Statistical Inference In Textbooks: Mathematical And Everyday Contexts

Israel García Alonso & Juan Antonio García Cruz
Universidad de La Laguna (Spain)

Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education, Vol. 2, pp. 257-264. Seoul, Korea.

STATISTICAL INFERENCE IN TEXTBOOKS: MATHEMATICAL AND EVERYDAY CONTEXTS

Israel García-Alonso & Juan Antonio García-Cruz

Universidad de La Laguna (Spain)

Various terms in the field of Statistical Inference and their presentation in secondary school text books are examined. A comparison of these terms in secondary school textbooks is carried out against their meanings in everyday use as well as in the mathematical context from two standard university textbooks from the field. We offer evidence that the meanings are not necessarily the same and that in some cases the definition which appears in the secondary school textbook is closer to its everyday use than to its mathematical one. Some implications for school textbook writers are derived.

THEORETICAL FRAMEWORK

Changes in mathematics secondary school curriculum, including statistics, have taken place in several Western European countries, such The Netherlands and Spain, since the late 70´s. Some statistical concepts which were previously introduced during the early years at the university level are now being taught at the secondary level, i.e. confidence intervals and hypothesis testing based on the normal distribution. Some recommendations have even been made to introduce basic concepts of statistical inference during earlier schooling (NCTM, 2000), but of course without the required sophistication and formalization seen at the university levels.

Ample research into the difficulties and obstacles that students encounter when facing statistical inference have also appeared lately. Vallecillos and Batanero (1997) identified difficulties in the learning and understanding of statistical inference in the university context, especially with respect to the concepts of significance levels, parameters, and a statistic, among others, in addition to a general understanding of the logic involved in hypothesis testing. Moreno and Vallecillos (2002), researched the secondary school setting. Their study of 15 and 16 year-old students showed that students had misconceptions about statistical and carried incorrect inferences. They identify *Representativeness* as the key concept which presents the most difficulty (Kahnemann et al., 1982). Specifically, they point out that *Representativeness* is characterized by the belief that small samples must reproduce the essential characteristics of the population from which it has been taken. Students also find Hypothesis Testing to be difficult. Vallecillos (1999) indicates that students possess different ideas about exactly what a hypothesis test is. García-Alonso and García-Cruz (2003) carried out a study using a sample of (n=50) students who sat for the University Entrance Examination. They concluded that most students (86%) were unable to completely carry out those problems in the exam which dealt with Statistical Inference, even though these exercises were no different than the typical

ones that were covered in their daily classes. Previous research into this subject confirms that an inherent difficulty is present in this topic based on the logical-mathematical point of view. However our research focuses on the use of language and the effect that possible obstacles present could have on the students understanding of the topic.

Language is an important tool in the construction of mathematical knowledge, and during lower secondary school, this language is no different than that used in an everyday context, except for some very specific terms. Higher levels of "technification" and abstraction are required in upper secondary school mathematics. Hence, a greater amount of specific terms are needed in their study, which is also the case for Statistical Inference. Shuard & Rothery (1984) indicated that two contexts are present in the mathematics classroom: an everyday context and a mathematical context. An everyday context refers to carrying out typical communication found in day to day settings while a mathematical context only deals with mathematical communication. These authors categorize the terms which appear in the mathematics classroom according to these groups: (1) those terms which have the same meaning in both contexts; (2) those terms whose meaning changes from one context or the other, and (3) those terms which are only seen in a mathematical context. Students should not have any problems understanding the terms found in category (1). On the other hand, all terms found in category (3) must be defined since they are not part of a students' vocabulary. The terms found in category 2 may be difficult for students to understand. The role of textbooks was included in Discussion Group 7 at the 30th PME Conference. The need to carry out research on European textbooks was identified, and one of the specific areas covered the need to understand how textbooks are prepared and used during the learning and teaching process (Pepin, Grevholm & Straesser, 2006).

Our research into textbooks is based on the "a priori" approach. Thus we review the terms related with statistical inference and analyze them within their context. They are then classified according to the categories given by Shuard & Rothery (1984)

We wish to address the following questions: How are statistical inference terms presented in upper secondary school textbooks? Are there differences between the mathematical meaning of these terms and the ones from everyday use? Are the definitions in the textbooks correct?

OBJECTIVES AND METHODOLOGY

The are two main objectives in our research:

1. To determine if the meanings of the terms found in Statistical Inference is the same in both everyday and mathematical contexts.
2. To compare these meanings with those developed in secondary school textbooks.

