



# Difficulties Greek Senior High School Students Identify in Learning and the Teaching of Statistics: The Case of Experimental and Private High Schools

Dimitris Ghinis  
 Konstantinos Korres  
 Sotiris Bersimis  
 University of Piraeus, Greece

*Journal of Statistics Education* Volume 17, Number 3 (2009), [www.amstat.org/publications/jse/v17n3/ghinis.html](http://www.amstat.org/publications/jse/v17n3/ghinis.html)

Copyright © 2009 by Dimitris Ghinis, Konstantinos Korres and Sotiris Bersimis all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

**Key Words:** Learning and Teaching of Statistics; Statistics' Syllabus; Difficulties in Learning and Teaching of Statistics.

## Abstract

The present paper examines the difficulties Greek senior high school students identify in learning Statistics and how these difficulties are related to the course's level of difficulty. Also it examines the difficulties students identify that teachers face while teaching Statistics, their suggestions for changes and how these difficulties and suggestions are related to the level of the students' satisfaction by the method of teaching. In the paper a case-study is presented, that was designed and realized at the Department of Statistics and Insurance Sciences of the University of Piraeus. In the study 163 students from Experimental and Private High Schools participated, all attending the 3<sup>rd</sup> grade of Greek senior high school.

## 1. Introduction

[Garfield and Ahlgren \(1988\)](#) state that one of the most interesting fields in which the methodology of teaching Statistics must focus is in depth analysis of the difficulties that students face in understanding statistical concepts and the stochastic way of thinking, as well as, the factors that are responsible for this fact.

Motivated by the above mentioned work, we designed an appropriate empirical study examining the difficulties that 3<sup>rd</sup> grade Greek high school students from Experimental and Private high schools face in understanding Statistics and in this article, we present the main findings.

The importance of in depth analysis of the difficulties that students face in understanding statistical concepts was also pointed out by many other authors (e.g. see [Hawkins and Kapadia \(1984\)](#), [Garfield and Ahlgren \(1988\)](#), [Batanero et al. \(1994\)](#), [Garfield \(1995\)](#), [Jones et al. \(1997\)](#) and [Richardson \(2001\)](#)).

Lately, [Cooper and Shore \(2008\)](#) studied students' misconceptions in interpreting center and variability of data represented via histograms and stem-and-leaf plots. Histograms and stem-and-leaves plots, beyond their value in interpreting center and variability, are strongly related to the foundational statistical concept of a distribution / cumulative distribution (which is further explored on the basis of Probability theory). Distribution is probably the most crucial concept that students have to understand in order to develop, what we call, a stochastic way of thinking.

In our study, we explored many probable components of stochastic thinking, through the analysis of the students' views, as well as other potential problematic points that students identify in the teaching process. The study focuses strictly on teaching Statistics (not on teaching Probability).

Most of the research methodology for teaching statistics (in Greece as well as worldwide) focused on teachers of Mathematics and their views on the subject of Statistics. It has been found that although teachers of Mathematics regard teaching Statistics as very necessary, they are not very fond of teaching it because of a lack of in-depth statistical knowledge (e.g. [Ghinis et al. \(2005\)](#)).

The novelty of the present paper, which examines difficulties Greek senior high school students from Experimental and Private high schools identify in learning and the teaching of Statistics, is the point of view. Specifically, the present work seeks to identify these difficulties from the students' side.

Moreover, this paper presents the main findings of the analysis of the students' answers to a questionnaire that focuses in two major axes: (a) the difficulties students identify in learning Statistics and how those difficulties are related to the course's level of difficulty and (b) the difficulties students identify that teachers face while teaching Statistics, their suggestions for changes in teaching and how those difficulties and suggestions are related to the level of the students' satisfaction with the method of teaching.

For a preliminary analysis on the views of high school students regarding difficulties in the subject of Statistics see [Ghinis and Korres \(2005\)](#).

The paper is organized as follows: In Section 2, information is presented regarding Greece's Statistics curriculum and the Statistics' Syllabus in High School. In Section 3, we give some framework details, as well as, the corresponding questionnaire and the statistical methods that were used. In Section 4, we present how the students' questionnaires were analyzed to see firstly, the difficulties they face, as well as, the strengths and weaknesses of the teaching process. Finally, in Section 5 we interpret the results of the analysis giving some brief concluding remarks.

## 2. Statistics' Syllabus in Greece

In this section, we give details about the Statistics curriculum in Greek Education in general as well as the Statistics syllabus in Secondary Education (High

school).

## 2.1. Curriculum concerning Statistics

Greek education is structured in two levels, the six-year primary education and the six-year secondary education, which is further divided into two three-year levels, Junior High School (Gymnasium) and High School (Lyceum). Students are admitted to Higher Education after they pass the formal state examinations and after their grades in some school subjects during their last year in high school are considered.

According to the detailed curriculum of Greek education, students are taught the basic concepts of Statistics and Probability at the 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> grade of primary education, the 2<sup>nd</sup> and 3<sup>rd</sup> grade of Junior High School and the 3<sup>rd</sup> grade of High School as part of the Mathematics' Curriculum. However, in practice the teaching of Statistics and Probability is almost always insufficient due to the extensive content of the yearly Mathematics' Syllabus and because teachers consider other subjects of Mathematics more important and useful for students' further knowledge. An exception to this is the teaching of Statistics at the 3<sup>rd</sup> grade of Greek High School (Lyceum).

Statistics has been taught at the 3<sup>rd</sup> grade of Senior High School since the academic year 1999–2000, as a part of the subject "General Mathematics", which is taught for two hours weekly. In this subject's framework, students are taught mainly material in Descriptive Statistics and the Theory of Probability. The content of this subject is assessed during the formal State Examinations and is weighted heavily for students who are interested in studying Economics.

Students from all three main programs of study available at Greek Lyceum (Theoretical, Practical and Technologic Direction) can take the program in Economic Studies, with subjects "General Mathematics" and "Economic Theory" given high weight at the formal State Examinations. This is the reason why examining the students' difficulties at this grade is very important. This is also the reason why the teaching of Statistics at the 3<sup>rd</sup> grade of Senior High School is given adequate coverage, opposed to the teaching of Statistics in Junior High School.

