

Chapter 11

STUDENTS' REASONING ABOUT THE NORMAL DISTRIBUTION ¹

Carmen Batanero¹, Liliana Mabel Tauber², and Victoria Sánchez³
*Universidad de Granada, Spain¹, Universidad Nacional del Litoral, Santa Fe, Argentina²,
and Universidad de Sevilla, Spain³.*

OVERVIEW

In this paper we present results from research on students' reasoning about the normal distribution in a university-level introductory course. One hundred and seventeen students took part in a teaching experiment based on the use of computers for nine hours, as part of a 90-hour course. The teaching experiment took place during six class sessions. Three sessions were carried out in a traditional classroom, and in another three sessions students worked on the computer using activities involving the analysis of real data. At the end of the course students were asked to solve three open-ended tasks that involved the use of computers. Semiotic analysis of the students' written protocols as well as interviews with a small number of students were used to classify different aspects of correct and incorrect reasoning about the normal distribution used by students when solving the tasks. Examples of students' reasoning in the different categories are presented.

THE PROBLEM

One problem encountered by students in the introductory statistics course at university level is making the transition from data analysis to statistical inference. To make this transition, students are introduced to probability distributions, with most of the emphasis placed on the normal distribution. The normal distribution is an important model for students to learn about and use for many reasons, such as:

¹ This research has been supported by DGES grant BS02000-1507 (M.E.C., Madrid).

- Many physical, biological, and psychological phenomena can be reasonably modeled by this distribution such as physical measures, test scores and measurement errors.
- The normal distribution is a good approximation for other distributions—such as the binomial, Poisson, and t distributions—under certain conditions.
- The Central Limit Theorem assures that in sufficiently large samples the sample mean has an approximately normal distribution, even when samples are taken from nonnormal populations.
- Many statistical methods require the condition of random samples from normal distributions.

We begin by briefly describing the foundations and methodology of our study. We then present results from the students' assessment and suggest implications for the teaching of normal distributions. For additional analyses based on this study see Batanero, Tauber, and Meyer (1999) and Batanero, Tauber, and Sánchez (2001).

THE LITERATURE AND BACKGROUND

Previous Research

There is little research investigating students' understanding of the normal distribution, and most of these studies examine isolated aspects in the understanding of this concept. The first pioneering work was carried out by Piaget and Inhelder (1951), who studied children's spontaneous development of the idea of stochastic convergence. The authors analyzed children's perception of the progressive regularity in the pattern of sand falling through a small hole (in the Galton apparatus or in a sand clock). They considered that children need to grasp the symmetry of all the possible sand paths falling through the hole, the probability equivalence between the symmetrical trajectory, the spread and the role of replication, before they are able to predict the final regularity that produces a bell-shaped (normal) distribution. This understanding takes place in the formal operations stage (13- to 14-year-olds).

Regarding university students, Huck, Cross, and Clark (1986) identified two erroneous conceptions about normal standard scores: On the one hand, some students believe that all standard scores will always range between -3 and $+3$, while other students think there is no restriction on the maximum and minimum values in these scores. Each of those beliefs is linked to a misconception about the normal distribution. The students who think that z -scores always vary from -3 to $+3$ have frequently used either a picture or a table of the standard normal curve, with this range of variation. In a similar way, the students who believe that z -scores have no upper or lower limits have learned that the tails of the normal curve are asymptotic to the abscissa; thus they make an incorrect generalization, because they do not notice that no finite distribution is exactly normal.

For example, if we consider the number of girls born out of 10 newborn babies, this is a random variable X , which follows the binomial distribution with $n = 10$ and $p = 0.5$. The mean of this variable is $np = 5$ and the variance is $npq = 2.5$. So the maximum z -score that could be obtained from this variable is $z_{max} = (10 - 5)/\sqrt{2.5} = 3.16$. Thus we have a finite limit, but it is greater than 3.

In related studies, researchers have explored students' understanding of the Central Limit Theorem and have found misconceptions regarding the normality of sampling distributions (e.g., Vallecillos, 1996, 1999; Méndez, 1991; delMas, Garfield, & Chance, 1999). Wilensky (1995, 1997) examined student behavior when solving problems involving the normal distribution. He defined *epistemological anxiety* as the feeling of confusion and indecision that students experience when faced with the different paths for solving a problem. In interviews with students and professionals with statistical knowledge, Wilensky asked them to solve a problem by using computer simulation. Although most subjects in his research could solve problems related to the normal distribution, they were unable to justify the use of the normal distribution instead of another concept or distribution, and showed a high epistemological anxiety.

Meaning and Understanding of Normal Distributions in a Computer-Based Course

Our research is based on a theoretical framework about the meaning and understanding of mathematical and statistical concepts (Godino, 1996; Godino & Batanero, 1998). This model assumes that the understanding of normal distributions (or any other concept) emerges when students solve problems related to that concept. The meaning (understanding) of the normal distribution is conceived as a complex system, which contains five different types of elements:

1. *Problems and situations from which the object emerges.* In our teaching experiments, students solved the following types of problems: (a) fitting a curve to a histogram or frequency polygon for empirical data distributions, (b) approximating the binomial or Poisson distributions, and (c) finding the approximate sampling distribution of the sample mean and sample proportion for large samples (asymptotic distributions).
2. *Symbols, words, and graphs used to represent or to manipulate the data and concepts involved.* In our teaching, we considered three different types of representations:
 - a) *Static paper-and-pencil graphs and numerical values of statistical measures*, such as histograms, density curves, box plots, stem-leaf plots, numerical values of averages, spread, skewness, and kurtosis. These might appear in the written material given to the students, or be obtained by the students or teacher.
 - b) *Verbal and algebraic representations of the normal distribution; its properties or concepts related to normal distribution*, such as the words *normal* and *distribution*; the expressions *density curve*, *parameters of the*

normal distribution, the symbol $N(\mu, \sigma)$, equation of density function, and so forth.

