



**HAL**  
open science

## How can histograms be useful for introducing continuous probability distributions?

Charlotte Derouet, Bernard Parzysz

► **To cite this version:**

Charlotte Derouet, Bernard Parzysz. How can histograms be useful for introducing continuous probability distributions?. ZDM, 2016, 48, pp.757 - 773. 10.1007/s11858-016-0769-9 . hal-03618056

**HAL Id: hal-03618056**

**<https://hal.science/hal-03618056>**

Submitted on 24 Mar 2022

**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.

## How can histograms be useful for introducing continuous probability distributions?

Charlotte DEROUET

Bernard PARZYSZ\*

Laboratoire de Didactique André Revuz

LDAR (EA 4434)

Université Paris Diderot et \*Université  
d'Orléans

### Abstract.

The teaching of probability has changed a great deal since the end of the last century. The development of technologies is indeed part of this evolution. In France, continuous probability distributions began to be studied in 2002 by scientific 12<sup>th</sup> graders, but this subject was marginal and appeared only as an application of integral calculus. With the high school reform recently implemented (in 2012 for grade 12), continuous probability distributions now have an important place in the scientific section. As induced by official texts, the use of histogram as a link between descriptive statistics and continuous probability distributions through its analogy with density curves, is indeed a promising path for learning. In order to answer this demand, which is indeed of general interest, we wondered how density function could be introduced for such a purpose. We began with a study of how textbooks deal with that question, which led us to propose an alternative introductory classroom activity. The reason for this proposition is that, contrary to what is currently implemented in textbooks, a specific mathematical working space (MWS) has to be brought into play, making various mathematical domains intervene in turn and articulating them with each other.

Key words: Probability. High school. Continuous distribution. Density. Mathematical Working Space.

### 1. Introduction

In French high school mathematics curricula as well as in many other countries, the probabilistic and statistical contents have been notably enlarged for some years. The reasons for including probability and statistics in teaching have been highlighted over the past 30 years (*e.g.* Holmes 1980; Hawkins et al. 1991; Vere-Jones 1995; Wild and Pfannkuch 1999; Gal 2002). There are nowadays two main aims that account for this: probability and statistics on the one hand get nearer to uses in social and professional life (Batanero et al. 2004a), and on the other hand, take advantage of the new possibilities linked with technologies that are now at students' disposal. This evolution has conveyed, for many teachers, a necessary update of knowledge in the domain, which has not always been very easy or well felt (Lattuati et al. 2012). In any event, new working spaces had to be worked out and brought into play in classrooms, most frequently associating several mathematical domains (Parzysz 2011). In particular, in the final year of high school (12<sup>th</sup> grade), several continuous probability distributions are now currently part of various tracks in high school, in particular the scientific one. The necessity for introducing such distributions at high school

level may rightfully be questioned, and several voices have been raised to propose “sending back to university [...] the study of the normal distribution, in order to focus in high school on discrete distributions (binomial, Poisson, geometric...) and use this probabilistic knowledge to make statistics too”<sup>1</sup> (Perrin 2015, p.63). But, either in high school or later, in most countries students will have to study continuous distributions, thus the question of their introduction remains. The understanding of normal distribution by students at university (Batanero et al. 2004a, 2004b) or by secondary teachers in an in-service program (Bansilal 2012, 2014) has been studied.

This article focuses on the introduction of probability distributions of continuous random variables, i.e., the notion of probability density. This research has two main goals:

- to analyse in which way French mathematics textbooks currently introduce density functions as well as histograms;
- to suggest an alternative introduction of these notions.

We bring to the fore the corresponding working spaces and suggest paths for its implementation in classrooms. More precisely, we begin with studying how the notion of density function is tackled in the eight French textbooks designed for scientific 12<sup>th</sup> graders<sup>2</sup>, and then we propose some ideas for introducing it at this level, in order to give it sense. Even if our research is based on French textbooks, most countries may of course be concerned with our research questions, since our proposal intends to solve a quite general teaching and learning problem, namely the introduction of the notion of density function and the use of histograms for that purpose.

## 2. Theoretical background

We locate our research within the framework of Mathematical Working Spaces (MWS), first proposed by Kuzniak (2011) as an extension of Geometrical Working Spaces (Houdement and Kuzniak 2006, Gómez-Chacón and Kuzniak 2015). The aim of this model is to describe the work put in play by people solving mathematical tasks. Here the word “work” means not only mathematical knowledge, but also cognitive activities. In this article we present only the elements of this framework useful for this article. The reader will find more details in the introductory article of the special issue.

The MWS framework (fig. 1) distinguishes two planes for mathematical work, epistemological and cognitive, which interact with each other (Kuzniak 2011).

The former is made up of three components:

- a set of signs: semiotic representations (Duval 1995) or *representamen* in Peirce’s sense<sup>3</sup>,
- a set of artefacts (instruments) designed to perform actions on these objects,
- a theoretical reference system.

---

<sup>1</sup> The translations of all French quotations are ours.

<sup>2</sup> In France, textbooks are an important help for teachers, who generally make use of several of them.

<sup>3</sup> In Peirce’s theory a sign is composed of three elements : a representamen (sign vehicle), an object (what is represented) and an interpretant (translation of the sign). For more details, see for instance the work of Atkin (2013).

The latter is composed of three processes: recognition (visualization<sup>4</sup>), construction and argumentation (proof). These two planes are articulated according to three dimensions:

- *semiotic*, based on the production of representations (in relation to semiotic representation registers of Duval 1995) and visualisation;
- *instrumental*, related to constructions determined by the instruments used (*e.g.* the use of software);
- *discursive*, linked to the theoretical reference system and the proof.

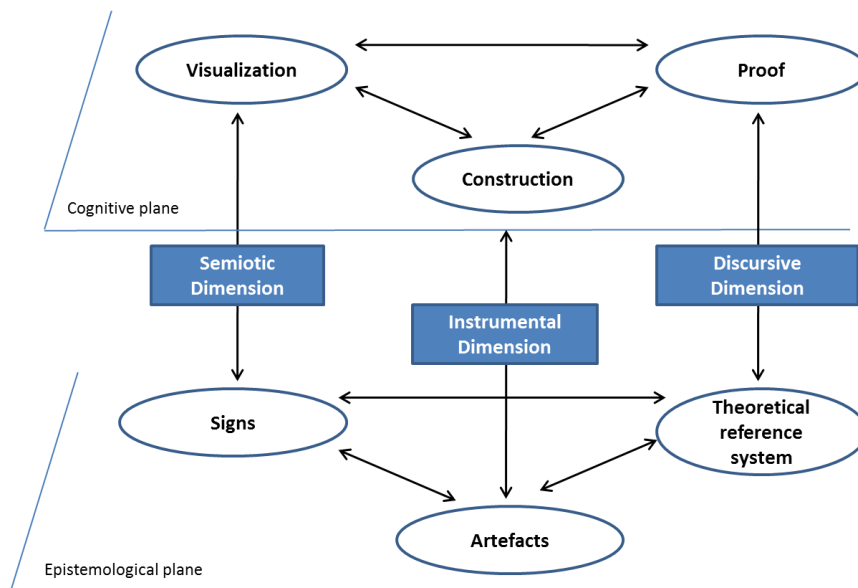


Figure 1. The MWS model (after Kuzniak 2011)

The associated hypothesis is that articulating various dimensions of MWS between them favours understanding mathematical knowledge.

Three levels of MWS are distinguished:

- reference MWS, describing the work institutionally aimed at ; this MWS is partly described at national level by texts (syllabi and other official documents);
- suitable MWS, describing the work planned by teachers to allow an effective setting in classrooms;
- personal MWS, depending on the student.

In this article our interest lies in the first two levels (reference and suitable MWS); in our case, the suitable MWS is as presented in textbooks.

Montoya and Vivier (2014) proposed a conception of MWS associated with a mathematical domain. The structure of the domain must be consistent from mathematical and epistemological points of view. Hence one may speak of geometrical MWS, algebraic MWS and so on. In the present article we shall meet with several mathematical domains: probability, calculus and statistics. We also consider a subdomain of calculus, namely integral

<sup>4</sup> Here this term does not mean just seeing, but processing and structuring the information brought by the signs (Kuzniak & Richard 2014).

calculus (IC), together with two subdomains of statistics: descriptive statistics (DS) and inferential statistics (IS). The global MWS coordinating the MWSs related to various (sub)domains can then be taken into account, as we shall see in this article.

Concerning the domain of probability we do not speak of subdomains but of paradigms. As redefined by Kuzniak (2011), who took it from Kuhn (1962), a paradigm is a set of beliefs, techniques and values shared by a scientific group. In his research work, Parzysz (2011) identifies several paradigms referring to the domain of probability:

- a ‘pseudo-concrete’ probabilistic paradigm (P1) constituting a first theoretical step starting from real situations, which is not very formalized (Henry 2001);
- a more formal probabilistic paradigm (P2) based on the notion of finite probability space, the theoretical horizon of which is an axiomatic system of the Kolmogorov type (P3 paradigm).