## 2.2. Statistics' Syllabus in High School

In Statistics, students are taught how to process statistical data and interpret critically statistical conclusions. The Syllabus includes the following subjects ([Pedagogical Institute of Greece, 2007](#)):

1. Basic concepts: The students are taught, through appropriate examples, basic statistical concepts such as population, variables (quantitative and qualitative), census and sample.
2. Presentation of Statistical data: The students are taught about frequency distributions and their graphical representations. Topics also include relative frequencies, cumulative frequencies and cumulative relative frequencies and how data from continuous variables can be grouped in class intervals.
3. Location measures and measures of variation: The students are taught how to compute the arithmetic mean, the weighted mean, the median, the quartiles and the mode (location measures) and the range, the interquartile range, the variance, the standard deviation and the coefficient of variation (measures of variation) of discrete and continuous variables.
4. The synthesis of basic concepts, graphical techniques and descriptive measures for solving real life problems and / or deciding which is the best solution, by using the empirical distribution of data and its properties (location, dispersion, symmetry, kurtosis etc).

Even though the Statistics' Syllabus supports the use of computers and / or calculators, they are rarely used or not used at all, since students who participate in the "General Mathematics" formal State Examinations are not allowed to use any kind of accommodating tools during their examinations.

In the Theory of Probability, through appropriate activities, students are taught how to understand the meaning of probability and how to use it in solving problems. The correspondence between the language of sets and the language of events is pointed out and is used when students are taught how to deal with problems. The Syllabus includes the following subjects ([Pedagogical Institute of Greece, 2007](#)):

1. Sample space – Events: The students are taught the concepts of random experiments, sample space and events. Since events are regarded as subsets of the sample space,  $\Omega$ , students are taught how to perform operations between sets and the corresponding operations between events.
2. The concept of probability: The students are taught the two basic definitions of probability, the classic definition by Laplace and the axiomatic definition by Kolmogorov. Also they are taught about rules of the calculus of probability and how to use those rules when dealing with problems of calculating probabilities of events.

## 3. Framework of the study

In this section, we refer to the way the study was conducted, the sample and the variables we set for exploration. The case-study was conducted during the second semester of the academic year 2004–2005 (March, 2005). The collection of data was carried out by undergraduate students of the Department of Statistics and Insurance Sciences, attending the course "Didactics of Statistics" (under the authors' supervision), through questionnaires filled in by a sample of senior high school students. We mention here that the Statistics' Syllabus we presented above has not changed since the school year 2004–05 that the case-study was conducted.

Additionally, as already stated, the study focused strictly on teaching Statistics (not on teaching Probability); this was the reason why the questionnaires were given to the students in March, at a time point when they hadn't been taught the material on Probability. Studying students' behavior and views in learning and teaching Probability or furthermore comparing Statistics to Probability, even though they are very interesting subjects, are not goals of the present study.

### 3.1. The Sample

The sample of the study was 163 students, 64 boys and 99 girls ([Figure 1](#)), from three different High Schools: Ionnidios School of Piraeus (sample of 90 students), Evangeliki School of Nea Smirni (sample of 38 students) and Hellenic–French School Jeanne D' Arc of Piraeus (sample of 35 students). Ionnidios School of Piraeus and Evangeliki School of Nea Smirni are "Experimental" high schools while the Hellenic–French School Jeanne D' Arc of Piraeus is a "Private" high school. The characterization "Experimental" describes a category of schools, in which, innovations are firstly introduced by the Greek State, before they are widely applied to the whole country's schools. In Experimental schools the choice of teachers is made using a certain qualification procedure and the students have in general quite good performance. "Private" schools have similar characteristics to "Experimental" schools, in the sense that the teachers are carefully selected while the students also have good performance.

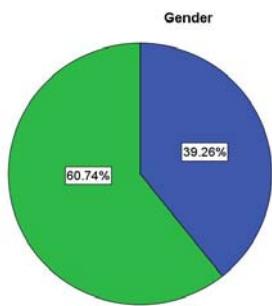
We selected the particular population of senior high schools for our study mainly due to the following reasons: (a) "Experimental" and "Private" senior high

schools are schools where changes in curriculums are introduced in advance or directly and these schools are evaluated regarding whether they are applying curriculums accurately. That evaluation ensures the homogeneity of the material taught and the methods used for all students that participated in the study. (b) Moreover, the selection of these types of schools was made by taking into account the special character of the study. The selection was based on the grounds of exploring the difficulties the students find in Statistics mainly due to the nature of the subject and the competences and skills it involves, thus, we had to eliminate factors such as difficulties due to insufficient studying or lack of interest for the course. These kinds of schools have students that generally accomplish good or excellent results in the formal state examinations, results that are above the mean results of randomly selected schools. Additionally, the teachers working in these schools are selected, as already stated, through a certain qualification procedure, so that in the study of the difficulties of the students or the difficulties that the students recognize in their teachers, factors as the inadequacy of the teachers in their cognitive or didactic skills are minimized.

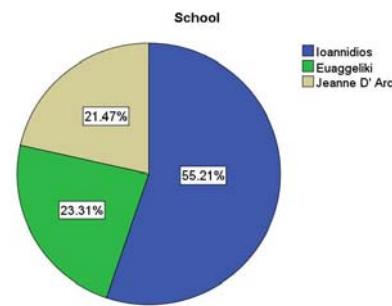
In Greece there are 18 Experimental senior high schools, 9 of which are located in the prefecture of Attica as well as 91 private senior high schools, 53 of which are located in the prefecture of Attica. The population of experimental senior high schools and private senior high schools in the prefecture of Attica includes 62 senior high schools. We selected randomly 3 of these high schools, involving the total of their students in the study. The total number of students in these 3 schools is 5% of the total student population in private and experimental high schools in Attica and 3% in all of Greece.

We should note that our results are statistically valid for the population from which we selected the specific sample, in particular the Experimental and Private senior high schools and cannot be generalized to the total of Greek senior high schools. Hence, we selected the particular population of senior high schools for specific reasons analyzed above. However, we believe that our results will not be significantly modified in the case of the total of Greek senior high schools, but that remains to be investigated with further research.

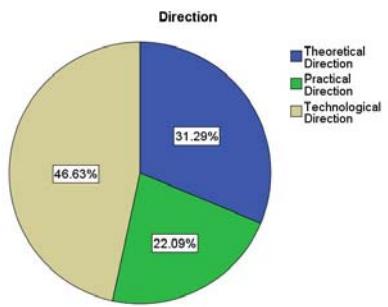
The students' distribution according to school is given in [Figure 2](#), while the students' distribution according to the direction of their studies is given in [Figure 3](#). In [Figure 4](#), we give the number of the students in the three different schools according to the direction of their studies.