- c) *Dynamic graphical representations on the computer*. The Statgraphics software program was used in the teaching. This program offers a variety of simultaneous representations on the same screen which are easily manipulated and modified. These representations include histograms, frequency polygons, density curves, box plots, stem-leaf plots, and symmetry and normal probability plots. The software also allows simulation of different distribution, including the normal distribution.
3. *Procedures and strategies to solve the problem*. Beyond the descriptive analyses of the variables studied in the experiment, the students were introduced to computing probabilities under the curve, finding standard scores, and critical values (computed by the computer or by hand).
4. *Definitions and properties*. Symmetry and kurtosis: relative position of the mean, median and mode, areas above and below the mean, probabilities within one, two and three standard deviations, meanings of parameters, sampling distributions for means and proportions, and random variables.
5. *Arguments and proofs*. Informal arguments and proofs made using graphical representation, computer simulations, generalization, analysis, and synthesis.

SUBJECTS AND METHOD

Sample and Teaching Context

The setting of this study was an elective, introductory statistics course offered by the Faculty of Education, University of Granada. The instruction for the topic of normal distributions was designed to take into account the different elements of meaning as just described. Taking the course were 117 students (divided into 4 groups), most of whom were majoring in Pedagogy or Business Studies. Some students were from the School of Teachers Training, Psychology, or Economics.

At the beginning of the course students were given a test of statistical reasoning (Garfield, 1991) to assess their reasoning about simple statistical concepts such as averages or sampling, as well as to determine the possible existence of misconceptions. An examination of students' responses on the statistical reasoning test revealed some errors related to sampling variability (representativeness heuristics), sample bias, interpretation of association, and lack of awareness of the effect of atypical values on averages. There was a good understanding of probability, although some students showed incorrect conceptions about random sequences.

Before starting the teaching of the normal distribution, the students were taught the foundations of descriptive statistics and some probability, with particular emphasis on helping them to overcome the biases and errors mentioned. Six 1.5-

hour sessions were spent teaching the normal distribution, and another 4 hours were spent studying sampling and confidence intervals. Students received written material specifically prepared for the experiment and asked to read it beforehand. Half of these sessions were carried out in a traditional classroom, where the lecturer introduced the normal distribution as a model to describe empirical data, using a computer with projection facility. Three samples ($n = 100, 1,000, \text{ and } 10,000$ observations) of intelligence quotient (IQ) scores were used to progressively show the increasing regularity of the frequency histogram and polygon, when increasing the sample size. The lecturer also presented the students with written material, posed some problems to encourage the students to discover for themselves all the elements of meaning described in section 3.2, and guided student discussion as they solved these problems.

The remaining sessions were carried out in a computer lab, where pairs of students worked on a computer to solve data analysis activities, using examples of real data sets from students' physical measures, test scores, and temperatures, which included variables that could be fitted to the normal distribution and other variables where this was not possible. Activities included checking properties such as unimodality or skewness; deciding whether the normal curve provided a good fit for some of the variables; computing probabilities under the normal curve; finding critical values; comparing different normal distributions by using standardization; changing the parameters in a normal curve to assess the effect on the density curve and on the probabilities in a given interval; and solving application problems. Students received support from their partner or the lecturer if they were unable to perform the tasks, and there was also collective discussion of results.

Assessing Students' Reasoning about the Normal Distribution

At the end of the course students were given three open-ended tasks, to assess their reasoning about the normal distribution as part of a final exam that included additional content beyond this unit. These questions referred to a data file students had not seen before, which included qualitative and quantitative (discrete and continuous) variables (See Table 1). The students worked alone with the Statgraphics program, and they were free to solve the problem using the different tools they were familiar with.

Each problem asked students to complete a task and to explain and justify their responses in detail, following guidelines by Gal (1997), who distinguished two types of questions to use when asking students to interpret statistical information. Literal reading questions ask students for unambiguous answers—they are either right or wrong. In contrast, to evaluate questions aimed at eliciting students' ideas about overall patterns of data, we need information about the evidential basis for the students' judgments, their reasoning process, and the strategy they used to relate data elements to each other. The first type of question was taken into account in a questionnaire with 21 items, which was also given to the students in order to assess literal understanding for a wide number of elements of the normal distribution

(Batanero et al., 2001). The second type of question considered by Gal (1997) was considered in the following open tasks given to students.

Task 1

In this data file, find a variable that could be fitted by a normal distribution. Explain your reasons for selecting that variable and the procedure you have used.

In this task the student is asked to discriminate between variables that can be well fitted to a normal distribution and others for which this is not possible. In addition to determining the student's criteria when performing the selection (the properties they attribute to normal distributions), we expected students to analyze several variables and use different approaches to check the properties of the different variables to determine which would best approximate a normal distribution. We also expected students to synthesize the results to obtain a conclusion from all their analyses. We hoped that student responses to this task would reveal their reasoning.

Task 2

Compute the appropriate values of parameters for the normal distribution to which you have fitted a variable chosen in question 1.