For the purpose of this research we needed to distinguish between a standard version of the P2 paradigm limited to discrete probability, and an extended one (P2+) including continuous probability. In this paper only P2+ will be used.

We think that changes of (sub)domains, especially between DS, P2+ and IC, are a valuable way to introduce density function but the question is how these changes are coordinated. The main aims of this work are to study the shifts of (sub)domains occurring when introducing the notion of probability density with histograms, to understand the transfer dynamics between the (sub)domains during introductory tasks and determine through which dimensions of the MWS such changes occur. MWS is quite relevant for these aims, because it enables one to consider within a single, global framework the changes of domains and the various semiotic registers; moreover, it also allows the detection of the cognitive processes ruling them.

### **3. Research methodology**

We began our research (§ 4) by studying the reference MWS through the official texts defining the syllabus for scientific 12<sup>th</sup> graders (MENJVA 2011), together with the documents (‘resource’ documents) intended to help teachers (MJENR 2002, MENJVA 2012). Then (§ 5) we studied the potential suitable MWS at play in the eight French textbooks published for this level (see list in Appendix 1).

We focused our attention on the preliminary activities of the chapter on density distributions, and more especially on the ones which dealt explicitly with the notion of density. The activities were studied, question after question, and we spotted the changes of (sub)domains, particularly the mathematical objects symbolising them. Then this information was transferred on to a diagram stressing the shifts between the various domains involved in the activity. Moreover, we took an interest in the dimensions of the MWS put into play in these shifts. This also led us to identify possible difficulties, or even mathematical errors, in the activities.

In order to make things clear it may be useful to detail the use of the framework with an example. In one textbook [G12T6], the introductory activity (p. 377) is shown in Appendix 3. It begins with a statement and we can notice that a graph is given: the histogram of the size of a sample of 50,000 men. There is a reminder that the area of a rectangle corresponds to the frequency on the interval. So, in question 1, when the student has to estimate a proportion (or

a frequency), a calculation of some areas is required. Hence we are now definitely within the domain of descriptive statistics. In the second question a shift of domain occurs: the curve (of the density function) is introduced as that of a function which corresponds with the histogram. The textbook assumes that “*for any large-size sample a histogram close to that curve would be obtained*”. The statement, together with the graph, show that we have now moved to the probability domain. To answer question 2 (determine a probability), it is necessary to evaluate the area below the curve, and hence calculate an integral: hence a second change of domain. The shift from the relative frequency associated with the area of a part of the histogram (area 1) to the probability associated with the area under the curve on the same interval (area 2) is not taken into account. Table 1 summarizes this information, and the diagram (fig. 2) is a means for understanding the chronology of the activity and the different domains met with. The mathematical work is mainly done on various signs (semiotic dimension) through changes of registers: for example shifting from the graphic register to the numeric register to get the frequency. Similarly, for the move from SD to P2+, the shift from histogram to curve is guided by visualization.

Part of the task	Mathematical domain	Mathematical object	Means
Introductory statement	DS	Histogram (given)	
Question 1	DS	Proportion/frequency	
Question 2	P2+	Curve (given)	
	P2+	Probability	
	IC	Integral	

Table 1. Example of use of the framework for the study of a set of tasks

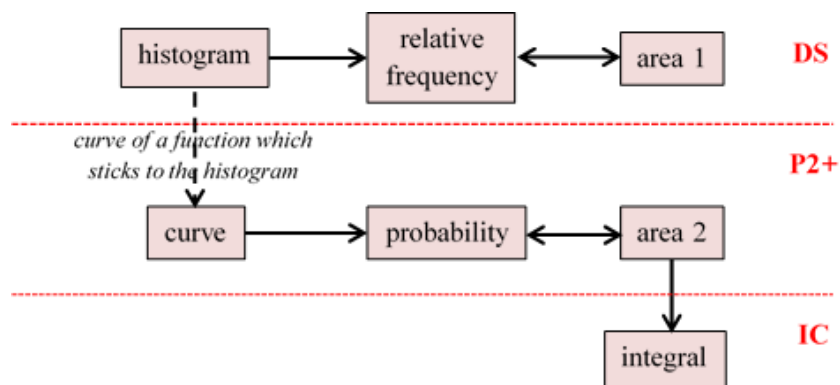


Figure 2. Diagram of the process of the activity proposed in [G12T6] (p.377)

After this study (§ 6), we developed an interest in histogram, which is used in most of the 12<sup>th</sup> grade textbooks for introducing the notion of density. This specific type of graph is explicitly taught at 10<sup>th</sup> grade level, so we studied the eleven textbooks for 10<sup>th</sup> graders (see Appendix 2) on that point, from a mathematical point of view, in order to have an insight into the knowledge of 12<sup>th</sup> graders about histograms and to try to explain some of the errors in their

textbooks. Then we based ourselves on our results to conceive a classroom situation about probability density aiming at making up for the weaknesses spotted in the textbooks. The former MWSs were enriched, and sometimes modified, in order to find a suitable MWS helping the students understand the notion of density, and fostering a renewed interest in the study of histograms. The proposed situation was experimented in classrooms, but we shall not discuss this aspect in this article.

#### 4. Reference texts

The current syllabus for French scientific 12<sup>th</sup> graders, in use from September 2012 on, recommends that teachers start introducing the notion of continuous random distribution with examples (MENJVA 2011, p. 12). It mentions three types of distributions: uniform, exponential and normal, but gives no indication as to how they should be presented, nor as to how the notion of density could be tackled (in particular: does this notion have to be limited to the above distributions or extended to any continuous random variable?). However, in the resource document put online for the 12<sup>th</sup>-grade teachers (MENJVA 2012), the normal distribution, introduced from the binomial distribution (de Moivre-Laplace theorem), deals, for this particular case, with the question of shifting from a discrete distribution to a continuous one, and thus it is concerned with the notion of density (although without naming it).

The corresponding process is schematized in fig. 3:

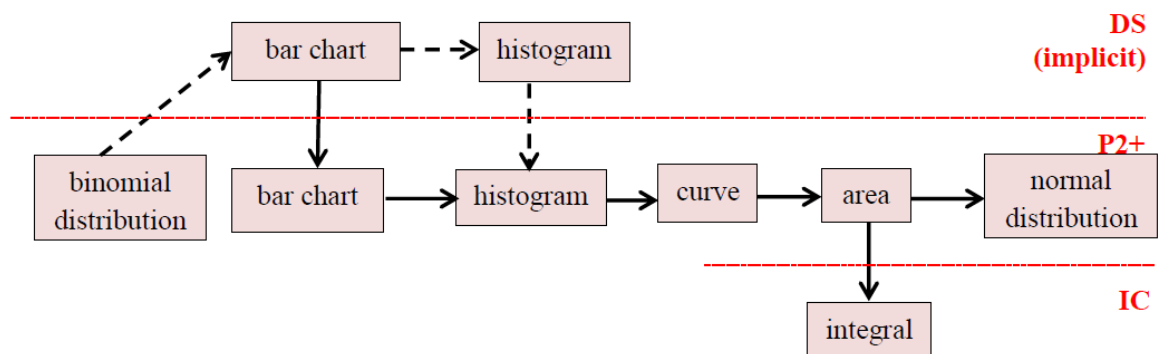


Figure 3. Diagram of the process of the activity proposed in the document (MENJVA 2012)

In the previous syllabus, which did not include normal distributions, an appendix to the 2002 resource document dealt with the question in general and suggested undertaking a modelling process based on a statistical series corresponding to a continuous variable (a sample of 50,000 heights of adult men): a histogram of relative frequencies (with equal intervals) was made, then a model approximating the histogram was sought for, i.e., “*a function  $f$ , the curve of which fits the histogram, the area under the curve being equal to 1*” (MJENR 2002, p. 143).

Contrary to the previous example, here the process appears clearly: at the beginning one is in DS, and the modelling issue makes one shift to P2+, according to the following schema (fig. 4):

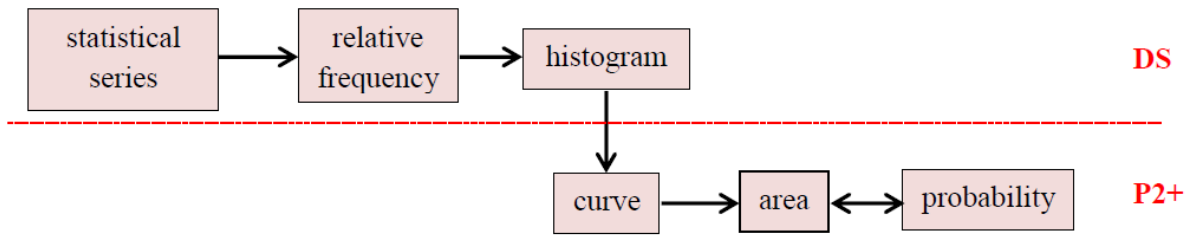


Figure 4. Diagram of the process of the activity proposed in the document (MJENR 2002)

Since we found the idea of using histograms in a process introducing the density of a continuous random variable interesting, we were willing to see which knowledge could be called up by 12<sup>th</sup>-graders about this subject.