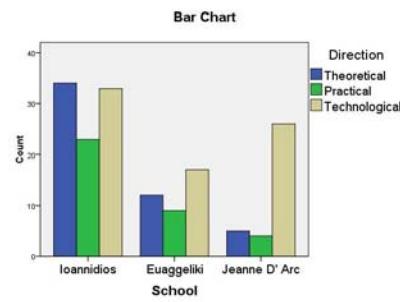
**Figure 1: Students' distribution according to gender**



**Figure 2: Students' distribution according to school**



**Figure 3: Students' distribution according to direction of studies**



**Figure 4: Students' distribution according to school and direction**

### 3.2. The Questionnaire

The questionnaire that was given to students includes two groups of statements. In each statement, students responded "Yes", when they agreed with the statement or "No" when they disagreed with the statement. We used statements via which students identify difficulties and make suggestions rather than scales of evaluation, because we regard students as not being experienced enough to grade the level of each difficulty. We regard the identification of the specific difficulties as a more valid process. Since the study refers not to evaluation or assessment by the students, but to the identification of difficulties they face in learning statistics and the difficulties their teachers face in teaching statistics, the subjectivity of the students' evaluations is minimized. The only questions that involve evaluation are the ones concerning the students' overall satisfaction with their teachers and the students' overall difficulty in learning Statistics. The students answered these two questions in a 1 – 3 Likert scale that also minimizes students' subjectivity, since the extent of satisfaction and difficulty are clearly distinguished by the three categories of the scale.

The questions in the questionnaire were designed and formed after we had extensively discussed with a group of students the problems and difficulties they face and distinguish in Statistics. Additionally, in order to avoid misconceptions when the questionnaires were handed out to the students, each individual question was explained and an example was given concerning what exactly each question means.

The first group includes eight statements by which students identified the difficulties they faced while learning Statistics. The second group includes 10

statements, by which students identified the difficulties that teachers face during teaching Statistics and the suggestions students make for the improvement of teaching.

The first group includes the statements (each statement is followed by a short explanation):

1. Difficulty in understanding basic statistical concepts (with this particular question we examine the degree of difficulty regarding basic statistical concepts, such as population, sample, measures of variation, distribution, etc.)
2. Difficulty in understanding assumptions and conclusions in statistical problems (with this particular question we examine if the students are capable of understanding the context of a statistical problem so as to distinguish which are the assumptions and which are the conclusions of the problem.)
3. Difficulty in designing a method of solution (with this particular question we examine whether the students, after having understood the assumptions and the conclusions of the problem, are able to decide on and develop a strategy for solving the problem, which includes the choice of the appropriate statistical methods.)
4. Difficulty in applying the appropriate statistical methodology and formulas (with this particular question we examine whether the students are able to apply the appropriate methods and formulas for solving the statistical problems.)
5. Difficulty in checking the validity of the method of solution (with this particular question we examine the ability of the students to re-evaluate their choice regarding the methodology of the problem's solution and its application.)
6. Difficulty in applying known statistical methodology in real life situations and problems (with this particular question we examine the ability of the students to model problems someone faces in real situations of everyday life. Modeling of problems is the gradual elimination of the real situation attributes that transforms the objective problem into a mathematical one.)
7. Difficulty in interpreting statistical results (with this particular question we examine the ability of the students to interpret charts, descriptive measures or even comparing distributions using their properties.)
8. Difficulty in performing mathematical operations for obtaining results (this particular difficulty is examined given the fact that in the Greece's educational system the use of calculators has not been established and particularly during the examination of Statistics for succeeding at the state examinations for entering higher education, the use of calculators is not permitted.)

As already mentioned, the second group includes 10 statements, by which students identified the difficulties that teachers face during teaching Statistics and the suggestions students make for the improvement of teaching. This is not an easy task to perform, since in order to succeed we had to eliminate the subjectivity and / or the probable immaturity of the students' views. The concerns mentioned above were overcome by the fact that the teacher that taught Statistics was the regular and the same teacher that had taught the students the first part on Mathematics in the course, thus the students were comparing the same teacher in different subjects. In that way, students were able to evaluate teachers in a "normalized" way, since even though they could not "measure" teachers' difficulties, they could assess the relative difference in the teachers' performance when teaching Mathematics and when teaching Statistics. Thus, if students believed that their teacher had more difficulty teaching the statistics material as compared with teaching the mathematics material, this would be identified as a perceived difficulty in teaching statistics. Additionally, in order to isolate the teacher's effect as a determining factor of difficulties in learning statistics, as perceived by the students, in each individual question the tasks the teacher aimed to achieve were further explained to the students, in order for the students to express their views concerning whether the teacher achieved those tasks and the difficulties he / she faced in achieving those tasks.

Thus, the second group includes the following statements:

1. Teacher's difficulty in presenting basic statistical concepts (with this particular question we examine if the students identify difficulty on behalf of the teacher in presenting basic statistical concepts taking into consideration the fact that teaching the course in Statistics requires a particular method and conceptual approach different from the one used in Mathematics.)
2. Teacher's difficulty in presenting and organizing data in the classroom (with this particular question we examine if the students identify difficulty on behalf of the teacher in presenting and organizing data in the classroom taking into consideration that the teachers teaching Statistics adapt their teaching to limited tools or teaching aids since the students attending the formal State Examinations organize their data using only the free spaces on their tests.)
3. Teacher's difficulty in analyzing data (with this particular question we examine if the students identify difficulty on behalf of the teacher in analyzing data in the classroom taking into consideration that the teachers teaching Statistics adapt their teaching to limited tools since the students attending the formal State Examinations process the problems' data using only their pens and the free spaces on their tests.)
4. Teacher's difficulty in presenting the interpretation of statistical results (with this particular question we examine if the students identify difficulty on behalf of the teacher in presenting the interpretation of diagrams, location and variation measures and other statistical results. For example difficulty in interpreting the diagram of the cumulative distribution function and equivalent difficulty in communicating to the students the particular interest the diagram has.)
5. Teacher's difficulty in teaching using real life projects (with this particular question we examine if the students identify difficulty on behalf of the teacher in teaching Statistics using projects).

The questions that follow consist of the students' proposals that aim to improve the teaching method for the course in Statistics. The five options given to the students are:

6. Suggestion for the teachers to use slides or other teaching aids in the classroom
7. Suggestion for the teachers to be better informed on the Science of Statistics
8. Suggestion for the teachers to provide more and / or more difficult exercises
9. Suggestion for the teachers to use computers in teaching
10. Suggestion for the teachers to better prepare students for the State Examinations

Also, there are two additional questions, which students graded on a 1 to 3 Likert-like scale, the level of difficulty of the lesson and their satisfaction with the method of teaching. Specifically:

1. "Grade your overall satisfaction with your teacher while he is teaching statistics?" With this question we tried to quantify the general perception of overall satisfaction of the students with their teacher (perception of teacher's overall performance when teaching statistics). This is not an easy task to perform, since students' perceptions are not always accurate. Nevertheless, as already stated, teachers that taught Statistics were the regular and the same teachers that had taught the students the first part on Mathematics in the course. This fact gives to students a standard for comparison of the same teacher in different subjects. In that way, we have a "normalized" index, in the sense that student does not give his perception on the exact teacher's ability but gives his perception on the relative difference of teacher's ability (comparing to teacher's ability when teaching Mathematics).
2. "Grade your overall difficulty in learning statistics?" With this question we tried to quantify the level of difficulty in understanding and handling subjects of statistics (overall perception of their difficulty in Statistics).