Questionnaires were administered to a group of 37 secondary school teachers who were responsible for teaching the last year of an upper secondary school mathematics course. Teachers were asked to identify the publisher of the textbook they were using.

The questionnaire also revealed a very interesting fact. Approximately 30 teachers in the survey used the textbook as a means to recall and/or review the statistical concepts that their students would be working with in class. This decision emphasizes the importance of the textbook as a key element in teaching training.

We chose the four most popular textbooks according to the survey. The number in parentheses represents amount of times this book was used among the four in the study. The four publishers are: Publisher 1 (P1): Anaya (n = 24); Publisher 2 (P2): SM (n = 6); Publisher 3 (P3): Santillana (n = 4); Publisher 4 (P4): Edelvives (n = 2).

We then selected all the terms related with Statistical Inference in each publisher´s version. The next step was to analyze the meaning of each term in everyday use as well as in the mathematical context. *The Diccionario de la Real Academia de la Lengua Española* (hereafter *Diccionario*) was consulted for the definitions of the terms in everyday context. Two university textbooks were used for the mathematics definitions: ME = Mendenhall (1982) and MO = Moore (2005).

Everyday context: The Spanish language is regulated by the Real Academia de la Lengua. The mission of the Real Academia is to collect all terms and new versions of their meanings which have been introduced into the language with the passing of time. The *Diccionario* is an important resource when studying the meanings of terms found and used in everyday language. *The dictionary also makes a reference to technical terms, "to introduce those words which originate from distinct fields of knowledge and also from professional activities whose current use (...) has exceeded its original meaning in another setting, and consequently has extended its use, either frequently or occasionally from either common language or in a cultural context".*

Mathematical context: The university textbooks are used to identify the mathematical context of the terms, since they are written with the understood mission to convey the definitions of these terms and technical concepts to students. As mentioned previously, Mendenhall (1982) and Moore (2005) were used in our study.

Two definitions for each term were taken from the *Diccionario*. The first definition is the one which appears first in the *Diccionario*. The second definition is the one which is closest to the mathematical context. The definitions from the *Diccionario*, the University textbooks and the secondary school textbooks which appear in this paper have been translated into English from the Spanish versions. The English translation of the statistical inference terms are provided to guide the reader through our research methodology.

Now that the meanings of the terms from both the everyday and mathematical contexts are known, the next step is to categorize them according to the criteria given

in Shuard & Rothery (1984). If the definitions given in the *Diccionario* and the university textbooks are the same or quite similar, then the term is placed in the first category *"same meaning in both contexts"*. The third category "s*pecific meaning in mathematical context"* is used for those terms that are only found in university textbooks. Category 2 "*different meanings in both contexts"* is the remaining category and is used for those terms whose meanings are not the same. The final stage of our research analyzed the treatment of these terms in the textbooks from the four publishers.

RESULTS AND DISCUSSION

27 terms related to statistical inference were analyzed. Each term was grouped into its corresponding category. Each term was analyzed according to the procedure described in the previous section. Some examples by category follow.

Category 1: *Same meaning in both contexts*

Four terms were found to belong to this category: *Statistics, Population, Individual* and *Sample size*. We use the *Population* term as an example of how we carried out our study:

<table>
<tr><td colspan="2">Term: POPULATION</td></tr>
<tr><td>Diccionario</td><td>University Textbooks</td></tr>
<tr><td>1. Act and effect of populating.
2. Set of individuals or things subject to a statistical evaluation by means of sampling.</td><td>ME.- Set of all measurements of interest to the person who obtains the sample.
MO.- An entire group of individuals of which we want to know certain information about is called a population.</td></tr>
<tr><td colspan="2">PUBLISHERS</td></tr>
<tr><td colspan="2">P1.- "A population or universe is the set of all individuals in our study".
P2.- "is the set of all elements that possess a specific characteristic. Populations are generally assumed to be very large."
P3.- "when a statistical study refers to a group, set or collection of elements, this collection is called the population."
P4.- "the homogeneous group of people, animals or things on which a study is to take place".</td></tr>
</table>

A comparison of the four definitions reveals how the P4 version requires homogeneity within the group. This condition is not necessary, since the purpose of the statistical study could be to study a homogeneous characteristic within the population, not a homogeneous group. Thus, we consider the inclusion of homogeneity in the definition to be misleading, and possibly causing confusion to the student.