The histogram is explicitly part of the syllabuses from 7<sup>th</sup> to 10<sup>th</sup> grades, and the resource document for middle school states that “*the syllabus specifies that the examples studied are limited to the case of classes having the same width*” and “*then a histogram can be read like a bar chart*” (MEN 2007, p. 6). In high school, the resource document for 10<sup>th</sup> grade reminds one that “*histograms, graphic representations, [...] are all tools which can be used by students to make questions about data arise, draw out study issues or sum up essential information...*” (MEN 2009, p. 4). This document also mentions “*the unavoidable help of computing tools*” (p. 3). So histograms appear very soon in secondary education, although under a peculiar form that does not make them different from bar charts, and the only work asked from students is—in the best case—to be able to extract numerical information from them. Thus this type of graph is not really emphasised and, if it happened to be used in 12<sup>th</sup> grade, a new ‘first encounter’ (in Chevallard’s sense, 2007) would have to be organised.

### 5. Suitable MWS proposed in 12th grade textbooks

We decided to investigate how the notion of probability density function is currently introduced in the eight textbooks for scientific 12<sup>th</sup> graders<sup>5</sup>, more especially in order to see if the process recommended by the 2002 resource document is taken up by some of them. For this study we consulted eight textbooks. All of them choose to deal with continuous uniform and exponential distributions before normal distributions. Thus probability density functions are met with before the introduction of normal distributions, and integral calculus is a prerequisite.

Four different approaches may be distinguished in their introduction of the probability density function<sup>6</sup>:

a) One textbook ([G12T8]) gives the definition first, without posing the problem:

*“In the case of a discrete random variable, the associated probability distribution [...] can be given by way of a general formula [...]. In the case of a continuous random variable a function defined on  $\mathbb{R}$ , called density, is defined.”* (pp. 394-395).

<sup>5</sup> See list in Appendix 1.

<sup>6</sup> Four textbooks deal with normal distributions in the same chapter as other continuous distributions, while in the other four it constitutes a specific chapter.



b) Another textbook ([G12T2]) uses the example of target shooting. A random variable is introduced: the abscissa (contained between 0 and 1) of the final position of a dart when shooting on special targets consisting in parts of the Euclidian plane contained between a curve<sup>7</sup> and the abscissa, the areas of which are equal to 1 (fig. 5).

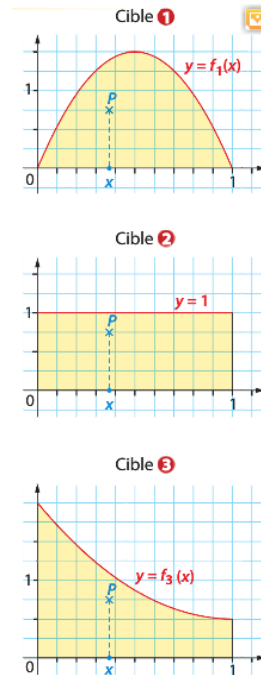


Figure 5. Introduction of probability density functions in [G12T2] (p.364)

Then the probability of events  $x \in [a; b] \subset [0; 1]$  is asked, without explicitly defining any random variable (although implicitly it is uniform). This leads to the introduction of the notion of density function as the function represented by the curve, and to conclude that “*the function, the curve of which is the upper edge of the target, is called **probability density** of the distribution defined in this way on  $[0; 1]$* ” (p. 364). This process is not satisfactory, since on the one hand it is ‘unnatural’ and, on the other hand, a great deal is implied, as for example the reason why the area of the surface under the curve has to be equal to 1 and why the distribution of the points of impact is uniform.

c) A third textbook ([G12T4]) uses a cumulative relative frequency polygon to introduce at first the cumulative distribution function (which does not belong to the syllabus) and then the frequency histogram, considering it as the ‘derivative’ of the cumulative frequency polygon.

d) The other five textbooks have a common approach, based on a relative frequency histogram with equal classes, in order to approximate it by a continuous curve in the course of the process. We present now three significant examples of textbooks, all of them assuming that students know how to integrate a positive function on a limited interval.

1) The first example is the activity of textbook [G12T6], that has been analysed in part 3. It is considered “*the curve of a function which ‘hugs’ the histogram*”. The work is based on the semiotic dimension.

<sup>7</sup> Three curves are considered in turn.

2) In order to refine the histogram, [G12T7] reduces the width of the intervals (dividing it by 10). Then (fig. 6) “a curve appears, similar to the one drawn on the graph. This curve represents a continuous function on  $[0; 6]$ , called **probability density** of the distribution of  $X$ .” (p. 322).

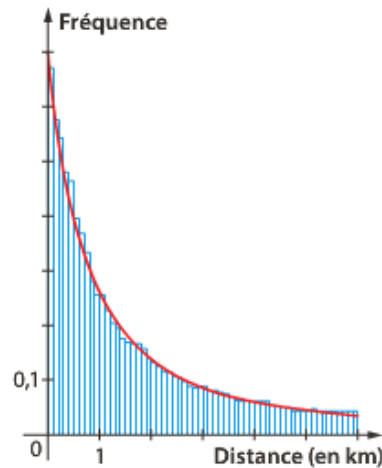


Figure 6. Histogram and associated density curve in [G12T7] (p. 322)

One may observe that, here as well, no random experiment which would allow the shift from relative frequency to probability is defined (fig. 7). Moreover, in the first part of the activity area 1 is assimilated to probability, which causes an unquestioned confusion between DS and P2+. Here as well the shifts of subdomains are taken into account by the semiotic dimension.

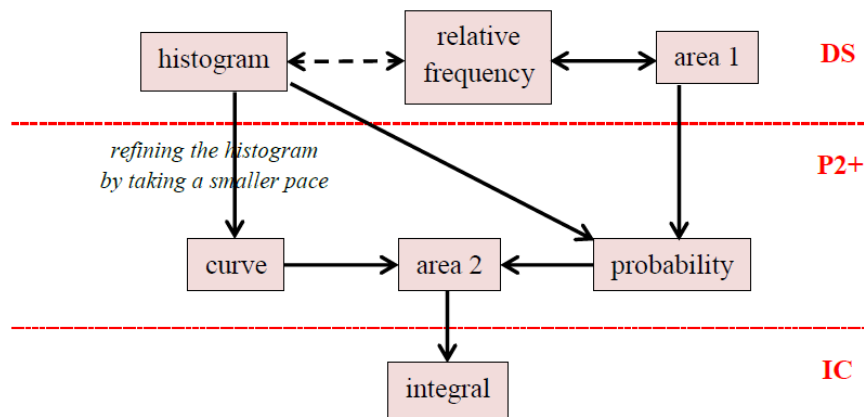


Figure 7. Diagram of the process of the activity proposed in [G12T7] (p. 322).

3) In the first activity, putting the stress on the instrumental dimension of MWS, [G12T5] uses a simulation—with spreadsheet—of a uniform random variable on  $[0; 10]$ , together with a ‘histogram’ given by the software<sup>8</sup>. Here each bar represents the relative frequency of an interval; this explains why the abscissae are named with intervals rather than with numbers. Here we find the kind of process proposed in the 2012 resource document for this level, which ‘takes advantage’ of the ambiguousness of the software naming ‘histogram’ a bar chart and allowing one to change the width of the bars. In the present case the variable is *really*

<sup>8</sup> Since the software does not provide histograms but only bar charts, one has to ‘cheat’ by widening the bars to make them contiguous

continuous—however this process betrays the difficulty of handling histograms. On the graph is also shown a “red curve, called ‘trend line’, which ‘smoothens’ the histogram” (p. 398) and which in fact represents the density of the uniform distribution. Here as well, the shift from relative frequency to probability is not defined.

The second activity is about the simulation—referring to *Geogebra* but not using it—of a sample of an exponential model of the operation time  $T$  of an electronic device. By gathering the data in classes, students are asked to estimate probabilities of  $P(X \leq a)$  and  $P(a \leq X \leq b)$  types, which are then compared with the integral of  $f$  between the same limits. Thus the process consists of starting from a probability density—although without defining or naming it—introduced in the software to simulate a continuous distribution generating a histogram corresponding to a particular sample which is superimposed on the density curve. Then a relative frequency is calculated, graphically interpreted as the area of a part of the histogram (area 1) and considered as an approximation of a probability. To end with, since the histogram approximates the density curve, this probability is identified as the area under the curve on the same interval (area 2), according to the following diagram (fig. 8).