In the forthcoming section, we will analyze students' responses in order to answer the main questions of the study as posed in the introduction. Here we must

remind the reader that the students expressed their views strictly regarding the material on Statistics.

### 3.3. Statistical methods used

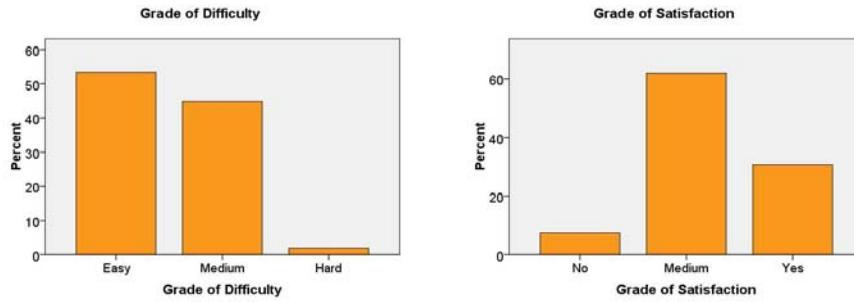
In general, the multi-methodological approach was used ([Cohen and Manion \(2000\)](#)). Methods from Descriptive and Inferential Statistics were used ([Tsantas, et al. \(1999\)](#) in combination with Multi-dimensional Analysis, in particular Multiple Correspondence Analysis ([Behrakis \(1999\)](#) and [Greenacre \(1984\)](#), [Benzécri \(1973, 1980\)](#), [Lebart et al. \(1977\)](#), [Lebart et al. \(1995\)](#) and [Dunteman \(1989\)](#)).

## 4. Findings of the study

### 4.1. Overall Evaluation

By observing [Figure 5](#), we can see that the majority of students (98.2%) consider the subject of Statistics of low difficulty (53.4%) or of medium difficulty (44.8%). Also, by observing [Figure 6](#) we can see that the majority of the students are relatively satisfied with their teachers.

In [Table 1](#), the difficulties that students identified in learning Statistics and difficulties students identified that teachers face in teaching Statistics are presented. In decreasing order these are the difficulty in applying known statistical methodology in real life situations and problems (36.2%), in designing a method of solution (26.4%), in checking the validity of a method of solution (24.5%), in performing the appropriate mathematical operations for obtaining results (22.7%), in applying the appropriate statistical methodology and formulas (17.8%) and smaller percentages the remaining difficulties.



**Figure 5: Students' characterization of difficulty of the lesson**

**Figure 6: Students' characterization of satisfaction by their teachers**

Additionally, the analysis revealed that students who characterized the subject as easy, identified as most difficult applying known statistical methodology in real life situations and problems (26.4%) and performing the mathematical operations for extracting results (21.8%). The group of students who characterized the subject as of medium difficulty, identified difficulties in applying known statistical methodology in real life situations (46.6%), in designing a solving procedure (39.7%), in checking the validity of a method of solution (30.1%) and in performing the operations for obtaining results (23.3%).

In the forthcoming analysis we explore whether there are any clusters of statements (variables) or students (objects) involving a common meaning and if there are whether this can be interpreted. In order to accomplish this, we applied the technique of Multiple Correspondence Analysis to our data set. The purpose of Multiple Correspondence Analysis is to find transformations of statements that are optimal, in the sense that the categories of variables are separated from each other as much as possible. By this, we deduce that students identifying the same difficulties (or making the same suggestions) will be plotted close to each other, while students identifying different difficulties (or making different suggestions) will be plotted as far apart as possible.

**Table 1: Students' Identifying Difficulties in learning and teaching Statistics**

| Statement  | Frequency | %     |
|--|-----------|-------|
| <b>Difficulties in learning Statistics</b>   |           |       |
| 1. Difficulty in understanding basic statistical concepts                                    | 6         | 3.7%  |
| 2. Difficulty in understanding assumptions and conclusions in statistical problems           | 22        | 13.5% |
| 3. Difficulty in designing a solving procedure   | 43        | 26.4% |
| 4. Difficulty in applying the appropriate statistical methodology and formulas               | 29        | 17.8% |
| 5. Difficulty in checking the validity of the solving procedure                              | 40        | 24.5% |
| 6. Difficulty in applying known statistical methodology in real life situations and problems | 59        | 36.2% |
| 7. Difficulty in interpreting statistical results  | 6         | 3.7%  |
| 8. Difficulty in performing mathematical operations for extracting results                   | 37        | 22.7% |
| <b>Difficulties and Suggestions in teaching Statistics</b>                                   |           |       |
| 1. Teacher's difficulty in presenting basic statistical concepts                             | 23        | 14.1% |
| 2. Teacher's difficulty in presenting and organizing data in the classroom                   | 47        | 28.8% |
| 3. Teacher's difficulty in analyzing data  | 33        | 20.2% |
| 4. Teacher's difficulty in presenting the interpretation of statistical results              | 23        | 14.1% |
| 5. Teacher's difficulty in teaching using real life projects                                 | 23        | 14.1% |
| 6. Suggestion for the teachers to use slides or other teaching aids in the classroom         | 25        | 15.3% |
| 7. Suggestion for the teachers to be better informed on the Science of Statistics            | 25        | 15.3% |
| 8. Suggestion for the teachers to provide more and / or more difficult exercises             | 32        | 19.6% |
| 9. Suggestion for the teachers to use computers in teaching                                  | 37        | 22.7% |
| 10. Suggestion for the teachers to better prepare students for the State Examinations        | 32        | 19.6% |

#### 4.2. Analyzing First Group's variables (Students' related variables)

The Multiple Correspondence Analysis for the variables of the first group (related to students' difficulties) showed that more than half of the variance in the data was accounted for by a three dimensional solution. The term dimension corresponds to a latent variable or in other words to a transformation of statements. Specifically, 24.1% of the dispersion was interpreted by the first dimension and 14.9% by the second. The three dimensions together provide an interpretation in terms of distances. If one of the original statements discriminates well, the students (objects) will be close to the categories to which they belong. Ideally, students in the same category will be close to each other. This means that they should have similar scores.

