INVERSE NUMERICAL FUNCTION CONCEPT AND COMPUTER SOFTWARE LEARNING
Or the necessity of distinguishing between D# and Eb

Catherine-Marie CHIOCCA
Assistant Professor of didactics of Mathematics
Ecole Nationale de Formation Agronomique,
Département Culture Langage Education et Formation
B.P. 87, 31326 Castanet-tolosan
tel : 00 33 5 61 75 32 82 ; Fax : 00 33 5 61 75 03 09
catherine-marie.chiocca@educagri.fr

ABSTRACT

We are in the context of a course on hypothesis tests and computing for French students taking plant biotechnology option for a professional degree.

It reports on a study in this context about statistical and numerical concepts and the tools used to determine them.

The question raised in this paper is:

how do the teacher's instructions to systematically use the graphic frame for determining the critical threshold with a table or with a spreadsheet, and the rejection area, explained as such, affect the students' work?

Keywords : obstacles, spreadsheet, inverse numerical function

1. Introduction

During the testing of a hypothesis or when looking for the critical value of a given risk threshold, two functions are involved: the random variable which follows a known distribution for the zero hypothesis and the inverse of the distribution function. Students appear to see the function for a given distribution as a 'black box' which transforms real numbers (values of the random variable) into probabilities which are also real numbers taken between 0 and 1. In this case, difficulties in doing from discrete to continuous representations (a point becomes a surface area) are encountered (Schneider 1991).

This paper recounts part of a study of phenomena produced when teaching the construction of statistical tests to third year university students.

2. Theoretical context

The work combines the following theoretical frameworks :

1 – The cognitive approach to instruments developed by Pierre Rabardel (Rabardel 1995). This especially involves investigating the instrumentation[1] of the spreadsheet for determining the significant value in a hypothesis test.

2 – The didactical engineering framework developed by Michèle Artigue (Artigue 1990). There is a dialectic between students' availability spreadsheets and their mathematical knowledge.

3. Discussion

The work is based on the hypothesis (developed during the author's work on her doctoral thesis) of the importance of teachers reflecting on and structuring knowledge discourses for learning purposes. The work is thus part of a wider study of meaningful reflection in order to clarify the concept of a probability distribution. An investigation of the kinds of obstacles[2] that have been encountered by students studing the Plant biotechnology leading to a professional university degree, gave the following main results:

-mathematical obstacles: there is some confusion between the availability of the direct function and the inverse numerical function.

-Didactical obstacles: some students do not use the trial and error possibilities offered by spreadsheet software.

-Computing obstacles: there is some confusion between the Gauss law inverts and the Gauss law spreadsheet function. Some students cannot manage without the teacher's help.

The conclusion was that it seems students are more likely to accomplish the task if they both have some theoretical knowledge **and** use the trial and error strategy.

A special course on this difficulty was designed to clarify the 'confusion' phenomenon when teaching the distribution function and its inverse function (both in mathematics and with the spreadsheet).

---

[1] Instrumentation is the aspect of the instrumental genesis process when the subject is developing an instrument. The instrument is assumed to consist of two elements: an artefact and one or several corresponding user 'schèmes'

[2] Chiocca 2002.

The question raised in this paper is:

how do the teacher's instructions to systematically use the graphic frame for determining the critical threshold with a table or with a spreadsheet, and the rejection area, explained as such, affect the students' work?

4. Observation protocol and description of the research setting

The experimental situation is build to enable observation of a few schemes for spreadsheet software, taken as an artefact. A clinical, rather than a statistical approach, to the experiment and the observation was chosen as this would provide a sufficiently rich data set to reveal the students' errors.

The study used the students' final examination work (paper and diskette), video recordings and transcriptions of these recording. The data collection protocol enabled us to focus on tasks and techniques (Chevallard 1992). The filmed session completed the analysis of tasks and techniques used by students.

The population: consisted of 12 university students studing Plant biotechnologyat the universitéy level. The students have a school graduate and two years of higher education. They are on average 22 years old. They had previous experience of applying inferential statistics during their studies. For two years they had been steeped in conformance and adjustment tests and variance analysis. Since the official instructions of the Ministry of Agriculture excluded teaching theoretical developments, they had no knowledge of statistical theory and only a little of probability. They will have to use such statistical tests in their professional work, which for most of them will begin as soon as they have obtained their degree (95% of successfully qualified students find work within one year). They were thus very keen to learn how to use spreadsheet software for implementing the test.