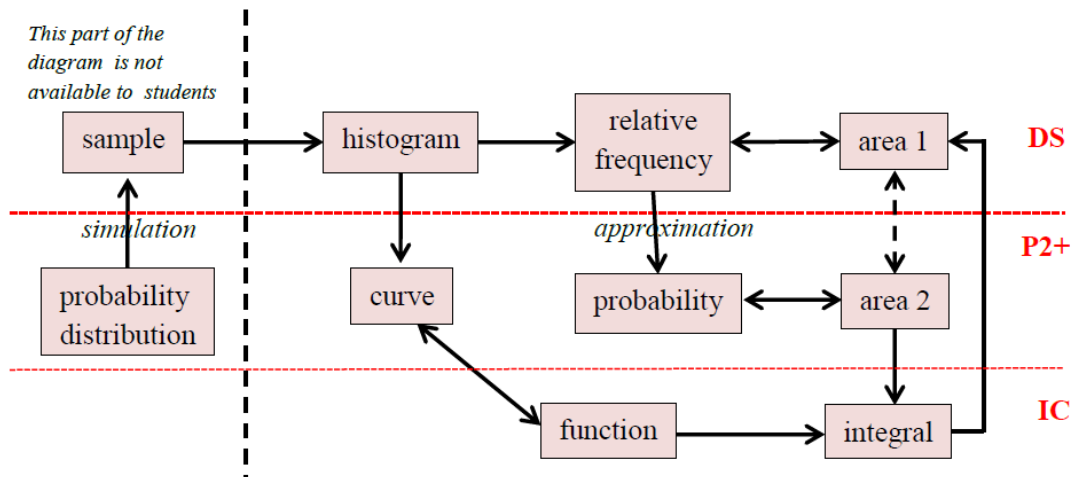


Figure 8. Diagram of the process of the second activity proposed in [G12T5] (p. 399)

In this case, the purpose of using simulation is to make the probabilistic situation more experimental. Simulation replaces the real data, which have to be modelled (Biehler 1991).

During the course of this process, the MWS brings successively into play the subdomain IC (continuous function) then DS (histogram), although in a way totally invisible to students. Then one moves to P2+ (probability of an event) and finally goes back to IC (integral calculus). The game of approximations is rather subtle—if not somewhat specious—for, starting from the probability distribution (P2+), a detour through the histogram associated with a large size sample (DS) is undertaken and the analogy between graphic representations is used to justify the use of an integral (IC) to calculate a probability. As we already did (Parzys 2011), we may also notice that—although in a less blatant way than in the previous examples—the transitions between P2+ and DS are not highlighted, as if the analogies brought into play were self-evident. This is however an essential element of the process, since a question is how to justify the use of a continuous model to account for a random phenomenon which is perceived

through discrete measures. In both cases the semiotic and instrumental dimensions are involved.

We may also notice that, in these three examples—but in the other textbooks as well—the shift to DS is operated from simulated large size samples and not from real data; thus the histogram is very close to the curve. This makes the approximation by the density curve easier, the latter being in fact subjacent to the simulation through which the data were obtained; but then the justification is limited to the semiotic dimension and the discursive dimension is lacking.

Only three out of the five textbooks which use a histogram to get to the density curve produce a (quick) reminder of this object. This leads one to think that this graphic representation is considered being fully part of the students' personal MWS. However none of them gives a correct representation of the histogram: as we already pinpointed about the [G12T5] textbook, but also noticeable in [G12T1] (fig. 9), the histogram is in fact a bar chart with wide bars and the relative frequency is read on the ordinate axis, hence not as the area of a rectangle.

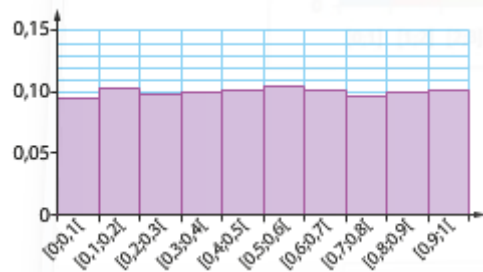


Figure 9. Example of ‘histogram’ being in fact a bar chart ([G12T1], p.378)

In other textbooks, in which the representation looks correct, we may also notice some mistakes in the name of the ordinate axis, sometimes called ‘relative frequency’, ‘percentage’... All these remarks converge towards contradicting the fact that histograms are available in the students' personal MWS, since they are not in the suitable MWS proposed by textbooks. In none of them is the word ‘density’ justified, or even approached; it is just a name given to a function, seeming to have no particular meaning.

To put it in a nutshell, using the relative frequency histogram to get to the density function was chosen by a majority of textbooks. However we could also notice that the shift between DS and P2+, by use of the analogies between histogram and density function and between relative frequency and probability, is not really addressed but is hardly motivated, if at all. Moreover, we have seen that histograms were most of the time defined or represented in an erroneous way, which adds an obstacle to the shift. Ortiz et al. (2002) also find that textbooks and curriculum documents do not offer enough support to secondary teachers.

In the MWSs developed in textbooks, the work undertaken on histograms (semiotic dimension) is considered transparent by the teachers, although this notion proves to be problematic among students. Moreover, the instrumental dimension, when making use of software, is not really used, and when it is it may be a source of difficulty for them.

## 6. Histograms in 10th grade textbooks

To get a better idea of the scientific 12th graders' personal MWS we have taken an interest in their virtually available knowledge on histograms, which belongs explicitly to the 10th

graders' syllabus: “*Represent graphically a statistical series (scatter plot, histogram, cumulative frequency curve)*” (BOEN 2009). For this we studied eleven textbooks published between 2009 and 2014, corresponding to the current syllabus (in use from 2009 on)<sup>9</sup>.

The first question is about the kind of statistical variables associated with a histogram: it is a quantitative variable, the values of which, being too numerous to be represented individually in a graph, are gathered in contiguous classes constituting a partition of a closed limited interval of  $\mathbb{R}$ . It is then a matter of representing the frequencies, or relative frequencies, of the classes. But how to bin (i.e., classify in bins) the range of values, that is, determining the classes, choosing their widths and their limits, is a question hardly taken into account in classrooms. However, a crucial point for the graph is linked with the choice of the limits: shall we stick to equal classes or consider the possibility of unequal ones? In order to make the notion of density—necessary for the study of continuous random variables—appear, the general case has to be taken into account. Our study shows that, although most textbooks mention histograms with unequal classes, they generally content themselves with giving the general definition or/and an example studied in an exercise. Thus the general case appears in fact as derived from the case of equal classes, rather than the contrary.

Similarly the conversions of registers (data  $\rightarrow$  table  $\rightarrow$  graph) are almost absent. With regard to the definition of histograms, six textbooks state that “*for each class, the area of each rectangle is proportional to the frequency, or the relative frequency, of the class*” ([G10T2], p. 162), but two of them indicate that “*the frequency of each class is proportional to the area of the corresponding rectangle*” ([G10T1], p. 233). This means that the former give a procedural definition in the ‘writing’ sense, allowing one to construct the graph from the table, whereas the latter give a ‘reading grid’ of the graph allowing one to build up the table from it. But, in fact, very few exercises give the students a real opportunity to work on the conversions between one register and another.

There is a real risk of confusion between histogram and bar chart, since a textbook, instead of graduating the abscissa axis like a real number axis as usual, names at times the classes by the symbol of the corresponding intervals, as would be done for the modes of a qualitative statistic variable<sup>10</sup> (fig. 10).

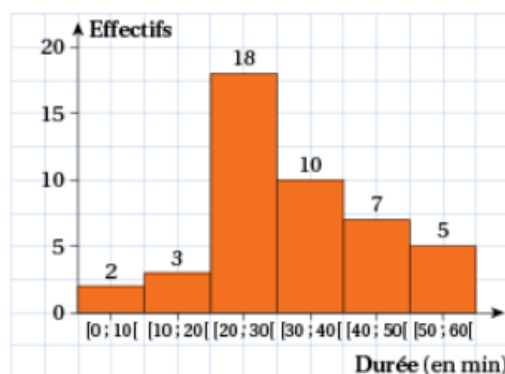


Figure 10. Example of error in graduating the horizontal axis ([G10T7], p. 281)

<sup>9</sup> See list in Appendix 2.

<sup>10</sup> We have already noticed it in 12<sup>th</sup> grade textbooks (cf., fig. 9).

The question of the ordinate axis is important in the process, because it corresponds to the relative frequency density, which is a crucial notion in shifting from discrete to continuous random variables. The way in which this is dealt with in textbooks clearly shows the authors' uneasiness (thus certainly teachers' as well). For example, in the histograms of the textbooks consulted, the widths of the classes may be equal or not, and in both cases there may be a vertical axis or not (table 2).

	<i>Equal widths</i>	<i>Unequal widths</i>
<i>Vertical axis</i>	[G10T1], [G10T3], [G10T4], [G10T5], [G10T6], [G10T9]	[G10T1] (fig. 11), [G10T8], [G10T10]
<i>No vertical axis</i>	[G10T2], [G10T11]	[G10T1], [G10T2], [G10T6]

Table 2. Classification of textbooks based on the studied characteristics of histograms

When the classes are equal, it is not incorrect to represent a vertical axis on which the relative frequencies are shown, for in this case relative frequency is equal to density (apart from a scale factor). But when they are unequal, such an axis should necessarily indicate frequency density, and not size ('effectif') or frequency as in textbook [G10T1] (fig. 11). But the word 'density' never appears, even when the quotient of the relative frequency by the width of the class is made explicit by a formula of the  $f_i/(b_i - a_i)$  type ([G10T10], [G10T4]).