In this question the students have to remember what the parameters in a normal distribution (mean and variance) are. We also expected them to remember how to estimate the population mean from the sample mean and to use the appropriate Statgraphic program to do this estimation. Finally, we expected the students to discriminate between the ideas of statistics (e.g., measures based on sample data) and parameters (e.g., measures for atheoretical population model).

Task 3

Compute the median and quartiles for the theoretical distribution you have constructed in Task 2.

The aim is to evaluate the students' reasoning about the ideas of median and quartiles for a normal distribution. Again, discrimination between empirical data distribution and the theoretical model used to fit these data is needed. We expect the student to use the critical value facility of Statgraphics to find the median and quartiles in the theoretical distribution. Those students who do not discriminate will probably compute the median and quartile from the raw empirical data with the summary statistics program.

The three tasks just described were also used to evaluate the students' ability to operate the statistical software and to interpret its results. Since the students were free to solve the tasks using any previous knowledge to support their reasoning, we could evaluate the correct or incorrect use of the different meaning elements (representations, actions, definitions, properties, and arguments) that we defined earlier and examine how these different elements were interrelated.

Each student worked individually with the Statgraphics and produced a written report using the word processor, in which they included all the tables and graphs needed to support their responses. Students were encouraged to give detailed

reasoning. Once the data were collected, the reports were printed and we made a content analysis. We identified which elements of meaning each student used correctly and incorrectly to solve the tasks.

In the next section we provide a global analysis for each question and then describe the elements of meaning used by the students.

RESULTS AND ANALYSIS

Students' Perception of Normality

In Table 1, we include the features of variables in the file and the frequency and percentage of students who selected each variable in responding to the first question. The normal distribution provided a good fit for two of these variables: *Time to run 30 m (December)* and *Heartbeats after 30 press-ups*. The first variable, *Time to run 30 m*, was constructed by simulating a normal continuous distribution. Normality can be checked easily in this variable from its graphical representation; the skewness and kurtosis coefficient were very close to zero, although the mean, median, and mode did not exactly coincide. *Heartbeats after 30 press-ups* was a discrete variable; however, its many different values, its shape, and the values of its different parameters suggested that the normal distribution could provide an acceptable fit.

Table 1. Description of the variables that students considered to fit a normal distribution well

Variable	Variable Features			Students choosing this variable (%)	
	Variable type	Skewness	Kurtosis		
Age	Discrete; three different values	0	-0.56	13, 13, 13	27 (23.1)
Height	Continuous Multimodal	0.85	2.23	156.1, 155.5, †	26 (22.2)
Heartbeats after 30 press-ups*	Discrete; many different values	0.01	-0.19	123.4, 122, 122	37 (31.6)
Time spent to run 30 m.(Dec.)*	Continuous	0.23	-0.42	4.4, 4.4, 5.5	12 (10.3)
Weight	Continuous Atypical values	2.38	9.76	48.6, 46, 45	4 (3.4)
Heartbeats at rest	Discrete; many different values	0.2	-0.48	71.4, 72, 72	6 (5.2)
Time spent to run 30 m. (Sep.)	Continuous	2.4	12.2	5.3, 5.2, 5	4 (3.4)
No answer					9 (7.2)

* Correct answer. The normal distribution is a good approximation for these variables

† Although *Height* had in fact three modes: 150, 155, 157, that were visible from the stem plot, this was noticeable only from the histogram with specific interval widths.

The variable *Height*, despite being symmetric, had kurtosis higher than expected and was multimodal, though this was noticeable only by examining a stem-and-leaf plot or histogram of the data.

Some of these students confused the empirical data distribution for *Age* (Fig. 1a) with the theoretical distribution they fitted to the data. In Figure 1b the data frequency histogram for *Age* and a superimposed theoretical normal curve are plotted. Some students just checked the shape of the theoretical density curve (the normal curve with the data mean and standard deviation) without taking into account whether the empirical histogram approached this theoretical curve or not.

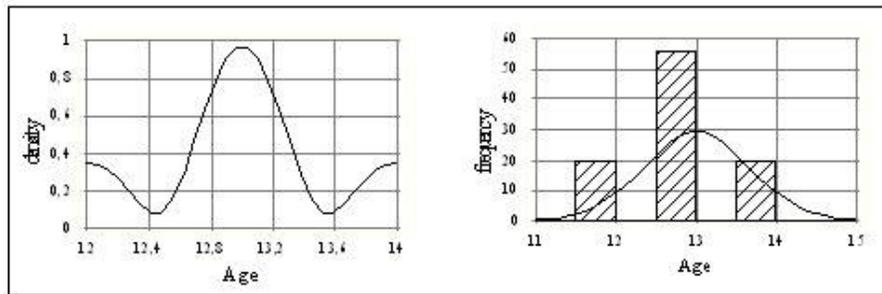


Figure 1. (a) Empirical density curve for *Age* (b) Theoretical normal curve fitted to *Age*.

Twenty-two percent of students selected a variable with high kurtosis (*Height*). In the following example, while the student could perceive the symmetry from the graphical representation of data, this graph was however unproductive as regards the interpretation of the standard kurtosis coefficient (4.46) that was computed by the student. The student did not compute the median and mode. We assume he visually perceived the curve symmetry and from this property he assumed the equality of mean, median, and mode.

Example 2

"I computed the mean (156.1) and standard deviation (8,93) and they approach those from the normal distribution. Then I represented the data (Figure 2) and it looks very similar to the normal curve. The values of mean, median and mode also coincide. Std Kurtosis = 4.46" (Student 2).

