

$$F(X) = \begin{cases} \emptyset \\ 0 \\ 1 \end{cases} \quad \text{pour} \quad x \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases} \quad (14)$$

Nous indiquons maintenant par $X(F)$ la variable ordonnée X , c'est-à-dire sa courbe de graduation, et par $F(X)$ sa fonction de répartition. Entre les fonctions correspondantes, que par commodité nous écrivons avec les notations du cas continu, qui inclut, par les conventions ordinaires, le cas discontinu, $x(F)$ et $F(x)$, on a la relation

$$x(F) = F^{-1}(x) \quad (15)$$

IV - CONVERGENCE "PRESQUE EXCLUSIVE" D'UNE SUITE DE VARIABLES REELLES -

Nous pouvons maintenant appliquer la (5) au cas général de la convergence d'une suite de variables réelles à une variable-limite principale.

Définition :

Nous disons que la suite de variables réelles $X_i (i=1, 2, \dots, n, \dots)$ a la variable X comme limite presque exclusive ou principale, si, étant donné un $\varepsilon > 0$ pour tous les X_i , sauf un sous-ensemble infini d'ordre < 1 , l'inégalité

$$| X_i(F) - X(F) | < \varepsilon \quad (16)$$

est satisfaite pour chaque valeur de F , sauf tout au plus un ensemble infini de ces valeurs, de dimension nulle.

Si le sous-ensemble de variables X_i , qui fait exception, est fini ou nul, la limite est unique et la convergence est ordinaire.

Soit la convergence ordinaire, soit la convergence principale peuvent avoir lieu "partout" ou "presque partout", selon que (16) est valable pour toutes les valeurs de F sans exception, ou avec l'exception d'un ensemble de dimension nulle.

La convergence presque exclusive partout est, par conséquent, définie par la même inégalité

$$\text{Freq} \{ |X_n(F) - X(F)| < \varepsilon \} > 1 - \eta \quad (17)$$

définissant la convergence des suites de nombres réels, la fréquence $1 - \eta$ étant relative à l'ensemble des n premiers X_1 de la suite des X_1 , pour $n > n_0$. Dans la convergence presque exclusive, presque partout, la fréquence $1 - \eta$ est relative à l'ensemble des X_1 , et pour un certain n assez grand, pour chaque X_1 satisfaisant (16), à l'ensemble des valeurs de F . Quand il y a une possibilité d'équivoque, il faut distinguer les différents cas, en mettant l'indication de la variable en question en bas, avant les parenthèses, c'est-à-dire qu'il faut écrire :

$$\text{Freq}_n \quad \text{ou bien} \quad \text{Freq}_{n,F} \quad \text{au lieu de} \quad \text{Freq}.$$

La notion Freq_F devrait indiquer la convergence ordinaire presque partout.

Il est facile de construire des exemples de suites de variables montrant ces types de convergence à une limite seule. Si, par exemple, X_n est constitué par l'ensemble des points de l'intervalle $\pm(1 + a_n)$ avec densité

$$y = C_n \sqrt{(1 + a_n)^2 - x^2}$$

la suite des X_n converge partout, à la variable, dont la distribution de la fréquence est

$$y = C \sqrt{1 - x^2}$$

exclusivement ou presque exclusivement, selon que les a_n sont des valeurs de la suite (7) ou de la (8).

D'une manière analogue, la distribution normale :

$$y = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-M)^2}{2\sigma^2}} \quad (18)$$

car

$$\lim_{\sigma \rightarrow 0} y = \begin{matrix} 0 \\ \infty \end{matrix} \quad \text{pour } x \begin{matrix} \neq \\ = \end{matrix} M$$

définit la variable normale tendant vers la presque constante $X=M$, définie dans tout le champ $\mp \infty$, exclusivement ou principalement, selon que l'on donne à σ_1 des valeurs d'une suite convergeant vers zéro comme (7) ou comme (8). Mais si nous considérons la constante M pour chaque σ suffisamment petit, il en résulte :

$$\text{Freq } \{ |X_\sigma - M| < \varepsilon \} > 1 - \eta \quad (19)$$

et par conséquent on peut dire aussi que la variable normale X_σ converge "presque partout" vers la moyenne M , de quelque façon le σ tende vers zéro. Particulièrement la variable de Bernoulli \mathcal{P}_n converge au sens ordinaire, avec l'augmentation de n , vers la presque constante P , définie dans l'intervalle $(0, 1)$.

Des considérations faites il s'ensuit que si nous mettons

$$\omega(X_n, X) = \max |X_n(F) - X(F)| \quad (20)$$

la condition nécessaire et suffisante, pour que X_n converge au sens ordinaire vers X , est que :

$$\lim_{n \rightarrow \infty} \omega(X_n, X) = 0 \quad (21)$$

et que, par conséquent, l'inégalité

$$\lim_{n \rightarrow \infty} \omega(X_n, X) > 0 \quad (22)$$

caractérise, avec (17), la convergence partout vers une limite principale, et la convergence presque partout vers une limite exclusive ou principale.

Nous observons que la convergence ordinaire, "presque partout" est une caractéristique des variables aléatoires théoriques, sujet du calcul des probabilités, et que les variables empiriques correspondantes convergent - par la loi empirique du cas - vers la valeur limite théorique, mais seulement "principalement". Par exemple, la distribution d'un nombre suffisamment élevé K de déterminations empiriques indépendantes de la variable de Bernoulli, faites chacune sur n épreuves, a pour $n \rightarrow \infty$, la "presque constante" P comme limite principale, car pour chaque valeur de n , ces K déterminations peuvent même toutes présenter des valeurs bien distantes de P , bien que ce fait soit absolument exceptionnel.