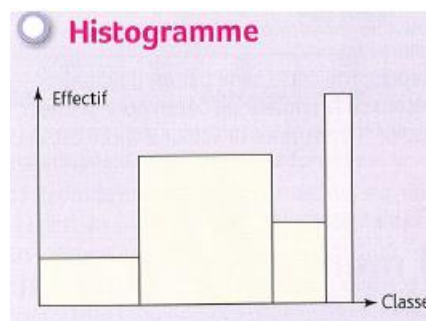


Figure 11. Example of error in naming the vertical axis ([G10T1], p. 233)

This difficulty certainly explains why some textbooks prefer not to tackle the question of the vertical axis and define a square, the area of which corresponds to a given relative frequency (fig. 12).

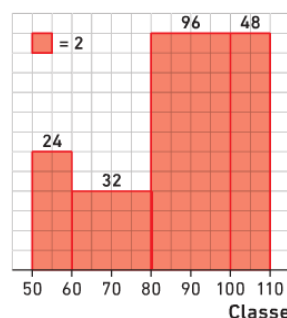


Figure 12. Example of textbook adding size on the histogram ([G10T2], p. 162)

To end with, probably in order to make the identification of the frequencies of the classes easier, most textbooks (7 out of 11) top the rectangles with the corresponding (relative) frequencies (fig. 12).

Another surprising point of our review is that we did not find any reason justifying the quotient process, which may at first look strange to students, although it is quite transparent for the authors.

On the whole, what can be remembered from this study is that the proportionality between area and relative frequency is generally mentioned but not justified, and the meaning of the vertical dimension (density) is completely absent. Making students work on the conversions of registers between table of data and graph is only marginally taken into account and seemingly only to set one's mind at ease.

Finally, the 10th grade textbooks do not really display any interest in histograms. At least three reasons may possibly explain this situation. The first one is that, if one confines oneself to equal classes, this representation somewhat duplicates the bar graph, which is easier to bring into play. The second one is that, if one wants to deal with unequal classes, a proportional process will have to be performed on the data, a technique that some students do not master well, even at that level. The third one is that the most widespread spreadsheets do not provide histograms directly. Thus it appears as a marginal object, the only use of which is to represent some statistical series. At this level there is no clue about a possible further use in probability, e.g., to make the introduction of continuous random variables easier.

A French study (Roditi 2009) shows that French teachers, in junior as well as senior high school, spend little time teaching histograms (about one hour each year) and that “almost half [of them] deem that teaching histograms does not contribute to the students' mathematical education” (p. 131).

This accounts for the weaknesses and errors spotted by our study, especially regarding the semiotic representation of histograms in the suitable MWSs proposed by the textbooks. They may just generate poor personal MWSs (for students and teachers alike), while we think that histogram is a key notion for teaching continuous probability.

## **7. Proposal of a suitable MWS**

In this section we propose a potential suitable MWS taking into account the previous reflections. The question is how to introduce the notion of density function retaining the use of histograms but aiming at giving it more meaning. As we could see above, this appears an interesting idea, but in the suitable MWS proposed by textbooks the very notion of density, which is essential for getting at continuous probability, is in fact a hidden knowledge at 12<sup>th</sup> grade level. Histograms enable one to introduce this idea in a rather ‘natural’ way, which encouraged us to use them for our approach:

Here the notion of (frequency or relative frequency) density appears fundamental, all the more so as one knows that in mathematical statistics continuous quantitative variables are modelled by random variables defined by their probability density function, for which a histogram of the variable constitutes a good estimation. (Chauvat 2002, p. 78)

### 7.1. Density histogram

Scientific 12<sup>th</sup> grade students have not met histograms since 10<sup>th</sup> grade, i.e., two years earlier. Moreover we could see that the ‘histogram’ object is not really present in junior high school and in 10<sup>th</sup> grade. Thus it is important and necessary to undertake working on histograms (in DS) before a probability course, at any rate more thoroughly than is currently done. Batanero et al. (2004b) also stress the importance of working on histograms previously, in their case with the aim of studying normal distributions: “it is important that students understand basic concepts such as probability, density curve (...) and histograms before they start the study of normal distribution” (p. 275). They also insist on

Interpreting areas in frequency histograms and computing areas in the cases when a change in the extremes of intervals is needed. This point is not specific to normal distribution or to the use of computers, and the student should have learnt it at secondary school level. (p. 273)

As Roditi (2009) specifies:

A reader educated in statistics knows that on a histogram the areas of the rectangles are proportional to the frequencies or relative frequencies, he/she knows that the ordinate axis is not an axis of frequencies or relative frequencies but [...] an axis of densities, he/she also knows that in the practice of mathematics a histogram is built up to get an approximation of a probability density curve. (p. 5)

In 12<sup>th</sup> grade, students have already come across the notion of density in other matters (geography, physics...). What seems easier to work on is the notion encountered in geography (population density of humans, animals...). Wikipedia considers it “a key geographic term”:

Population density is a measurement of population per unit area or unit volume; it is a quantity of type number density. It is frequently applied to living organisms, and particularly to humans. [...] Population density is population divided by total land area or water volume, as appropriate.

For humans, population density is the number of people per unit of area, usually quoted per square kilometer or square mile. (Wikipedia)

It is possible to bring to the fore an analogy between population density and relative frequency density. Then for a continuous statistical variable it becomes clear that, for a given population, a relative frequency corresponding to two unequal intervals cannot be represented in the same way, if only from a visual point of view: a relative frequency corresponds to a given proportion of the population and then the colored patch which represents it must correspond to this same proportion of the patch representing the whole population, for fear of being accused of deception.

Besides, a link can be established with the sector diagram (circular or semi-circular), which is particularly suitable for a qualitative variable: in a sector diagram the center angle—and so the area of the surface occupied—is proportional to the relative frequency. Since apparently sector diagram is not problematic with students, it would be possible at 10<sup>th</sup> grade level to exploit the analogy between the two representations (table 3). In both cases one wants the area of the



surface (sector or rectangle) to be proportional to the relative frequency, therefore a similar process has to be brought into play<sup>11</sup>, the main difference being that, in the case of a qualitative variable, the sectors can be displayed in any order, since the classes have to be ordered on the real axis.

Sector diagram	Histogram
issue	class
↓	↓
relative frequency	relative frequency
↓	↓
sector area	rectangle area
↓	↓
central angle	height

Table 3. Comparison between sector diagram and histogram

In the case of histograms, since each coloured patch is a rectangle, the height of which is  $h_i$  and the base is the width  $a_i$  of the class, the area of this patch,  $a_i h_i$ , must be equal (apart from a scale factor) to  $f_i$ , the relative frequency of the class, and this imposes  $h_i = f_i/a_i$  (and not  $f_i$ ).

This approach, quite possible at junior high school, is a means to give sense to histograms, and beyond that to the notion of density. For this, however, it is necessary to work on examples in which the classes have different widths, in order to emphasize what is intended to be represented and how to do it.

This kind of work can be carried out individually, for instance as homework. For 12<sup>th</sup> graders, it allows them to come back to this graphic representation and make them understand that:

- the relative frequency corresponds to the area of the rectangle;
- the ordinate axis (height of rectangles) does not represent the relative frequency of each class but its relative frequency density;
- the sum of the areas of all the rectangles is equal to 1.

Another important point is to bin the range of values in different ways, i.e. represent the same set of data with various histograms, varying the limits and the widths of the classes in order to bring to the fore that, even if the graphs are different, their global appearance stays roughly similar. It is also useful to reflect on how to do it, especially for intervals in which the frequencies are low.

---

<sup>11</sup> The fact that the total area must be equal to 1 will be re-used later in probability.

When the notion of histogram has been re-worked and has become part of the personal MWS of the students, it is now possible to introduce the notion of probability density. Several questions arise then: What types of activities can be proposed to students? What types of probabilistic situations can be chosen?

How to give a sense of probability to the notion of density? One can see that the bet of introducing continuous distributions at the end of high school is daring, to say the least. Here the point of view of modelling is unavoidable. Continuous distributions are then tools enabling one, in the corresponding models, to calculate the probability of some types of events. These continuous models are chosen and their parameters determined (model hypotheses) to fit pretty well with a described discrete reality (working hypotheses), based on heuristic considerations and statistical estimations. The values of the probabilities calculated are proven rather close to the stabilized frequencies of the associated events, obtained through statistical observation. (Henry 2003, p. 6)

Our choice is to use a modelling point of view, as recommended by Henry. Indeed modelling work leaves the possibility of working on open problems, in which the research is not guided and the student has to find the ‘good’ questions by him/herself. We think that two types of modelling situations—not artificial as is often the case in textbooks—can be considered:

- ‘unknown’ probabilistic situations which can be simulated using software, as for instance the Xenakis distribution (difference of two uniform distributions);
- situations using real sets of statistical data.

## 7.2. A classroom situation

In order to illustrate the second possibility, we propose hereafter a situation enabling one to introduce the notion of probability density function from a problem including raw data, a kind of situation that is not much exploited in textbooks but which proves to be quite motivating for students, since they are confronted to a ‘real’ problem leading them to choose a suitable probabilistic model.