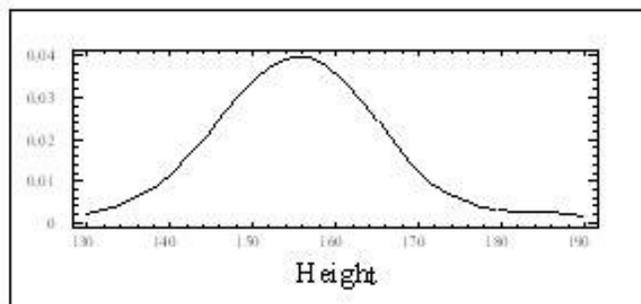


Figure 2. Density trace for Height.

Finding the Parameters

Table 2 displays the students' solutions to question 2. Some students provided incorrect parameters or additional parameters such as the median that are not needed to define the normal distribution. In Example 3, the student confuses the tail areas with the distribution parameters. In Example 4, the student has no clear idea of what the parameters are and he provides all the summary statistics for the empirical distribution.

Example 3

"These are the distribution parameters for the theoretical distribution I fitted to the variable pulsation at rest:

area below 98.7667 = 0.08953

area below 111.113 = 0.25086

area below 123.458 = 0.5" (Student 3)

Example 4

"Count=96, Average = 123.458, Median = 122.0, Mode = 120.0, Variance = 337.682, Standard deviation = 18.3761, Minimum = 78.0, Maximum = 162.0, Range = 84.0, Skewness = 0.0109784, Std. Skewness = 0.043913, Kurtosis = -0.197793, Std. Kurtosis = -0.395585, Coeff. of variation = 14.8845%, Sum = 11852.0" (Student 4).

These results suggest difficulties in understanding the idea of parameter and the difference between theoretical and empirical distributions.

Table 2. Frequency and percentage of responses in computing the parameters

Response	Number and Percentage
Correct parameters	60 (51)
Incorrect or additional parameters	18 (15)
No answer	39 (33)

Computing Percentiles in the Theoretical Distribution

Table 3 presents a summary of students' solutions to question 3. About 65% of the students provided correct or partly correct solutions in computing the median and quartiles. However, few of them started from the theoretical distribution of critical values to compute these values. Most of the students computed the quartiles in the empirical data, through different options such as frequency tables or statistical summaries; and a large proportion of students found no solution. In the following example the student is using the percentiles option in the software, which is appropriate only for computing median and quartiles in the empirical distribution. He is able to relate the idea of median to the 50th percentile, although he is unable to relate the ideas of quartiles and percentiles. Again, difficulties in discriminating between the theoretical and the empirical distribution are noticed.

Example 5

"These are the median and quartiles of the theoretical normal distribution for Age. The median is 13. Percentiles for Age: 1.0% = 12.0, 5.0% = 12.0, 10.0% = 12.0, 25.0% = 13.0, 50.0% = 13.0, 75.0% = 13.0, 90.0% = 14.0, 95.0% = 14.0, 99.0% = 14.0" (Student 1)

Table 3. Frequency and percentages of students' solutions classified by type of distribution

	Type of distribution used		
	Theoretical	Empirical	None
Correct	21 (17.9)	29 (24.8)	
Partly correct	9 (7.7)	14 (12.0)	1 (0.9)
Incorrect	2 (1.7)	17 (14.5)	4 (3.4)
No solution			20 (17.1)

Students' Reasoning and Understanding of Normal Distribution

Besides the percentage of correct responses to each question, we were interested in assessing the types of knowledge the students explicitly used in their solutions. Using the categorization in the theoretical framework we described in Section 2, we analyzed the students' protocols to provide a deeper picture of the students' reasoning and their understanding of normal distributions. Four students were also interviewed after they completed the tasks. They were asked to explain their procedures in detail and, when needed, the researcher added additional questions to clarify the students' reasoning in solving the tasks. In this section we analyze the results, which are summarized in Table 4 and present examples of the students' reasoning in the different categories.

Symbols and Representations

Many students in both groups correctly applied different representations, with a predominance of density curves, and a density curve superimposed onto a histogram. Their success suggests that students were able to correctly interpret these graphs, and could find different properties of data such as symmetry or unimodality from them as in Example 6, where there is a correct use of two graphs to assess symmetry.

Example 6

“You can see that the distribution of the variable weight is not symmetrical, since the average is not in the center of the variable range (Figure 3). The areas over and below the centre are very different. When comparing the histogram with the normal density curve, this skews to the left” (Student 5).

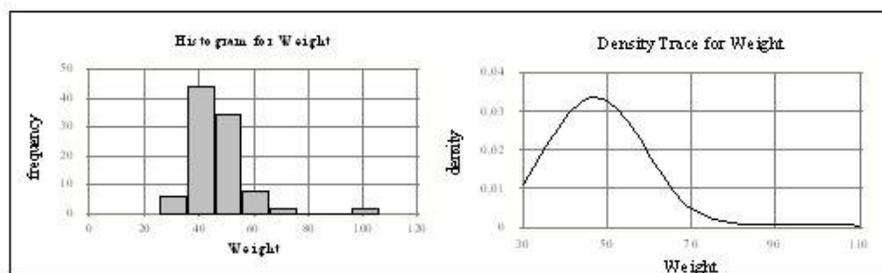


Figure 3. Histogram and density trace for Weight.