After examining the model summary, we looked at the students' scores. Using various plots we may determine the homogeneity of the data set. The data set is quite homogeneous (one outlier was found which was removed). After looking at the object scores, we looked at the discrimination measures (Figure 7). For each variable (statement) a discrimination measure is computed for each latent variable (dimension) which also represents the variance of the quantified variable in that dimension. It has a maximum value of 1. Large discrimination measures correspond to a large spread among the categories of the variable and consequently, indicate a high degree of discrimination between the categories of the original statement along that latent variable. The average of the discrimination measures for each dimension (latent variable) equals the percentage of variance accounted for that dimension. By this, it is obvious that the dimensions are ordered according to average discrimination.

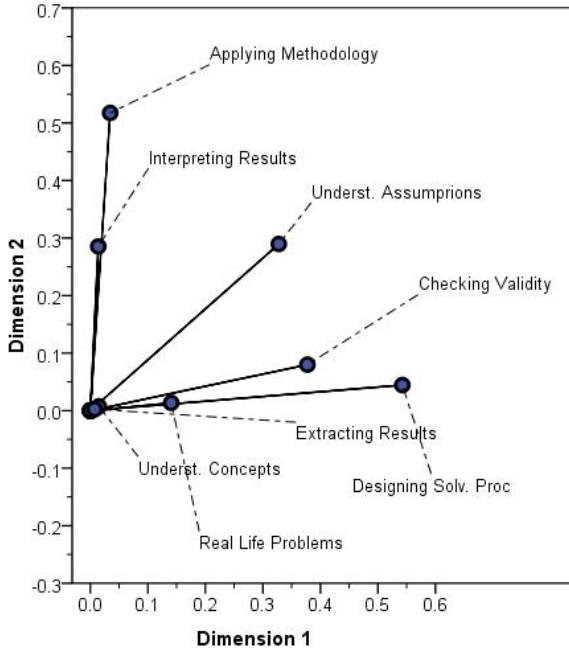


Figure 7: Discrimination measures (Students' related variables)

The discrimination measures plot showed that the first dimension is related to the following variables:

1. Difficulty in designing a method of solution (weight: 0.542)
2. Difficulty in checking the validity of a method of solution (weight: 0.377)
3. Difficulty in understanding assumptions and conclusions in statistical problems (weight: 0.328)

These variables have large discrimination measures on the first dimension and small discrimination measures on the second dimension. Clearly the first dimension is related to students' difficulty in stochastic thinking.

Regarding the second dimension, we observe that it is strongly related to students' difficulty in applying the appropriate mathematical procedures and interpreting results:

1. Difficulty in applying the appropriate statistical methodology and formulas (weight: 0.517)
2. Difficulty in interpreting statistical results (weight: 0.285)

Finally, as far as the third dimension, we observe that it is strongly related to students' difficulty in solving real life statistical problems:

1. Difficulty in performing mathematical operations for obtaining results (weight: 0.468)
2. Difficulty in applying known statistical methodology in real life situations and problems (weight: 0.246)

Here we must note that the variable (statement) "Difficulty in understanding basic statistical concepts" is not related to any of the three dimensions. This is because only a few students stated that they have difficulty in understanding basic statistical concepts.

The conclusion of this section is that according to students, their difficulties may be ranked as: firstly, difficulties in stochastic thinking secondly, difficulties in applying the appropriate mathematical procedures and interpreting results and finally, difficulties in solving real life statistical problems.

#### 4.3. Analyzing Second Group's variables (Teachers' related variables)

The Multiple Correspondence Analysis for the variables of the second group (related to teachers' difficulties as identified by the students) indicated that more than half of the variance in the data was accounted for by a three dimensional solution.

Furthermore, 35.5% of the variance was interpreted by the first dimension, 11.7% by the second and 10.3% by the third. After examining the model summary, we looked at the object scores in order to identify the existence of possible objects clusters and / or some outliers. The data were found to be quite homogeneous and no outliers were identified.

Then, we looked at the discrimination measures (Figure 8). The discrimination measures plot showed that the first dimension is related to the variables:

1. Suggestion for the teachers to better prepare students for the State Examinations (weight: 0.678).
2. Suggestion for the teachers to be better informed on the Science of Statistics (weight: 0.638).
3. Suggestion for the teachers to use slides or other teaching aids in the classroom (weight: 0.584).
4. Suggestion for the teachers to provide more and / or more difficult exercises (weight: 0.552).
5. Teacher's difficulty in teaching using real life projects (weight: 0.514).
6. Suggestion for the teachers to use computers in teaching (weight: 0.463).

These variables have large discrimination measures on the first dimension and small discrimination measures on the second dimension. Thus, for these variables, the categories are spread far apart along the first dimension only.

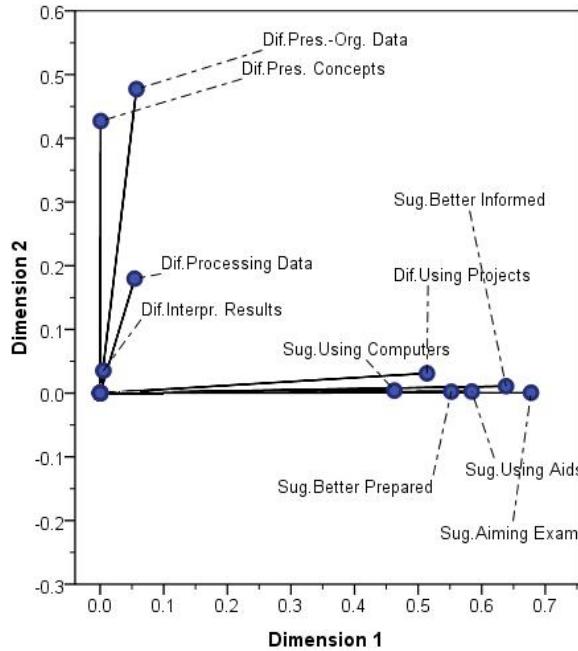


Figure 8: Discrimination measures (Teachers' related variables)

Regarding the second dimension we notice that this dimension is related to the following variables – statements:

1. Teacher's difficulty in presenting and organizing data in the classroom (weight: 0.477).
2. Teacher's difficulty in presenting basic statistical concepts (figure's label: "Dif. Pres. Concepts", weight: 0.427).

The third dimension is related only to the variable "Teacher's difficulty in presenting the interpretation of statistical results" with weight equal to 0.748.

#### 4.4. Analyzing the relation between the Latent variables and the Characteristics

In this subsection, we explore the possible interactions of the latent variables (dimensions) that were extracted in the previous two sub-sections with gender, direction, school, as well as the total grade of difficulty of statistics and total satisfaction index of the students by their teachers.