This task can be given to scientific 12<sup>th</sup> graders in the initial phase of the study of continuous probability distributions. In the present case the data are the initial years of the eruptions of the Aso volcano (on Kyushu island in Japan), one of the most active in the world. Textbook [G12T5] gives these data (some were missing but we were able to complete them) for the period between the 13<sup>th</sup> and 19<sup>th</sup> centuries (p. 415). In the textbook, these data are worked out with the explicit purpose of “*exploring real numerous statistical data and model them by a continuous distribution*”, some of these distributions having already been studied. We now suggest using these data before the students meet continuous distributions.

Students have at their disposal 90 initial years of eruptions of the volcano between 1229 and 1897 – hence they start in the DS subdomain – and the first question is:

*The last eruption of the volcano started in 2014. How can you estimate the probability for the next one to begin 14 or 15 years later?*

This question stands in the inferential statistical subdomain (IS); it asks to give an estimation based on the data by using available statistical tools (including software).

We can notice that this activity can be proposed to an entire class, alternating phases of individual work and phases of pooling. The use of the computer is a very important support in the work of the class. In this case we chose the GeoGebra software, most frequently used by French teachers. We think that the use of software is of interest in teaching statistics (Biehler 1991; Ben-Zvi 2000; Forbes et al. 2014) and the transition toward continuous probability as well. Similar research has been published for normal distribution (Batanero et al. 1999). In the present case, since the potentialities of the software are of interest for educational purpose, we think that it might be worthwhile to leave the teacher to manage its use at his/her convenience.

Several points have to be considered in the study process of this problem:

- 1) In the first place, the students have to pinpoint the continuousness of the situation, which is not self-evident, since the data (given as integers) are discrete. The waiting time for an eruption is calculated as the difference between the initial years of two successive eruptions; thus it must be understood that a waiting time of 2 means that an eruption starts during year  $n$  and the next one starts during year  $n + 2$ . This implies a necessary discussion among the students on what this number really means (it is in fact an interval, but which one?), to make them become aware of the continuousness of the statistical variable and operate a choice for binning the values. Several share-outs of the data are possible: for instance a waiting time  $k$  may be understood as strictly comprised between  $k$  and  $k + 1$ , or between  $k - 1$  and  $k$ , or between  $k - \frac{1}{2}$  and  $k + \frac{1}{2}$ , or... They may also consider that the data are not precise enough and ask to know the months (for most of the eruptions this information is available).
- 2) To answer the question, they may remain in the DS subdomain and calculate the relative frequency of the waiting time between two successive eruptions in the interval  $[14; 15]$ , but the table provides nothing about this interval, and so the corresponding frequency is nil, although there is no reason to think that such a waiting time is impossible. This shows that it is not possible to assimilate probability with relative frequency: another way has to be looked for.
- 3) They may think of grouping the data in classes and representing them, not by a bar chart but by a histogram of relative frequencies (in SD). But such a representation is of no use to answer the question, although if you choose large enough widths for the classes it is possible to determine an approximation of the relative frequency, which now will be strictly positive but different from the empirical frequency ( $= 0$ ). This leads the students back to point 2. However they have been able to observe that, for any initial grouping of the data as well as for any set of classes (either equal or not), the associated histograms are visually close to one another.
- 4) The global look of the histograms may lead—with or without help from the teacher—to search for a ‘trend curve’, as written in some textbooks, or a curve ‘that smooths the upper edge of the histogram’ as others say. A change of viewpoint on data is necessary: the students must be able to shift from a local to a global view on data (Ben-Zvi and Arcavi, 2001). This is the modelling viewpoint in which empirical and theoretical distributions coexist (Wilensky 1997; Wild 2006; Pfannkuch and Reading, 2006). This curve, linked with the histogram, must necessarily verify some conditions, since it represents a function:

- defined on an interval in which the random variable can take its values
- positive in this interval
- the area under the curve of which is equal to 1.

In scientific 12<sup>th</sup> grade, the continuousness of the function—although actually unnecessary—is part of the conditions. The above properties rest on the analogy with the properties of histograms.

The random variable here being the waiting time between two eruptions, its values belong theoretically to  $[0; +\infty[$ . Thus the students have to find a function defined on this interval, close to the histogram and the area under which is 1. The MWS is now in the graphical register of the domain of calculus. It can make use of a software like GeoGebra which makes it possible, with cursors, to test a very large number of functions in a short time. The function  $x \rightarrow 1/x$  and other similar ones are visually potential candidates, but they are quickly eliminated because of the area condition. The functions of the  $x \rightarrow ke^{-\alpha x}$  type, studied at that level, appear as better candidates. After research on the conditions it seems that the functions  $f(x) = \lambda \exp(-\lambda x)$  with  $\lambda \in [0,10; 0,17]$  are suitable (fig. 13).

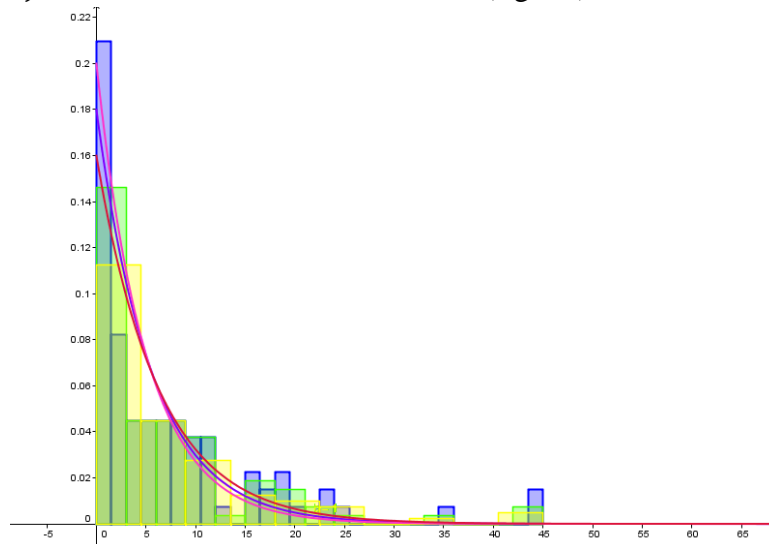


Figure 13. Smoothing by exponential curves

To verify that the area under the curve is equal to 1 (improper integrals are not studied in 12<sup>th</sup> grade) a first approach can be to use GeoGebra, which provides approximations (fig. 14).

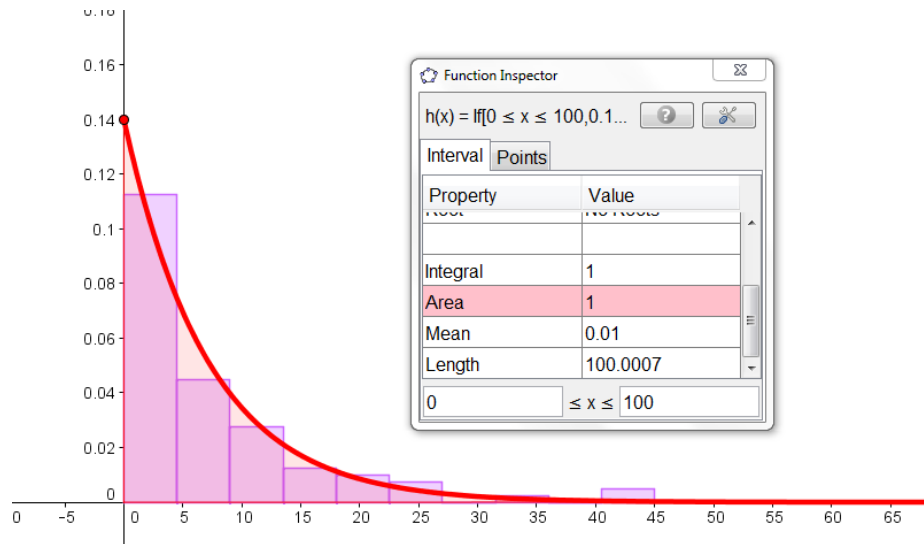


Figure 14. Estimation of the area under the curve

The process which has just been described can be schematized as a diagram (fig. 15).