5. The experimental situation

1 – *Elements of didactical engineering*

The teaching sessions 18 hours in all, were divided between work in the classroom using the tables found in most statistics manuals, and work in the computer room using spreadsheet features which appear to eliminate the need for paper tables.

first, experimental exercises with paper, pencil and spreadsheet software on the distribution law were given to the students to work on direct and reverse reading of a paper table and also on the availability of spreadsheet functions: Student's law and Student's law inverts.

When constructing a test of a hypothesis, the diagrams representing the density curves of usual laws, the abscissa of critical thresholds for a given probability (a surface area beneath a curve) and reject areas (IR intervals) are institutionalised[3].

2 – *Task*

The task to be accomplished by the student was as follows:

The examination consisted of four test constructions: two tests for comparing means with the same variance values and two independence tests using two random variables.

The researcher (teacher) asks the students to draw the reject areas for the zero hypothesis. The point was to see if they could produce such a drawing after 18 hours of class.

[3] Cf. the institutionalised diagram are shown in appendix

The students were allowed to use any documents they wanted, plus a pocket calculator, their course notes, statistical manuals and particularly the two artefacts: paper tables and the spreadsheet software.

3 - *A priori* obstacles

Mathematical:

The numbers written in the cells of the paper table are probabilities, in other words, surface areas under the density curve and are thus obtained by means of the distribution function. Given the probability, they are then told to look for the abscissa value. The work involves using the inverse function.

The risk threshold is the image obtained by the distribution function of a certain value called the critical threshold, which is thus the antecedent obtained, by the distribution function for the risk threshold. Learning these two concepts (antecedent, image, for a numerical function) is difficult. In France, students are taught the concept of image and antecedent for a numerical function three years before the final leaving examination and it is not explicitly reviewed in the official programme for later classes. However, teachers do review these concepts even if they do not do any explicit work on them.

Didactical:

Several critical values had been calculated when students were learning how to construct a test. Difficulties were encountered when, in the same session, the direct function was used (for instance when calculating the actual risk) and its reciprocal function (for finding the critical values which determined the reject area). It is thus likely that the same difficulty will occur for the learning task. Arguments frequently used in favour of spreadsheet software are that it eliminates the need for paper tables. A table is no longer necessary for finding a critical value and a reject area. But, without the paper table, there is no longer any graphic representation.

Since students are only able to use a little probability and statistical theory[4], assuming that they have any[5], it is difficult to explain meaningfully the search for reject areas and critical thresholds.

The word *zone* in French usually represents a surface area rather than a segment. It is thus likely to cause confusion between the reject area and the risk threshold (an interval or a set of IR intervals and probability, area under the curve). In English, the use of the word area would probably cause even more confusion.

Computing:

The spreadsheet help button provides instructions but no graphic representation.

In French, the spreadsheet specifies that LOI.STUDENT.INVERSE (the same for LOI.NORMALE.INVERSE) returns the value of a random variable according to Student's law as a function of the probability and the number of degrees of freedom. In French, the value of a random variable (which is a numerical function) corresponds to the numerical value taken by the random variable as a numerical function (for a continuous law) and not to the antecedent of a value for the distribution function.

The LOI.STUDENT INVERSE function requires giving the total risk even in the case of a bilateral test.

The names of laws or their reciprocal laws are not homogenous. The spreadsheet offers KHIDEUX.INVERSE and INVERSE.LOIF and not LOI. KHIDEUX.INVERSE and LOIF.INVERSE for the same model as LOI.NORMALE. INVERSE or LOI.STUDENT.INVERSE.

---

[4] As used by Aline Robert: a knowledge, which can be used, is knowledge that a student can look for to use without it being suggested explicitly or implicitly.

[5] As used by Aline Robert: available knowledge is knowledge, which a student knows how to use if told that it is that knowledge which he needs to solve the problem.

6. Some results

Depending on the individuals, the task took at the least half an hour to solve with no assistance from another person.