Category 2: *Different meanings in both contexts*

The following terms are included in this category: *Calculated Mean*, *Sample, Estimation, Infer, Distribution, Probability, Representative, Risk*.

<table>
<tr><td colspan="2">Term: SAMPLE</td></tr>
<tr><td>Diccionario</td><td>University Textbooks</td></tr>
<tr><td>1. Part of a product or merchandise which allows the quality of the goods to be known.
2. Part or chosen portion of a set by methods which allow it to be considered representative of it.</td><td>ME.- A sample is a subset of selected measurements from the population of interest.
MO.- A sample is that part of the population that we are currently studying with the objective of obtaining information.</td></tr>
<tr><td colspan="2">PUBLISHERS</td></tr>
<tr><td colspan="2">P1.- “a subset drawn from the population. A study of the sample helps to infer characteristics of the entire population”. “However, if the sample is incorrectly chosen, (it is not representative)…”
P2.- “a subset of the population”. If “a study is going to be reliable, it is critical that the selection of the sample be correct, so that it is clearly representative of the population”.
P3.- “part of the population, carefully selected, which is subject to scientific observation as a representation of the same population. Its purpose is to obtain valid results for the entire population”. In addition, “a sample is considered valid when it fulfils the definition of (…) being representative”
P4.- “Subset of the population”. “An appropriate selection” should be made.</td></tr>
</table>

This term has equivalent definitions in both university textbooks. However, the definitions in the *Diccionario* emphasize how the subset must be representative of the complete set. This requirement does not appear in the university textbooks**.**

Exactly what is representative? Two definitions are found in the *Diccionario*:

1. *To recall something with words or figures that the imagination remembers*.
2. *To be the image or symbol of something, or to imitate it perfectly*.

According to this definition, if a sample is said to be representative, it is indicating that the sample must “perfectly imitate” or “be an image” of the population. These definitions lead to an incorrect idea of the term, and lead to *Representativeness,* described in Kahneman et al. (1982). *Representativeness* is defined when the student expects that small samples reflect all of the population properties. We know that its similarity with the population does not validate our sample, but instead its selection method.

It is also important to remark how the *Diccionario* indicates the need for the sample to be representative. This fact shows how an incorrect meaning in the everyday context applied to mathematics can produce an incorrect understanding of this technical term, and create barriers for students in their understanding.

Category 3: *Specific meaning in mathematical context*

Category 3 contains the most terms of all the groups. The list includes: *Statistic, Parameter, Random sampling, Sample mean, Population Mean, Confidence level, Standard deviation, Significance level, Inductive statistics, hypothetical-deductive statistic, margin of error, Normal, Bias, Efficiency, Sample proportion.*.

Term: **STATISTIC**	
Diccionario	University Textbooks
1.Belonging to or related to statistics. 2.Person who exercises the statistics profession.	**MO.** A statistic is a number that can be calculated based on sample data without needing to use any unknown parameter. We typically usually use the term statistic to estimate an unknown parameter.
PUBLISHERS	
P1. & P4.- Undefined. **P2.-** *"a numerical value that describes a characteristic of the sample".* **P3.-** "*[values or measures] that characterize a sample".*	

The mathematical meaning of the term Statistic is not given in the *Diccionario*. Therefore it is included in the third category. When statistic is defined, it is done with references to population parameters instead of statistics. Given its importance in this concept, (statistical inference begins with this point), it is significant that some publishers exclude its definition, and that others introduce it later on in the text.

Term: **CONFIDENCE LEVEL**	
Diccionario	University Textbooks
Undefined	**ME.-** The probability that a confidence interval will include the estimated parameter. **MO.** A "C" confidence level represents the probability that the interval will contain the true value of a parameter in repeated sampling.
PUBLISHERS	
P1.- "*Starting with a sample of size n we can estimate the value of a parameter of the population (...) Resulting in an interval in which we are confident that this parameter will be included. (...) Finding the probability that such a thing occurs. This probability is called the confidence level".* **P2.-** "*the probability that an estimator for an interval covers the true value of the parameter which is being estimated. It is normally represented by 1 - α."* **P3.-** "*the level of confidence that we have that the population mean will belong to the interval is 1-α".* **P4.-***"Calculate the two values where we expect that the searched-for parameter will be found with a certain confidence level, which we will call 1-α, where α is the pre-determined risk level".*	

Note that publishers P2 and P3 use the probability that the estimator is included in the interval, but later on introduce new terminology to complete the definition, namely the value: 1-α. The value of α has not yet been defined. Publisher P4, however, states that α is the risk level (significance level). In other words, the definition of one term is given using another term which has not been previously defined.