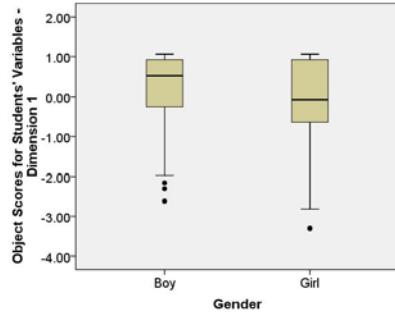
Because of the fact that the 6 latent variables (dimensions) were not normally distributed, we used appropriate non parametric tests in order to check if they varied with changes in the aforementioned factors (gender, direction, school, total grade of difficulty of statistics, total satisfaction index of the students by their teachers). The summarized results are given in [Table 2](#). In each cell we give the corresponding p-value of the appropriate hypothesis test. As we can see in [Table 2](#) the first dimension which is related to students' difficulty in stochastic thinking, depends on the gender of the student, the direction which has been chosen by the student, as well as the difficulty index.

Table 2: Interactions of the dimensions with gender, direction, school, the total grade of difficulty of statistics and total satisfaction index of the students by their teachers

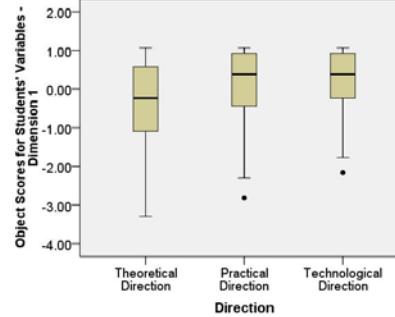
| Dimensions / Factors |             | Factor    |           |        |                  |                    |
|----------------------|-------------|-----------|-----------|--------|------------------|--------------------|
|                      |             | Gender    | Direction | School | Difficulty Index | Satisfaction Index |
| Students             | Dimension 1 | 0.029 (*) | 0.022 (*) | 0.201  | 0.001 (*)        | 0.707              |
|                      | Dimension 2 | 0.786     | 0.616     | 0.098  | 0.177            | 0.804              |
|                      | Dimension 3 | 0.544     | 0.524     | 0.567  | 0.234            | 0.465              |
| Teachers             | Dimension 1 | 0.132     | 0.565     | 0.051  | 0.588            | 0.250              |
|                      | Dimension 2 | 0.316     | 0.210     | 0.543  | 0.040 (*)        | 0.399              |
|                      | Dimension 3 | 0.784     | 0.574     | 0.756  | 0.324            | 0.635              |

Clearly, from [Figures 9a, 9b, and 9c](#) we can see that difficulty in stochastic thinking has lower objective scores (corresponding to greater difficulty) for female

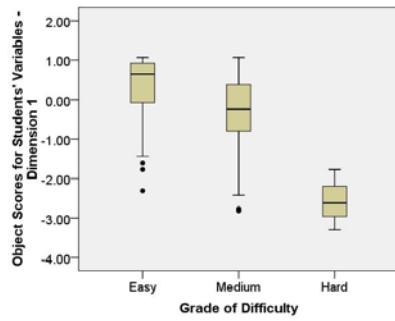
students, for students who have chosen the theoretical direction and finally for students that characterize Statistics as a difficult school subject. Even the fact, that statistical significant differences appeared, we may comment that the differences as displayed in [Figures 9a](#) and [9b](#) regarding the association of the students' difficulty in stochastic thinking to gender and direction are not dramatic (for example the gap between boys and girls is not too large), while additionally a lot of variation is present (student to student deviations). Finally, the other two students' dimensions did not give evidence of such dependences. Here, we should note that in [Figures 9a](#), [9b](#), [9c](#) and [9d](#), the scale corresponds to standardized scores (dimensions' scores), as well as, that smaller values correspond to higher percentages. Moreover, from [Table 2](#) and [Figure 9d](#), we can see that students' difficulty index on Statistics depends on the second teachers' dimension, which is related mainly to teachers' difficulties (as perceived by students) in presenting and organizing data in the classroom.



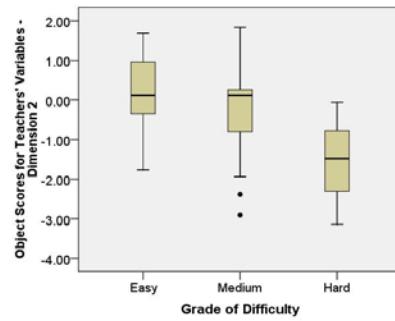
**Figure 9.a:** Students' difficulty in stochastic thinking and gender



**Figure 9.b:** Students' difficulty in stochastic thinking and direction



**Figure 9.c:** Students' difficulty in stochastic thinking and difficulty index



**Figure 9.d:** Teachers' difficulty in presenting and organizing data in the classroom and difficulty index

As already pointed out, even though dramatic differences are not observed, female students appear to find stochastic thinking a bit more difficult than male students. It is true that the magnitude of the difference between males and females (see [Figure 9a](#)) is not large. Thus, we believe that the question whether a difference of that magnitude is meaningful cannot be easily answered and for that reason it has to be further investigated.

However if we remove the outliers which can be easily observed in the [Figure 9a](#), the difference in the level of understanding of stochastic thinking among the two genders is getting larger. Another interesting point is that either removing or not removing these outliers, females' score values are uniformly smaller than males' score values as much in Technological Direction, as in Practical Direction and in Theoretical Direction. This means that males have greater scores than females independently of the direction of the studies. Finally, the same attitude is observed in the case of the school. This means that males have greater scores than females independently of the school.

In conclusion, in this sub-section we may characterize the findings in [Figures 9c](#) and [9d](#), as expected, as the more difficult a student finds the material, the more difficult he/she finds stochastic thinking and the more he/she perceives difficulty in his/her teachers. Hence, the difficulties that a student encounters, are strongly related to the difficulties they experience with the material, as well as, to their perceptions of the difficulties the teacher has in presenting the material, specifically, the difficulties students have with the material are positively associated with their perceptions of the difficulty of the material and their perception of the difficulties the teacher has in presenting the material.

## 5. Conclusions

### 5.1. Characterizing the Difficulties of Statistics and Tracing Students' Types of Difficulty

By using classical statistical analysis, it was revealed that in general students consider the subject of Statistics as easy or of medium difficulty and identify the difficulties they face mainly in the procedure of solving problems, in applying known statistical methodology to unfamiliar, real-life situations and problems and in performing the mathematical operations for extracting results. They do not identify difficulties in understanding basic statistical concepts nor in interpreting results.

As the Multiple Correspondence Analysis revealed, the first axis is related to students' difficulty in stochastic thinking while the second axis is strongly

related to students' difficulty in applying the appropriate mathematical procedures to obtain and interpret results.