V - CAS PARTICULIER DES VARIABLES DEPENDANTES -

Nous considérons maintenant, d'après les résultats précédents, la signification de l'inégalité

$$\text{Prob} \{ | X_n - X | < \varepsilon \} > 1 - \eta \quad (23)$$

prise par Fréchet comme la définition de la convergence en probabilité.

Si X_n , et X sont indépendantes, X est nécessairement une constante ou une presque-constante.

Tandis que s'il y a une connexion entre les valeurs des deux variables, alors une partie seule des valeurs de X_n peuvent être associées à un x déterminé, et si l'on remplace aussi Prob par Freq, comme nous avons fait à (1), (23) devrait être écrite plus correctement de la manière suivante :

$$\text{Freq} \{ | X_n(x) - x | < \varepsilon \} > 1 - \eta \quad (24)$$

pour x variable en tout le champ d'existence de X .

Cela signifie que pour chaque x constant, la variable $X_n(x)$ converge vers la constante certaine x . C'est-à-dire $X_n(x)$ est de la forme

$$X_n = x + y_n(x) \quad (25)$$

et en définitive

$$X_n = X + Y_n(x) \quad (25')$$

où $Y_n(x)$ représente un ensemble de fonctions du paramètre x tendant toutes vers zéro quand n tend vers l'infini.

On peut alors l'écrire aussi :

$$\text{Freq} \{ | Y_n(x) | < \varepsilon \} > 1 - \eta \quad (26)$$

et ainsi que l'équivalente (24), elle rentre comme cas particulier dans la forme plus générale (17) de la convergence presque exclusive, que Fréchet appelle "convergence légale", et que nous pensons qu'elle puisse exprimer le concept essentiel de la convergence en probabilité.

VI - CONCLUSIONS -

A plus de 30 ans de distance de la 1ère édition du "Calcul des Probabilités" par Castelnovo, où pour la première fois on affirme clairement que le théorème de Bernoulli ne peut pas remplacer la loi empirique du cas dans la légitimation des applications du calcul des probabilités, les auteurs de traités, qui s'adonnent à cette science, continuent à chérir - du moins chez les autres - l'illusion, pas encore déracinée, que le calcul des probabilités étudie les variables aléatoires empiriques. Cela donne au calcul des probabilités l'appât dangereux des choses miraculeuses et impose au statisticiens la mise en lumière des anneaux de passage entre les théories abstraites, la statistique théorique et les applications que les mathématiciens peuvent même négliger.

Si la connexion de la théorie mathématique des probabilités avec les applications est constituée par la loi empirique du cas, l'anneau de passage, pas aperçu toujours, entre le calcul des probabilités abstrait et la statistique théorique, est, selon nous, l'opération aléatoire définissant en bloc et "a priori" la variable aléatoire comme un ensemble de résultats équivalents possibles. La distribution de ceux-ci, comme elle est représentée par la courbe de graduation ou par la fonction de répartition, ou bien par la fonction de densité, définit parfaitement la variable aléatoire, et deux variables aléatoires sont égales, si elles ont la même fonction de répartition. L'opinion contraire de Fréchet est justifiée par la considération de deux variables aléatoires, ayant la même fonction de répartition, qui présentent de différentes déterminations associées, comme, par exemple, il arriverait si elles étaient définies par l'opération de jet d'un seul dé et ensuite par la lecture de deux faces différentes déterminées.

Mais en ce cas, les deux variables sont dépendantes et l'opération aléatoire définissant l'une des deux, Y comme dépendant de l'autre, X , comprend comme composante essentielle l'opération $\vartheta(X)$ qui définit le X , selon le schéma

$$\vartheta(Y) = \vartheta(X). \vartheta(x/Y)$$

et l'on ne peut supposer sa fonction de répartition sous la forme de $F(Y)$, mais sous la forme de $F(x/Y)$.

Rien n'empêche de définir égales deux variables Y et X associées et d'écrire :

$$Y(x) = X \quad \text{ou} \quad X(y) = Y$$

quand

$$Y(x) = x$$

ou

$$X(y) = y$$

il serait peut être mieux d'écrire

$$X = Y ;$$

mais cela ne signifie point que c'est impossible de définir égales deux variables simples indépendantes et d'écrire :

$$X = Y$$

quand les fonctions de répartition de celles-ci sont égales.

Il est évident que deux variables simples égales ne peuvent pas être coïncidentes et afin qu'elles deviennent associées, il faut introduire de nouvelles conditions qui ne sont pas nécessaires aux variables indépendantes.

L'opinion de V. Mises semble donc être justifiée, c'est-à-dire que les variables aléatoires sont tout à fait définies par leurs fonctions de répartitions, à condition, évidemment, que les variables doubles soient interprétées comme définies par la fonction de répartition double et ainsi de suite.

L'avantage d'une proposition du Calcul des Probabilités, qui soit fondée explicitement sur l'ensemble des résultats possibles de l'opération aléatoire, ou "univers des épreuves" (1), réside, selon nous, dans le fait qu'elle pourvoit la Statistique théorique d'un modèle concret des phénomènes empiriques, sujets à des variations non contrôlables une à une, et encore dans le fait qu'elle pose clairement les limites infranchissables du modèle mathématique, qui ne peut représenter la suite des phénomènes empiriques obéissant aux lois du cas - ou que l'on suppose qu'ils obéissent - mais elle représente seulement les ensembles constituées par ces suites et par leurs parties continues.

Dans notre étude, nous présentons la convergence des suites vers une limite principale comme le modèle mathématique de la convergence vers une limite des collectives empiriques.

(1) V. Castellano : "Statistica e Matematica", Discours d'ouverture de XVIIIe Réunion de la Société Italienne de Statistique, Roma, 1958.