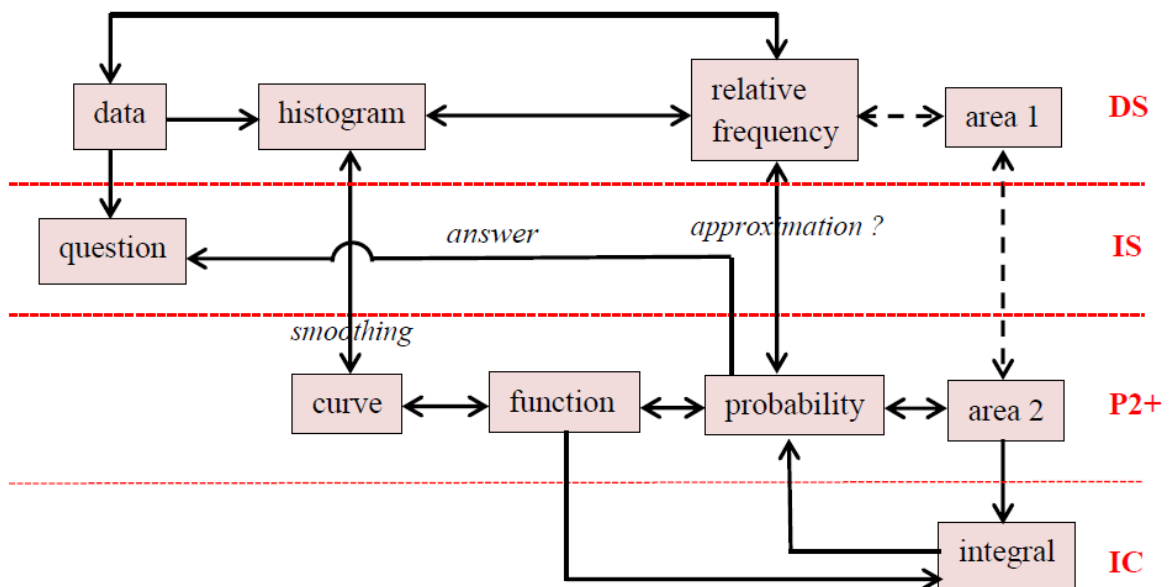


Figure 15. Diagram of the process of the proposed activity

Such a situation enables a questioning on the continuous nature of a random variable and on what a model is; it leads in a 'natural' way to the need of a 'trend curve' (density curve). The support of the analogy between relative frequency histogram (representing empirical density) and density curve (representing theoretical density) helps give a real meaning to the word 'density' and justify the link between DS and P2+.

The aim of this situation is to favour the semiotic and instrumental dimensions by fostering a tight link between representation registers and use of software. Notwithstanding a significant part of the mathematical work regarding visualisation and fitting curves, the discursive dimension is requested to apply constraints on the curve.

## 8. Conclusion

As stated by Batanero et al. (2004b) for normal distribution, density distributions are “a very complex idea that requires the integration and relation of many different statistical concepts and ideas” (p. 274). In this study we have brought to the fore that histograms, although they are statistical graphic representations normally studied over four years in high school (from 7<sup>th</sup> to 10<sup>th</sup> grade), have only a minor place in the personal MWS of French students in the domain of descriptive statistics. At its very best, it only includes:

- conversions of registers (tables ↔ graphs)
- placing points on the real axis (limits of classes)
- proportional calculating (heights of rectangles).

However the process suggested here, which confers to histograms a genuine role in the students’ mathematical education—namely an introduction to the study of density distributions—makes it necessary to undertake a much more thorough study than the one currently done in classrooms. Various other tasks will have to be performed:

- iconic: spot variations, symmetries, comparing areas, recognizing known forms...
- graphical: choosing classes, interpreting rectangles as statistical data, gathering classes...

In scientific 12<sup>th</sup> grade, when studying density distributions, such tasks aim mainly at:

- putting to the fore the effects of changing the classes on a histogram, and thus showing the importance of choosing them adequately;
- making explicit the conditions that a density function must verify;
- making the students used to recognizing and distinguishing various forms of density curves, and interpreting them in terms of probability;
- providing them with types of model distributions and enabling them to choose the one(s) that seem best in accordance with the distribution.

The proposed situation highlights the interest of using software in introducing this new notion. It is a way to get students to focus on searching for a model, and not on constructing histograms or functions, which would increase the amount of work.

This new way of approaching continuous distributions in 12<sup>th</sup> grade requires that teachers think of another suitable MWS than the usual one, because

Teachers need to know precisely the knowledge to be taught in reference to the scholarly knowledge, i.e. to know mathematical definitions and also the reasons which are at their origin and the questions which the knowledge answers to. (Roditi 2009, p. 137)

Then we hope that a new ‘first encounter’ of histograms in 12<sup>th</sup> grade may help teachers to give sense to the notion of density, in order to prepare the study of continuous probability distributions in an effective way.

Through this study we could see that in most textbooks the introduction of the notion of density rests on the notion of histogram, this kind of graph being indeed quite an interesting

means to give sense to the notion of density. But we could also pinpoint that it was underused in current teaching. Hence, taking an interest in this point, we showed that the corresponding working space was quite poor. Besides we could spot some errors in the textbooks, which have some consequence for teachers... and consequently for students. This poor MWS around histograms implies that they cannot be efficiently used as a support to give sense to density functions.

This is the reason why we recommend a more thorough and rigorous study of histograms, allowing later on the reliance on them for introducing density functions, the activity presented in section 7 being an example illustrating our point of view.

Having shown the interest of using histograms in the learning of probability, we hope that teachers will now give it more attention than they usually did, when they understand that it is a crucial element for giving sense to probability density, through an assumed and conscious back and forth movement between two mathematical domains: statistics and probability.

### Appendix 1. List of 12<sup>th</sup> grade textbooks

- [G12T1] Barros, J.-M., Bénizeau, P. & Morin, J. (dir.). (2012). *Transmath. Term S. Enseignement spécifique*. Paris: Nathan.
- [G12T2] Beltramone, J.-P. (dir.). (2012). *Décllic mathématiques TS. Enseignement spécifique et de spécialité*. Paris: Hachette Education.
- [G12T3] Bruneau, F., Choquer-Raoult, A., Cocault, M., Ferré, F., Hanouch, B., Joffrédo, T., Lavancier, F., Mauxion, H. & Simon, D. (2012). *Term S Maths repères. Enseignement spécifique et de spécialité*. Paris: Hachette Education.
- [G12T4] Deschamps, C. (dir.). (2012). *Symbole maths Term S*. Paris: Belin.
- [G12T5] Le Yaouanq, M.-H. (dir.). (2012). *Math'x. Term S. Enseignement spécifique*. Paris: Didier.
- [G12T6] Malaval, J. (dir.). (2012). *Hyperbole. Mathématiques. Term S. Enseignement spécifique*. Paris: Nathan.
- [G12T7] Poncy, M., Bonnafet, J.-L. & Russier M.-C. (dir.). (2012). *Indice. Maths. Term S. Enseignement spécifique*. Paris: Bordas.
- [G12T8] Sigward, E. (dir.). (2011). *Odyssée. Mathématiques. Tle S. Enseignement spécifique*. Paris: Hatier.

### Appendix 2. List of 10<sup>th</sup> grade textbooks

- [G10T1] Antibi, A. (dir.). (2010). *Math 2de. Travailler en confiance*. Paris: Nathan.
- [G10T2] Barbazo, E. (dir.). (2014). *Mathématiques 2de*. Paris: Hachette Education.
- [G10T3] Beltramone, J.-P. (dir.). (2014). *Décllic mathématiques 2de*. Paris: Hachette Education.
- [G10T4] Malaval, J., & Courbon, D. (dir.). (2010). *Hyperbole. Mathématiques. 2de*. Paris: Nathan.
- [G10T5] Gauthier, R., & Poncy, M. (2009). *Indice. Maths. 2de*. Paris: Bordas.
- [G10T6] Chesné, J.-F., & Le Yaouanq, M.-H. (dir.). (2010). *Math'x. 2de*. Paris: Didier.
- [G10T7] Sigward, E. (dir.). (2010). *Odyssée. Mathématiques. 2de*. Paris: Hatier.
- [G10T8] Bouvier, J.-P. (dir.). (2010). *pxel. Maths. 2de*. Paris: Bordas.
- [G10T9] Choquer-Raoult, A., Cocault, M., Hanouch, B., & Joffrédo, T. (2010). *2de maths repères*. Paris: Hachette Education.
- [G10T10] Deschamps, C. (dir.). (2010). *Symbole maths 2e*. Paris: Belin.
- [G10T11] Barra, R. (dir.). (2010). *Transmath. 2de*. Paris: Nathan.



### Appendix 3. Introduction activity of the textbook [G12T6] (p.377).

#### objectif

Approcher la définition de la loi de probabilité d'une variable continue.

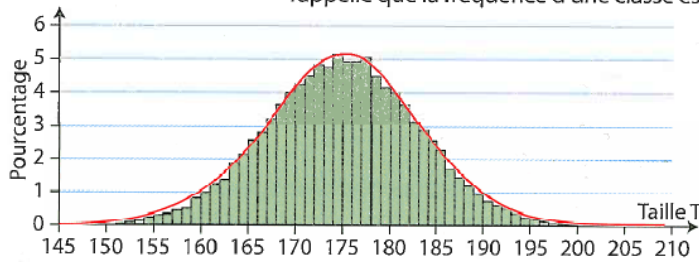


#### La notion de densité d'une loi de probabilité

Un commerçant souhaite vendre un stock de vêtements dont la taille convient à des hommes adultes mesurant entre 1,70 m et 1,77 m.

On se propose de déterminer avec quelle probabilité il pourra satisfaire un client qui entre au hasard dans son magasin.