Among numerical representations, the use of parameters (mean and standard deviation) was prominent, in particular to solve task 2. Statistical summaries were correctly applied when students computed the asymmetry and kurtosis coefficients, and incorrectly applied when they computed the median and quartiles, since in that question the students used the empirical distribution instead of the theoretical curve (e.g., in Example 5). Few students used frequency tables and critical values. We conclude that graphical representations were more intuitive than numeric values, since a graph provides much more information about the distribution, and the interpretation of numerical summaries requires a higher level of abstraction.

Actions

The most frequent action was visual comparison (e.g., Examples 2, 6), although it was not always correctly performed (such as in Example 2, where the student was unable to use the graph to assess the kurtosis). A high percentage of students correctly compared the empirical density correctly with the theoretical normal density (e.g., Example 6). However, 40% of the students confused these two curves.

Table 4. Frequency of main elements of meaning used by the students in solving the tasks

Elements of Meaning	Correct Use	Incorrect Use
Symbols and Representations		
<i>Graphical representations</i>		
Normal density curve	45 (38.5)	1 (0.9)
Over imposed density curve and histogram	30 (25.6)	
Normal probability plot	6 (5.1)	
Cumulative density curve	2 (1.7)	
Histogram	37 (31.6)	
Frequency polygon	12 (10.3)	
Box plot	2 (1.7)	
Symmetry plot	1 (0.9)	
<i>Numerical summaries</i>		
Critical values	29 (24.8)	4 (3.4)
Tail areas	3 (2.6)	5 (4.3)
Mean and standard deviation (as parameters in the distribution)	48 (41.0)	3 (2.6)
Goodness of fit test	2 (1.7)	2 (1.7)
Steam-leaf	5 (4.3)	
Summaries statistics	59 (50.4)	47 (40.2)
Frequency tables	26 (22.2)	
Percentiles	9 (7.7)	9 (7.7)
Actions		
Computing the normal distribution parameters	50 (42.7)	18 (15.4)
Changing the parameters	10 (8.5)	2 (1.7)
Visual comparison	56 (47.9)	49 (41.9)
Computing normal probabilities	13 (11.1)	1 (0.9)
Finding critical values	28 (23.9)	68 (58.1)
Descriptive study of the empirical distribution	39 (33.3)	8 (6.8)
Finding central interval limits	14 (12)	
Concepts and properties		
Symmetry of the normal curve	40 (34.2)	13 (11.1)
Mode, Unimodality in the normal distribution	32 (27.4)	16 (13.7)
Parameters of the normal distribution	51 (46.3)	16 (13.7)
Statistical properties of the normal curve	27 (26.1)	3 (2.6)
Proportion of values in central intervals	13 (11.1)	1 (0.9)
Theoretical distribution	48 (41.0)	50 (42.7)
Kurtosis in the normal distribution; kurtosis coefficients	27 (26.1)	1 (0.9)
Variable: qualitative, discreet, continuous	50 (42.7)	65 (55.6)
Relative position of mean, median, mode in a normal distribution	35 (29.9)	5 (4.3)
Skewness and standard skewness coefficients	34 (29.1)	1 (0.9)
Atypical value	5 (4.3)	
Order statistics: quartiles, percentiles	32 (27.4)	63 (53.8)
Frequencies: absolute, relative, cumulative	13 (11.1)	
Arguments		
Checking properties in isolated cases	18 (15.4)	3 (2.6)
Applying properties	58 (49.6)	7 (6.0)
Analysis	32 (27.4)	5 (4.3)
Graphical representation	58 (49.6)	36 (30.8)
Synthesis	26 (22.2)	4 (3.4)

For example, regarding the variable of *Age* (Figure 1a), the empirical density curve is clearly nonnormal (since there is no horizontal asymptote). The students who, instead of using this empirical density, compared the histogram with the normal theoretical distribution (Figure 1b) did not perceive that the histogram was not well fitted to the same, even when this was clearly visible in the graph.

A fair number of students correctly computed the parameters, although a large percentage made errors in computing the critical values for the normal distribution (quartiles and median, as in Example 5). Even when the computer replaces use of the normal distribution tables, it does not solve all the computing problems, since the students had difficulties in understanding the idea of critical values and in operating the software options. Finally, some students performed a descriptive study of data before fitting the curve.

Concepts and Properties

Students correctly used the different specific properties of the normal distribution as well as the definition of many related concepts. The most common confusion was thinking that a discrete variable with only three different values was normal (e.g., Examples 1, 5). This was usually because students were unable to distinguish between the empirical and the theoretical distribution. Other authors have pointed out the high level of abstraction required to distinguish between model and reality, as well as the difficulties posed by the different levels in which the same concept is used in statistics (Schuyten, 1991; Vallecillos, 1994).

An interesting finding is that very few students used the fact that the proportion of cases within one, two, and three standard deviations is 68%, 95%, and 99%, even when we emphasized this property throughout the teaching. This suggests the high semiotic complexity required in applying this property where different graphical and symbolic representations, numerical values of parameters and statistics, concepts and properties, and actions and arguments need to be related, as shown later in Example 7.

The scant number of students who interpreted the kurtosis coefficient, as compared with the application of symmetry and unimodality, is also revealing. Regarding the parameters, although most students used this idea correctly, errors still remain. Some students correctly compared the relative position of the measures of central position in symmetrical and asymmetrical distributions, although some of them just based their selection on this property and argued it was enough to assure normality.

Arguments

The use of graphical representations was predominant in producing arguments. In addition to leading to many errors, this also suggests the students' difficulty in producing high-level arguments such as analysis and synthesis. Most students just applied or checked a single property, generally symmetry. They assumed that one necessary condition was enough to assure normality. This is the case in Example 7,

where the student correctly interprets symmetry from the symmetry plot and then assumes this is enough to prove normality.