Publishers P2 and P3 present an example which barely explains the meaning of a confidence level. None of the proposed activities emphasize this concept. Publisher P1 offers comments about the meaning of the concept, but Publisher P4 does not even offer an example which would help explain the meaning of a confidence level.

Recall that confidence level is a new term for secondary school students. A clear understanding of its meaning is not easy, since the concept of probability is dropped and instead levels are introduced, where the meaning of confidence appears opposite to the meaning of significance level, which represents an α error. This set of concepts and their meanings are not properly separated nor explained in a convenient manner, thereby complicating their understanding by students.

CONCLUSIONS

Our research includes the selection of specific terms related to confidence intervals and their classification according to three categories, given their meaning in two different contexts: everyday and mathematical.

The first category includes the terms found in the textbooks with the *same meaning* in both contexts. Most of the publishers provided the correct definition of the terms, although there are some examples where the definition has been significantly altered.

The second category is made up of terms with *different meanings* in both contexts. Hence, it is possible that a student can incorrectly learn the mathematical concept. We observed that most publishers use the everyday definition of the technical term, not the mathematical one.

The third category is made up of terms *only* found in the mathematical context. Some of the textbooks do not use the correct definitions. We also found examples where terms were introduced and never referred to later on in the textbook. The presentation of the material in this way can be confusing for students, or indicates that the textbook is inconsistent, since it is not continuously using terms which were introduced at a given moment.

Overall we have noticed how some definitions that appear in the textbooks do not correspond to their mathematical meaning but instead to the one in their everyday use. There are times when the definition of the term is incorrect, or it does not emphasize the key elements in the definition. Textbooks which include these errors are of great concern to students and teachers if we take into account that most professors use these textbooks to review the concepts related to Statistical Inference,

and then prepare their classes, as confirmed in the questionnaire we administered at the beginning of our study.

Incorrect definitions of technical terms, or their complete absence in the textbook has been confirmed in all four publishers. Sometime the textbook altered the meaning of the term. These differences in the definitions are not easily found, and the errors would be impossible to detect for students or anyone without a previous knowledge of statistics. We have confirmed that the terms found in the textbooks at this level are, for the most part, technical, and therefore the language used in their presentation should be consistent with their content. It is desirable that the definitions of these technical terms be correctly defined, leaving no doubt about their meaning.

In our opinion the results of our research are useful for writers of textbooks, leading to more coherent efforts by the writers, and offering a final product which is an important tool for both teacher and student alike.

BIBLIOGRAPHY

R.A.E (2001): *Diccionario de la lengua española*. Author: Madrid.

García-Alonso, I. & García-Cruz, J.A. (2003): "Algunos resultados sobre la actuación de los alumnos en las cuestiones de estadística en la P.A.U.". *Actas de las XI Jornadas sobre el Aprendizaje y la Enseñanza de las Matemáticas (JAEM)*, 733-738.

Kahnemann, D., Slovic, P. & Tversky, A. (1982): *Judgement under uncertainty: Heuristics and biases.* Cambridge University Press.

Mendenhall, W. (1982): *Introducción a la probabilidad y la estadística*. Wadsworth Int.

Moore, D. (2005): *Estadística aplicada básica.* Antoni Bosch Editor, Barcelona.

Moreno, A. J. & Vallecillos, A. (2002): "Exploración heurística y concepciones iniciales sobre el razonamiento inferencial en estudiantes de secundaria". *Educación Matemática.* 14(1), 62-84.

N.C.T.M. (2000): *Principles and standards for school mathematics.* Reston, Virginia.

Pepin, B; B. Grevholm & Straesser. (2006). DG07: The mathematics textbook- A Critical Artefact? In Novotná et al. *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education,* Charles University in Prague, Faculty of Education, Vol 1, p. 193.

Shuard, H. & Rothery, A (Eds) (1984): *Children reading mathematics*. London: Murray.

Vallecillos, A. & Batanero, C. (1997): "Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios". *Recherches en Didactique des Mathématiques,* 17(1), 29-48.

Vallecillos, A. (1999): "Some empirical evidences on learning difficulties about testing hypotheses". *Proceedings of the 52nd Session of the International Statistical Institute.* Vol. 2 Tome LVIII, 201-204. The Netherlands: ISI.