The students produced two sorts of diagrams: a category of 'right' diagrams and another showing that the risk threshold and the critical threshold are confused. In other words, for image and antecedent by means of the distribution function, 6 out of the 12 students suggested false diagrams, 4 of them 'right' diagrams, in other words, diagrams which matched the expectations of the teacher-researcher, 2 of them suggested no diagrams in spite of the explicit instructions of the examination. The two students who did not produce any diagrams only attended 6 out of the 18 course hours. One may conclude that they did not spend enough time in the course to understand the didactical contract concerning these diagrams.

All of the students found the right value for the critical threshold, 10 out of 12 students found it by reading the paper table (by reading it backwards is what they said) whether the diagrams were right or wrong. One of the two students who used the spreadsheet made a false diagram, the other no diagram. Is this because the spreadsheet, while eliminating the need for the paper table, blocks out the graphic diagram at the same time?

The student who did not make any diagram appears to be in some kind of transition with respect to the use of the two artefacts: paper table and spreadsheet. Indeed, in the spreadsheet, he displayed the critical value by using the Student's law inverts function while at the same time specifying 'after reverse reading in the table'.

7. Concluding comments

Systematic use by the teacher of the graphic frame for determining critical thresholds and reject areas in this didactical framework, leads to one out of two students producing false diagrams.

On the other hand, most of the students were not able to use the spreadsheet features to look for critical thresholds.

There is a need to work on the concepts of antecedents and image. In the 'succession' of functions between the random variable and the distribution function, a distinction should be made between the image by the random variable and the antecedent by the distribution function. This would then make it easier to understand when distinguishing the real risk from the critical threshold.

These two conceps could be compared to the concept of coma in music: on certain instruments such as the violin, D# et Eb are not confused.

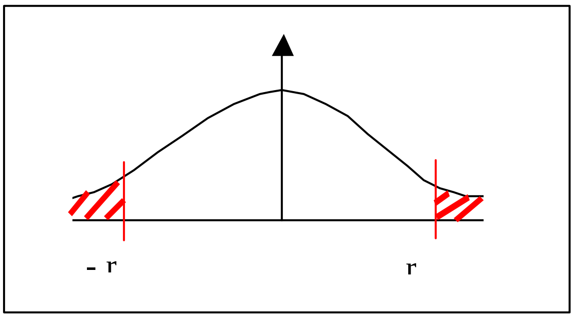
REFERENCES

- Artigue M. (1990), *Ingénierie didactique*, Recherches en Didactique des Mathématiques, 9 (3), 231-308. Ed. La pensée sauvage.
- Chevallard, Y., (1992) *Concepts fondamentaux de la didactique : perspectives apportées par une approche anthropologique*. recherches en didactique des mathématiques , 12 (1), 73-112. Ed. La pensée sauvage.
- Chiocca C-M. (1995), *From the classroom discourse of mathematics teachers to the mathematical representations of their students: an essay on didactical reflection*, These Paris VII.
- Chiocca C-M., (2002) CERME2 electronics proceedings
- Portugais J. (1996) *Esquisse d'un modèle des intentions didactiques*, Actes des 2ndes journées didactique de la Fouly, Editions Interactions didactiques
- Actes de la Xième Université d'été de Didactique des Mathématiques, Ed. IUFM de Caen, p201 à 235 :
  - Rabardel P. (1999) *Eléments pour une approche instrumentale en didactique des mathématiques,*
  - Lagrange J-B. (1999) *Les instruments de calculs formels,*
  - Laborde C. (1999) *L'activité instrumentée par des logiciels de géométrie dynamique,*

- Rabardel, P., (1995). *Les hommes et les technologies*. Paris : Armand Colin
- Robert A. (1988) *grille d'analyse des exercices de manuels scolaires* DIDIREM
- Schneider, M., (1991). *Un obstacle épistémologique soulevé par des "découpages infinis" des surfaces et des solides*. Recherches en didactique des mathématiques, 11 (2.3), 241-294
- Tavignot P. (1997), *Macro-systèmes de protocoles dans le cadre théorique de la transposition didactique*, Actes des premières journées de la Fouly, Edition Interactions Didactiques

Appendix

Figure : bilateral test



The most frequent error consists of writing the following comment:
the rejected area is in the shaded area.