Example 6

“We can graphically check the symmetry of *Time spent to run 30 Mts. in December* with the symmetry plot (Figure 4), as we see the points approximately fit the line; therefore the normal distribution will fit these data” (Student 6).

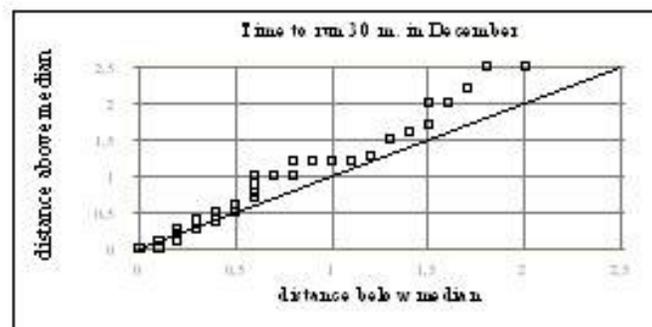


Figure 4. Symmetry plot.

In other cases the students checked several properties, although they forgot to check one of the conditions that is essential for normality, such as in the following interview, where the student studied the type of variable (discrete, continuous), unimodality, and relative position of mean, median and mode. However, he forgot to assess the value of the kurtosis coefficient, which is too high for a normal distribution (Student 7):

Teacher: In the exam you selected *Time to run 30 Mts. in December* as a normal distribution. Why did you choose that variable?

Student: I first rejected all the discrete variables since you need many different values for a discrete variable to be well fitted to a normal distribution. Since the two variables *Time to run 30 Mts. in December* and *Time to run 30 Mts. in September* are continuous I took one of them at random. I just might also have taken *Time to run 30 Mts. in September*. Then I realized the variable has only one mode, the shape was very similar to the normal distribution, mean and median were similar.

Teacher: Did you do any more analyses?

Student: No, I just did those.

A small number of students applied different elements of meaning, and carried out an analysis of each property. Seven percent of them produced a final synthesis, such as the following student.

Example 8

"The variable *Heartbeats after 30 press-ups* is what I consider best fits a normal distribution. It is a numerical variable. The variable is symmetrical, since both the histogram and the frequency polygon (Figure 5) are approximately symmetrical. On the other hand the skewness coefficient is close to zero (0.0109) and standard skewness coefficient falls into the interval $(-2, +2)$. We also observe that the kurtosis coefficient is close to zero (-0.1977) which suggests the variable can fit a normal distribution.

Furthermore, we know that in normal distributions, mean median and mode coincide and in this case the three values are very close (Mean = 123.4; Mode = 120; Median = 122). Moreover there is only one mode. As for the rule 68,95,99.7 in the interval $(\mu - \sigma, \mu + \sigma) \Rightarrow (105.08, 141.82)$ there are 68.75% of the observations, in the interval $(\mu - 2\sigma, \mu + 2\sigma) \Rightarrow (86.81, 160.19)$ there is 95.84% and in the interval $(\mu - 3\sigma, \mu + 3\sigma) \Rightarrow (68.34, 178.56)$ we found 100% of the data. These data are very close. Therefore you can fit a normal distribution to these data" (Student 8).

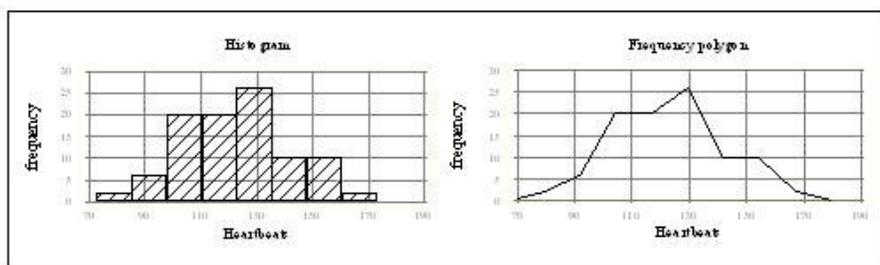


Figure 5. Histogram and frequency polygon for *Heartbeats after 30 press-ups*.

In this answer, the student relates the property of symmetry (concept) to the histogram and frequency polygon (representations). He is able to compute (action) the skewness and kurtosis coefficients (numerical summaries) and compares their values with those expected in normal distributions (properties and concepts). He also applies and relates the property of relative positions of central tendency measure and central intervals in a normal distribution, being able to operate the software (action) in order to produce the required graphs and summaries, which are correctly related and interpreted. This type of reasoning requires the integration of many different ideas and actions by the student.

Other students provided incorrect variables, even when they were able to use the software and to correctly produce a great number of different graphs. In Example 9 the student is able to plot different graphs and compute the quartiles. However, he is neither able to extract the information needed to assess normality from these graphs nor capable of relating the different results with the concepts behind them. No arguments linking these different representations or supporting his election are given. Moreover, he did not relate the high kurtosis coefficient to a lack of normality. The graphs and statistics produced are presented in Figure 6.

Example 9

“I selected *Height* since the normal distribution is used to describe real data. And describing the students’ height is a real biological problem. This is also a quantitative variable and normal distribution describes quantitative variables” (Student 9).