Another interesting point is that the first students' dimension which is related to students' difficulty in stochastic thinking depends on the gender of students, the direction which students have selected, as well as the difficulty index. Specifically, difficulty in stochastic thinking has been identified at a greater level by female students, by students who have chosen the theoretical direction and finally for students that characterize Statistics as a difficult school subject.

## 5.2. Characterization of Satisfaction with the method of Teaching

The majority of the students have a positive attitude towards the teachers' teaching methods. However, there is a percentage of 25.8% who are not satisfied. Students identify the difficulties that teachers face mainly in presenting and organizing data in the classroom and in processing of data. Their most important suggestions as to the improvement of teaching include the use of computers, the orientation of the content of Statistics towards their better preparation for the State Examinations, the use of slides and other teaching aids and the teachers being better informed about Statistics.

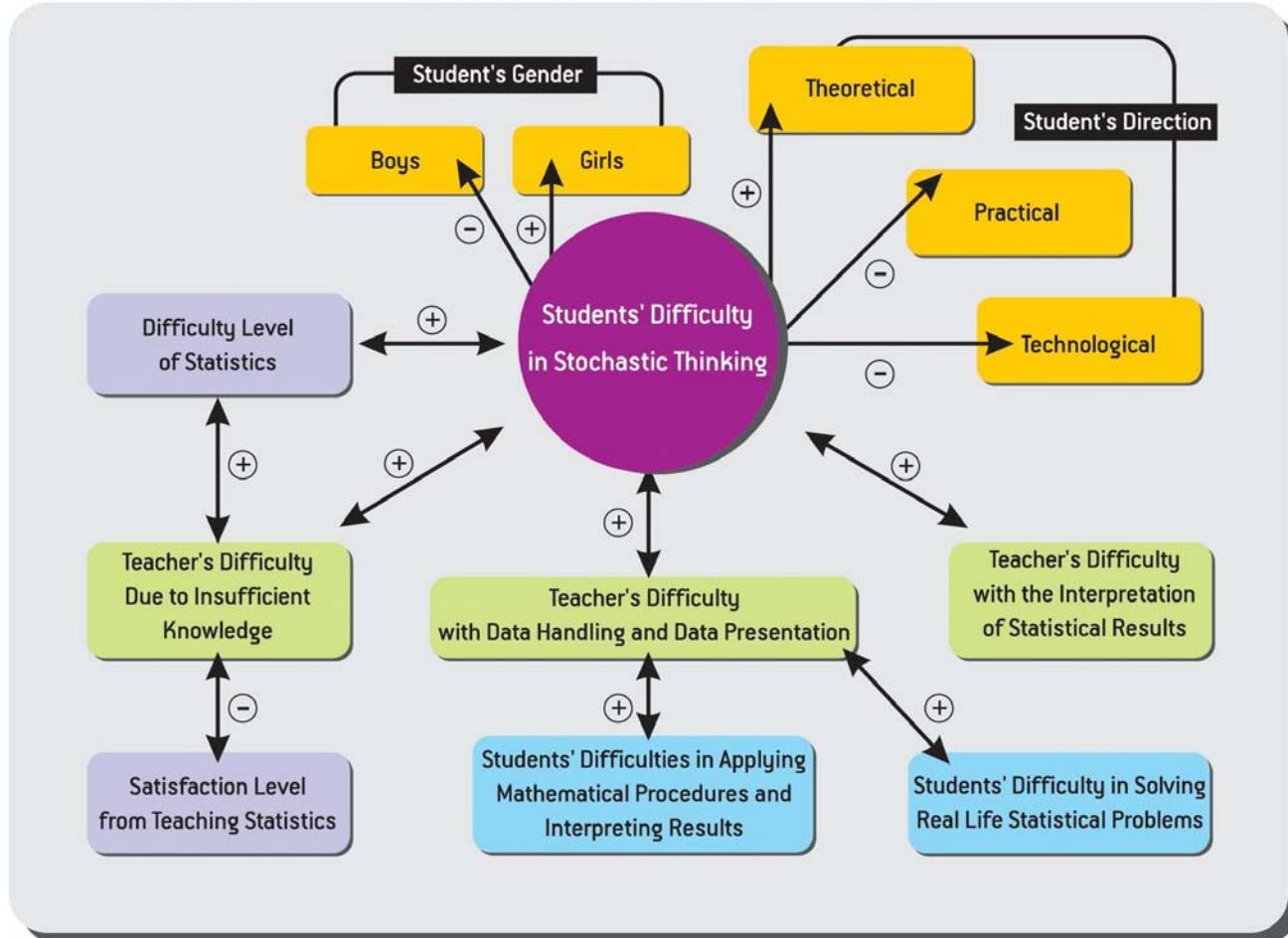
The same conclusions have also been obtained by applying Multiple Correspondence Analysis to our data. The first axis is related to the students' suggestions for changes in the method of teaching and is not related to the satisfaction with the method of teaching. This reveals a positive and energetic attitude on the part of the students. The second axis is related to students' satisfaction with the method of teaching. As was clear, students' satisfaction is negatively associated with perceptions of teachers' difficulties with analyzing data and interpreting results.

## 5.3. Framework of the Relations Revealed by the Analysis

In [Figure 10](#), the full framework of the relations revealed by the analysis is given. Based on this network we may give some suggestions for the improvement of the teaching of Statistics. As we can see, the most significant problem students identify is related to stochastic thinking. Furthermore, students' "Difficulty in Stochastic Thinking" is directly related to students' characterization of the level of difficulty of the lesson, as well as to the level of dissatisfaction with the teaching.

Even though dramatic differences are not observed, female students appear to find stochastic thinking a bit more difficult than male students. However the question whether a difference of that magnitude is meaningful cannot be easily answered and for that reason we believe it has to be investigated by further research.

Moreover, students who are in the theoretical direction find stochastic thinking more difficult. This result may be attributed to the mathematical and science background of the students in theoretical direction. Specifically, students in practical and technologic direction attend, both in 2nd and 3rd grade of Greek Lyceum, direction courses in Mathematics and Science, with more advanced content than Core Mathematics and Science courses that all students (including students from theoretical direction) in 2nd and 3rd grade attend. From the above discussion, it is clear that the students who are more familiar to Mathematics and Science can develop stochastic way of thinking more easily.



**Figure 10: Framework of the relations revealed by the analysis**

Thus, one significant intervention must focus on supporting students' stochastic thinking. This can be done by incorporating activities into teaching, such as active discovery learning activities, games, small scale research tasks, or directed projects with real life problems.