Pour cela, voici l'histogramme obtenu pour un échantillon de 50 000 hommes adultes. On rappelle que la fréquence d'une classe est donnée par l'aire du rectangle correspondant.



1. Estimer graphiquement la proportion des individus dont la taille est entre 1,70 m et 1,77 m.

2. On a tracé la courbe d'une fonction qui « épouse » l'histogramme et on admet que pour tout échantillon de grand effectif, on obtiendrait un histogramme proche de cette courbe.

Comment calculer à l'aide de cette fonction la probabilité de l'événement  $(1,70 \leq T \leq 1,77)$ ? Quelle doit être l'aire sous la courbe de  $f$ ? On dit que  $f$  est la **densité** de la loi  $P$ .

## References

- Atkin, A. (2013). Peirce's Theory of Signs. *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition). Edward N. Zalta (ed.).  
<http://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/>  
 Accessed 2016/2/3.
- Bansilal, S. (2012). Understanding student engagement with normal distribution problems. *Pythagoras*, 33(1), 37-49.
- Bansilal, S. (2014). Using an apos framework to understand teachers' responses to questions on the normal distribution. *Statistics Education Research Journal*, 13(2), 42-57.
- Batanero, C., Tauber, L., & Meyer, R. (1999). From data analysis to inference: A research project on the teaching of normal distributions. *Bulletin of the International Statistical Institute: Proceedings of the Fifty-Second Session of the International Statistical Institute* (Tome LVIII, Book 1, pp. 57–58). Helsinki, Finland: International Statistical Institute.
- Batanero, C., Godino, J. D., & Roa R. (2004a), Training teachers to teach probability. *Journal of Statistics Education*, 12(1).  
<http://www.amstat.org/publications/jse/v12n1/batanero.html>  
 Accessed 2016/2/3.
- Batanero, C., Tauber, L. M., & Sánchez V. (2004b), Students' reasoning about the normal distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 257-275). Dordrecht: Kluwer.
- Ben-Zvi, D. (2000). Towards understanding the role of technological tools in statistical learning. *Mathematics Thinking and Learning*, 2(1&2), 127–155.
- Ben-Zvi, D., & Arcavi, A. (2001). Junior high school student's construction of global views of data and data representations. *Educational Studies in Mathematics*, 43, 35–65.
- Biehler, R. (1991). Computers in probability education. In R. Kapadia and M. Borovcnik (Eds.), *Chance Encounters - Probability in Education*, (pp. 169-221). Dordrecht: Kluwer.
- Birgé, L., & Rozenholc, Y. (2006). How many bins should be put in a regular histogram. *ESAIM: Probability and Statistics*, 10, pp. 24-45.
- Chauvat, G. (2002). Quelques graphiques de plus ! Montrer et voir en statistiques. *Repères-IREM*, 47, 75-92.
- Chevallard, Y. (2007). Passé et présent de la théorie anthropologique du didactique. In L. Ruiz-Higueras, A. Estepa, & FJ García (Eds.), *Sociedad, Escuela y Matemáticas. Aportaciones de la Teoría Antropológica de la Didáctica* (pp. 705-746). Jaén: Universidad de Jaén. Servicio de Publicaciones.
- Duval, R. (1995). *Sémiosis et pensée humaine*. Berne: Peter Lang.
- Forbes, S., Chapman, J., Harraway, J., Stirling, D., & Wild, C. (2014). Use of data visualization in the teaching of statistics: a New Zealand perspective. *Statistics Education Research Journal*, 13(2), 187-201.

Gal, I. (2002). Adult's statistical literacy. Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.

Gómez-Chacón, I., and Kuzniak, A. (2015). Spaces for geometric work: figural, instrumental, and discursive geneses of reasoning in a technological environment. *International journal of science and mathematics education*, 13, 201-223.

Hawkins, A., Jolliffe, F., and Glickman, L. (1991), *Teaching statistical concepts*, London: Longman.

Henry, M. (1999). L'introduction des probabilités au lycée : un processus de modélisation comparable à celui de la géométrie. *Repères-IREM*, 36, 15-34.

Henry, M. (2003). Des lois de probabilités continues en terminale S, pourquoi et pour quoi faire ? *Repères-IREM*, 51, 5-25.

Holmes, P. (1980). *Statistics in your world*. Slough: Foulsham Educational.

Houdement, C., & Kuzniak, A. (2006). Paradigmes géométriques et enseignement de la géométrie. *Annales de didactique et des sciences cognitives*, 10, 5-53.

Kuhn, T. S. (1962). *La structure des révolutions scientifiques*. Flammarion.

Kuzniak, A. (2010). Un essai sur la nature du travail géométrique en fin de la scolarité obligatoire en France. *Annales de didactique et de sciences cognitives*, 16, 9-24.

Kuzniak, A. (2011). L'espace de Travail Mathématique et ses genèses. *Annales de didactique et de sciences cognitives*, 16, 9-24.

Kuzniak, A., & Richard, P. R. (2014). Spaces for mathematical work: viewpoints and perspectives. *Revista Latinoamericana de Investigación en Matemática Educativa*, 17 (4-I), 17-27.

Ministère éducation nationale (MEN). (2007). *Ressources pour les classes de 6è, 5è, 4è, et 3è du collège. Organisation et gestion de données au collège.*

[http://cache.media.eduscol.education.fr/file/Programmes/17/5/doc\\_acc\\_clg\\_organisation\\_donnes\\_109175.pdf](http://cache.media.eduscol.education.fr/file/Programmes/17/5/doc_acc_clg_organisation_donnes_109175.pdf)

Accessed 2016/2/3.

Ministère de l'éducation nationale (MEN). (2009). *Ressources pour la classe de seconde. Probabilités et statistiques.*

[http://cache.media.eduscol.education.fr/file/Programmes/17/9/Doc\\_ressource\\_proba\\_stats\\_109179.pdf](http://cache.media.eduscol.education.fr/file/Programmes/17/9/Doc_ressource_proba_stats_109179.pdf)

Accessed 2016/2/3.

Ministère de l'Éducation nationale, de la jeunesse et de la vie associative (MENJVA). (2011). Programme de l'enseignement spécifique et de spécialité de mathématiques. Classe terminale de la série scientifique. *Bulletin Officiel spécial*, n°8 (13 octobre 2011).

[http://cache.media.education.gouv.fr/file/special\\_8\\_men/98/4/mathematiques\\_S\\_195984.pdf](http://cache.media.education.gouv.fr/file/special_8_men/98/4/mathematiques_S_195984.pdf)

Accessed 2016/2/3.

Ministère de l'Éducation nationale, de la Jeunesse et de la Vie associative (MENJVA) & DGESCO. (2012). *Ressources pour la classe terminale générale et technologique. Probabilités et statistique.*

[http://cache.media.eduscol.education.fr/file/Mathematiques/11/5/LyceesGT\\_ressources\\_Math\\_T\\_proba-stat\\_207115.pdf](http://cache.media.eduscol.education.fr/file/Mathematiques/11/5/LyceesGT_ressources_Math_T_proba-stat_207115.pdf)

Accessed 2016/2/3.

Ministère de la Jeunesse, de l'Éducation nationale et de la Recherche (MJENR). (2002). *Mathématiques. Classe terminale. Série économique et sociale. Série scientifique.* Paris: Centre national de documentation pédagogique.

Montoya Delgadillo, E., & Vivier, L. (2014). Les changements de domaine dans le cadre des Espaces de Travail Mathématique. *Annales de didactique et de sciences cognitives*, 19, 73-102.

Ortiz, J. J., Cañizares, M. J., Batanero, C., and Serrano, L. (2002). An experimental study of probabilistic language in secondary school textbooks. Contributed paper to the *International Conference On Teaching Statistics 6 (ICOTS 6)*, Cape Town, July 2002.

Parzysz, B. (2011). Quelques questions didactiques de la statistique et des probabilités. *Annales de Didactique et de Sciences cognitives*, 16, 127-147.

Parzysz, B. (2016). Les probabilités dans l'enseignement secondaire, d'hier à demain. In Robert, A. (Ed.), *Formation pour l'enseignement des probabilités et de la statistique au lycée.* (Chapter 1). Besançon: Presses Universitaires de Franche-Comté (to appear).

Perrin, D. (2015). Remarques sur l'enseignement des probabilités et de la statistique au lycée. *Statistique et enseignement*, 6(1), 51-63.

<http://www.statistique-et-enseignement.fr>

Accessed 2016/2/3.

Roditi, E. (2009). L'histogramme : à la recherche du savoir à enseigner. *Spirale. Revue de recherches en éducation*, 43, 129–138.

Vere-Jones, D. (1995). The coming of age of statistical education. *International Statistical Review*, 63(1), 3-23.

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33(2), 171-202.

Wild, C. (2006). The concept of distribution. *Statistics Education Research Journal*, 5(2), 10-26.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 221-248.

Asosan eruption database (2015). *Smithsonian Institution. National Museum of Natural History.*

<http://www.volcano.si.edu/volcano.cfm?vn=282110>

Accessed 2016/2/3.