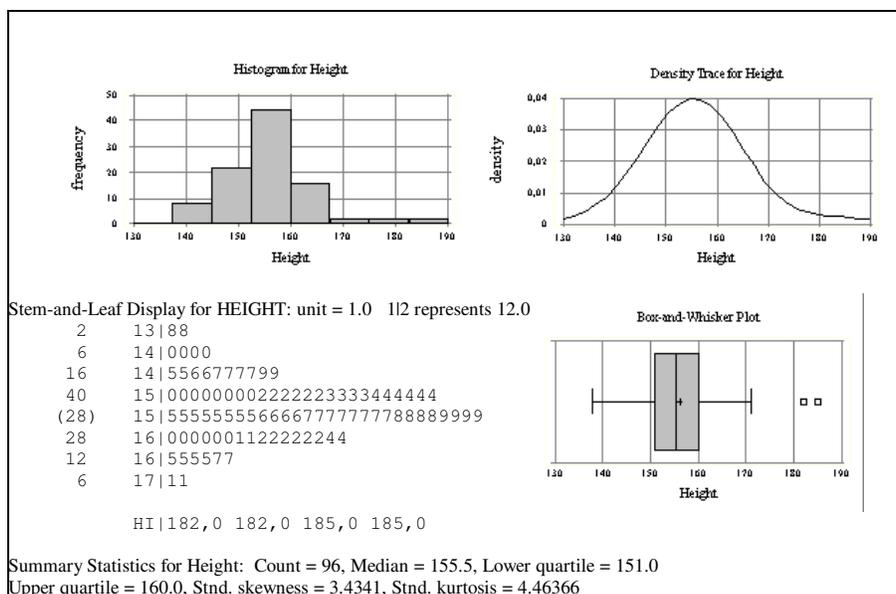


Figure 6. Graphical representations and statistical summaries for *Height*.

Discussion

Many students grasped the idea of model, and showed a good understanding of the usefulness of models, density curves, and areas under the normal curve. Our analysis of the various actions, representations, concepts, properties, and arguments used by the students in solving the tasks suggests that many students were able to correctly identify many elements in the meaning of normal distribution and to relate one to another. Some examples are as follows:

- Relating concepts and properties. For example, relating the idea of symmetry to skewness coefficient or to relative position of mean, median, and mode in Examples 6, 7, and 8.
- Relating graphical representations to concepts. For example, relating the empirical histogram and density curve shapes to the theoretical pattern in a normal curve (e.g., in Example 8).

- Relating the various graphic representations and data summaries to the software options and menus they need to produce them (relating representations and actions in all the examples).
- Relating the definition and properties of normal distribution to the actions needed to check the properties in an empirical data set (e.g., in Example 8).
- There was a good understanding of the idea of mean and standard deviation and its relationship to the geometrical properties of the normal curve (e.g., Example 2).

There was also a clear disagreement between the personal meaning of normal distribution acquired by the students and the meaning we tried to teach them. Here we describe the main difficulties observed:

1. Perceiving the usefulness of theoretical models to describe empirical data. This is shown in the following transcript (Student 10):

Teacher: Now that you know what the normal distribution is, can you tell me what it is useful for or in which way you can apply the normal distribution?

Student: For comparing, isn't it? For example to compare data and tables, it is difficult to explain. ... You have some data and you can repeat with the computer what we did in the classroom.

2. 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 the normal distribution or to the use of computers, and the student should have learned it at the secondary school level. However, in the following interview transcript, the student is not aware of the effect of interval widths on the frequency represented, which is given by the area under the histogram (Student 10):

Teacher: How would you find the frequency in the interval 0–10 in this histogram?

Student: The frequency is 5, this is the rectangle height.

Teacher: What about the frequency for the interval 10–30?

Student: It is 10, that is the height of this rectangle.

3. Interpreting probabilities under the normal curve. The graphical representation of the areas under the normal curve is the main didactic tool for students to understand the computation of probabilities under the curve and, at the same time to solve different problems involving the normal distribution. However, for some students with no previous instruction, this computation was not easily understood and performed.
4. We also observed difficulties in discriminating between empirical data and mathematical models, interpreting some statistical summaries and graphs,

and a lack of analysis and synthesis ability to relate all these properties when making a decision (Student 11).

Teacher: When you computed the median and quartiles in question 3, which data did you use: the theoretical normal distribution you fit to the data or the real data?

Student: I ... I am not very sure. Well, I used the data file ...

5. There was a great deal of difficulty in discriminating between the cases where a discrete quantitative variable can and cannot be fitted by a normal distribution (e.g., in Example 5) and even in distinguishing between the different types of variables.
6. Other students misinterpreted the skewness coefficient or assumed that the equality of mean, median and mode was enough to show the symmetry of the distribution, accepted as normal a distribution with no horizontal asymptote, made a rough approximation when formally or informally checking the rule ($\mu - k\sigma$, $\mu + k\sigma$), accepted too many outliers in a normal distribution, or misinterpreted the values of kurtosis.

Even when most of the students were able to change from the local to the global view of data (Ben-Zvi & Arcavi, 2001) in taking into account the shape of graphs as a whole, the idea of distribution as a property of a collective, and the variability of data, there is still a third level of statistical reasoning many of these students did not reach. This is the *modeling* viewpoint of data, where students need to deal at the same time with an empirical distribution as a whole (therefore, they need to adopt a global viewpoint of their data) and the mathematical model (the normal distribution in our research). In this modeling perspective, students need to concentrate on the different features of the data set as a whole and on the different features of the model (type of variable, unimodality, skewness, percentage of central cases, horizontal asymptote, etc., in our case). In addition to understanding the model as a complex entity with different components, they should be able to distinguish the model from the real data, to compare the real data to the model, and to make an accurate judgment about how well the model fits the data.