Since the more difficult students find the material, the more difficult they find stochastic thinking and the more they perceive difficulties in their teachers, an intervention could be to emphasize on clearer explanations, in order for these students to understand the material. Additionally, the use of discovery learning activities could help these students change their point of view for the course, adopting a more active attitude, an attitude of experimentation for the course's content and maybe change their perception of the grade of difficulty for the course.

Another intervention could be the early identification of the students who find the material of the course difficult, so additional support is provided to them (supporting lessons and/or extra material that these students can work with in order to strengthen their weaknesses). In this context, technology can provide essential support in identifying weak students using intelligent tutoring platforms ([Varlamis and Bersimis, 2006](#)).

Also technology can provide essential support and aid, with asynchronous communication systems as platforms of asynchronous teleconferencing, via which students can work with extra material that they can download from their home and send back to their teacher after they have finished ([Korres, 2007](#)). Also the use of Tutoring Systems or Intelligent Tutoring Systems, even though they are rare in the Greek market, could also provide essential aid in supporting students identified to find the material taught difficult.

Additionally, due to the fact that students identify that their teachers have difficulties due to insufficient knowledge and this is directly related to students' "Difficulty in Stochastic Thinking", another intervention must focus on supporting teachers' training, by getting them involved in further educational programs.

In Greece the case of elementary education teachers' training in teaching Statistics is described in [Ghini et al. \(2005\)](#). The existing teacher training programs for teachers of Mathematics, who also teach Statistics in Secondary Greek education, are focused mainly on teaching Mathematics and/or teaching using Computers. The teaching of Statistics in initial or continuing teacher training is barely covered. Mathematics' teachers in order to be trained in teaching Statistics more effectively, have to attend a postgraduate program on "Didactics of Mathematics" or "Studies on Education", or find material on the subject searching themselves in libraries or the Internet. Therefore the design and implementation of training programs on the teaching of Statistics for both pre service and in service teachers is urgent. Programs of training in teaching Statistics for Secondary Education teachers could have the form of workshops in which after a brief discussion of some theoretical concepts, teachers would work on certain real life projects (such as estimating the relation between the body mass index and the total cholesterol of some patients or the relation of the number of traffic law violations and the number of cars accidents etc).

---

## Acknowledgments

The authors would like to thank Bill Notz, Editor of the Journal of Statistics Education, for his valuable comments which substantively improved the final version of the paper.

---

## References

Batanero C., Godino J.D., Vallecillos A., Green D.R. and Holmes P. (1994). "Errors and difficulties in understanding elementary statistical concepts". *International Journal of Mathematics Education in Science and Technology*, **25** (4), 527–547.

Behrakis Th. (1999). *Multidimensional Data Analysis: Methods and Applications (in Greek)*. Livanis Publications, Athens.

Benzécri J. P. (1973). *L' Analyse des Données, Tome 2: L'Analyse des Correspondances*. Dunod, Paris.

Benzécri J. P. (1980). *Pratique de l' Analyse des Données. Abrege theoreque – Cas mod'ele*. Dunod, Paris.

Cohen L. and Manion L. (2000). *Research Methods in Education (4<sup>th</sup> Edition)*, Routledge, London and New York.

Cooper L.L. and Shore F.S. (2008). "Students' Misconceptions in Interpreting Center and Variability of Data Represented via Histograms and Stem-and-leaf Plots". *Journal of Statistics Education*, **16** (2).

Dunteman G. H. (1989). *Principal Components Analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-069, Beverly Hills.

Garfield J. and Ahlgren A. (1988). "Difficulties in learning basic concepts in Probability and Statistics: Implications for Research". *Journal for Research in Mathematics Education*, **19** (1), 44–63.

Garfield J. (1995). "How students learn statistics". *International Statistical Review*, **63** (1), 25–34.

Ghini D. and Korres K. (2005). *Views of 3<sup>rd</sup> Grade High School students regarding difficulties in the subject of Statistics (in Greek)*. Proceedings of the 22<sup>nd</sup> Hellenic Conference of Mathematics Education of the Hellenic Mathematical Society.

Ghini D., Chadjipantelis T. and Bersimis S. (2005). "Teaching statistics using directed projects in elementary school: the Greece's case". *Teaching Statistics*, **27** (1), 2–7.

Greenacre M. J. (1984). *Theory and Applications of Correspondence Analysis*. Academic Press, London.

Hawkins A.S. and Kapadia R. (1984). "Children's conceptions of probability – A psychological and pedagogical review". *Journal of Educational Studies in Mathematics*, **15** (4), 349–377.

Jones G.A., Langrall C.W., Thornton C.A. and Mogill A.T. (1997). "A framework for assessing and nurturing young childrens thinking in probability". *Educational Studies in Mathematics*, **32** (2), 101–125.

Korres K. (2007). *A Teaching Approach of Mathematics and Science courses using new technologies (in Greek)*. Doctoral Dissertation. University of Piraeus, Department of Statistics and Insurance Sciences.

Lebart L., Morineau A. and Piron M. (1995). *Statistique exploratoire multidimensionnelle*. Dunod, Paris.

Lebart L., Morineau A. and Tabard N. (1977). *Techniques de la description statistique*. Dunod, Paris.

Pedagogical Institute of Greece (2007). *Instructions for the syllabus of lessons in Gymnasium and Lyceum for the school year 2007–08 (Part B: Mathematics) (in Greek)*. Athens: Organization of Publication of Teaching Books.

Richardson A. (2001). "The world of chance". *Teaching Statistics*, **23** (2), 61–64.

Tsantas N., Moysiadis Ch., Bayiatis D. and Chadjipantelis T. (1999). *Data Analysis using statistical packages (in Greek)*. Zitis Publications, Thessaloniki.

Varlamis I. and Bersimis S. (2006). Providing shortcuts to the learning process. In *Recent Progress in Computational Sciences and Engineering*, Edited by Th. Simos & G. Maroulis, Lecture Series on Computer and Computational Sciences, Volume 7A, Brill Academic Publishers, ISBN-13: 9789004155428.

---

Dimitris Ghinis, PhD  
 Department of Statistics and Insurance Sciences  
 University of Piraeus  
 Greece  
 Email: [dghinis@unipi.gr](mailto:dghinis@unipi.gr)

Konstantinos Korres, PhD  
 Department of Statistics and Insurance Sciences

University of Piraeus  
Greece  
Email: [kostas\\_korres@yahoo.gr](mailto:kostas_korres@yahoo.gr)  
Tel.: 0030 6977 640 107

Sotiris Bersimis, PhD  
Department of Statistics and Insurance Sciences  
University of Piraeus  
Greece  
Email: [sbersim@unipi.gr](mailto:sbersim@unipi.gr)

---

[Volume 17 \(2009\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)