There was also difficulty in using secondary menu options in the software—which, however, are frequently essential in the analysis. Finally, the students showed scant argumentative capacity, in particular regarding analysis and synthesis (e.g., in Example 9).

IMPLICATIONS FOR TEACHING NORMAL DISTRIBUTIONS

The main conclusion in this study is that the normal distribution is a very complex idea that requires the integration and relation of many different statistical concepts and ideas. Recognizing this complexity, our work also suggests that it is possible to design teaching activities that facilitate the learning of basic notions

about normal distribution. Since the learning of computational abilities is no longer an important objective, an intuitive understanding about basic concepts is possible for students with moderate mathematical knowledge, whenever we choose appropriate tasks.

Working with computer tools seemed to promote graphical understanding, as students in our experiment easily recognized and used many different plots (such as density curves, histograms, etc.) to solve the problems proposed. Moreover, they also showed a good understanding of many abstract properties, such as the effect of parameters on the density curve shape, and made extensive use of graphs as part of their argumentation. This suggests the essential role of computers to facilitate students' exploration of these properties and representations.

It is important that students understand basic concepts such as probability, density curve, spread and skewness, and histograms before they start the study of normal distribution; its understanding is based on these ideas. They should also be confident in the use of software before trying to solve problems related to the normal distribution, since they often misinterpret or confuse results from different software options.

The student's difficulties in discriminating between theoretical models and empirical data suggest that more activities linking real data with the normal model are needed. Simulating data from normal distributions and comparing them with real data sets might also be used as an intermediate step between mathematical model and reality. As a didactic tool it can serve to improve students' probabilistic intuition, to teach them the different steps in the work of modeling (Dantal, 1997), and to help them discriminate between model and reality. Simulation experiences and dynamic visualization can contribute, as analyzed by Biehler (1991), to provide students with a stochastic experience difficult to reach in the real world.

Finally, it is important to take into account the different components of meaning and understanding when assessing students' learning. Computer-based assessment tasks in which students are asked to analyze simple data sets and provide a sound argument for their responses—such as those presented in this paper—are a good tool to provide a complete picture of students' understanding and ways of reasoning.

REFERENCES

- 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., Tauber, L., & Sánchez, V. (2001). *Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos* (Meaning and understanding of normal distributions in an introductory data analysis course). *Cuadrante*, 10(1), 59–92.
- 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 & M. Borovcnick (Eds.), *Chance encounters: Probability in education* (pp. 169–211). Dordrecht, The Netherlands: Kluwer.

- Dantal, B. (1997). Les enjeux de la modélisation en probabilité (The challenges of modeling in probability). In *Enseigner les probabilités au lycée* (pp. 57–59). Reims: Commission Inter-IREM Statistique et Probabilités.
- delMas, R. C., Garfield, J. B., & Chance, B. (1999). *Exploring the role of computer simulations in developing understanding of sampling distributions*. Paper presented at the *Annual Meeting of the American Educational Research Association*, Montreal, Canada.
- Gal, I. (1997). Assessing students' interpretations of data: Conceptual and pragmatic issues. In B. Phillips (Ed.), *Papers on Statistical Education presented at ICME-8* (pp. 49–58). Swinburne, Australia: Swinburne University of Technology.
- Garfield, J. B. (1991). Evaluating students' understanding of statistics: Developing the statistical reasoning assessment. In R. G. Underhill (Ed.), *Proceedings of the 13th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 1–7). Blacksburg, VA: Comité organizador.
- Godino, J. D. (1996). Mathematical concepts, their meaning and understanding. In L. Puig & A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 417–424). Valencia: Comité organizador.
- Godino, J. D., & Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. In A. Sierpiska & J. Kilpatrick (Eds.), *Mathematics Education as a research domain: A search for identity* (pp. 177–195). Dordrecht: Kluwer.
- Huck, S., Cross, T. L., & Clark, S. B. (1986). Overcoming misconceptions about z-scores. *Teaching Statistics*, 8(2), 38–40.
- Méndez, H. (1991). *Understanding the central limit theorem*. Ph.D. diss., University of California, University Microfilm International number 1-800-521-0600.
- Piaget, J., & Inhelder, B. (1951). *La gènesis de l'idée de hasard chez l'enfant*. Paris: Presses Universitaires de France.
- Rubin, A., Bruce, B., & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics* (pp. 314–319). Voorburg, The Netherlands: International Statistical Institute.
- Schuyten, G. (1991). Statistical thinking in psychology and education. In D. Vere-Jones (Ed.), *Proceedings of the III International Conference on Teaching Statistics* (Vol. 2, pp. 486–489). Dunedin, Australia: University of Otago.
- Vallecillos, A. (1996). *Inferencia estadística y enseñanza: Un análisis didáctico del contraste de hipótesis estadísticas* (Statistical inference and teaching: A didactical analysis of statistical tests). Madrid: Comares.
- Vallecillos, A. (1999). Some empirical evidence on learning difficulties about testing hypotheses. *Bulletin of the International Statistical Institute: Proceedings of the Fifty-Second Session of the International Statistical Institute* (Tome LVIII, Book 2, pp. 201–204). Helsinki, Finland: International Statistical Institute.
- Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223–265.
- Wilensky, U. (1995). Learning probability through building computational models. In D. Carraher y L. Meira (Eds.), *Proceedings of the 19th PME Conference* (Vol. 3, pp. 152–159). Recife, Brazil: Organizing Committee.
- Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33, 171–202.