

STATISTICAL DATA ANALYSIS IN PSYCHOLOGY, GENERATION OF SELF-EXAMINATION QUESTIONNAIRES FOR STUDENTS: APPROACH TO THE ITEM DISTRIBUTIONS STUDY THROUGH CHARACTERISTIC CURVES

Joan Guàrdia-Olmos, Maribel Peró-Cebollero, Montserrat Freixa-Blanxart, Jaume Turbany-Oset, Amàlia Gordóvil-Merino and María Carrera-Fernández
Department of Methodology of Behavioral Sciences, University of Barcelona, Spain
jguardia@ub.edu

The aim of this contribution is to present an approach to item analysis in a self-examination questionnaire for students in Statistical Data Analysis for Psychology. We have worked with a sample of 413 students from the Faculty of Psychology from the University of Barcelona during the 2008-09 academic year. They answered three different closed questionnaires (20 to 30 items each) about general and specific statistics topics. The study of the results was made according to the item response theory. The third questionnaire (tools of inference statistics) was the best one in discrimination rates, while the second one (probability models) presents a balanced distribution of items in terms of their rates of difficulty. Finally, the first questionnaire seems to be the easiest one (descriptive statistics). The conclusions show that self-examination can be a useful system of student assessment but not for the entire topic in a course on statistics for psychology students.

INTRODUCTION

The teaching staff of the Social Sciences degrees is aware of the low academic performance of the students in this subject, academic performance referring to the fact of passing or failing the course. Specifically, it is usual to find in the different Faculties of Psychology across Spain that the percentage of students that pass the Statistics course on their first attempt is around 60% in the best case. There are several papers that try to determine the factors that influence the students' low performance in technical-methodological courses and Statistics in particular. For example, some works study the relationship of academic performance with variables of the motivational kind, personality or psychosocial factors (Castejón, Navas & Sampascual, 1996; Garrido & Rojo, 1966; Pérez-Sánchez & Castejón, 1996; Rocas, González & Touron, 1997; or García & Fumero, 1998). Other authors point out other factors related to the undergraduates' academic performance in methodological courses such as SAT grades, performance in high school and itinerary chosen during high school (García, Alvarado & Jiménez, 2000), as well as attendance and participation in in-person courses (García et al., 2000; Huberty, 2000). Gay and Ginsburg (1994) or Garfield (1994) remark how useful it is to bear in mind the students' attitudes and beliefs toward Statistics and Smith (1998), Boyle (1999), and Gardner and Hudson (1999) point out how interesting it is to carry out practical works where different statistical techniques must be applied while learning Statistics.

Given our concern for the students' low performance in the *Data Analysis in Psychology* course, our group has taken several actions throughout the years (Guàrdia, et al., 2002; 2006). The results of these works show the common dispersion effect usually found in the study of academic performance. Other more active actions implied developing material for the students in an electronic format, such as the Descriptive Statistics CD-ROM (Guillén, et al. 2001), whose efficacy in learning Descriptive Statistics is reflected in the work by Peró et al. (2004). In the last years, action in this field has focused more directly on the actual teaching, specifically, collaborative work in groups of four students has been developed in order to solve different problems proposed by the teachers throughout the course. This strategy has implied an increase in the students' academic performance, with a 70% of the students passing the course on their first attempt (Guàrdia, Freixa, Turbany and Peró, 2008 and Guàrdia, Peró, Freixa, Turbany and Gordóvil, in press 2009a) (<http://www1.ub.edu/gid-estadistica/enllacos.htm>). As a complement to all this, some proposals of self-evaluation questionnaires have been generated by using computerized platforms (Guàrdia, Peró, Freixa, Turbany and Gordóvil, in press 2009b).

At this point, it is necessary to comment that the self-evaluation processes in the students' autonomous learning is a matter not yet solved. That is why this paper intends to show a psychometric study of a self-evaluation questionnaire generated from the necessary psychological

measure in order to guarantee that the items planned are adequate. This will be carried out by applying the item response theory. The choice of the IRT as a favorable vehicle is justified not only by the very features of the analysis technique, but also by how scarcely the IRT has been used for this type of performance tests. There are precedents in the academic field (Guàrdia et al., 2009), but none addressed to self-evaluation tests or systems in the field of Statistics in university degrees.

METHOD

The sample studied comprised 447 students registered in five groups of the morning shift of the *Data Analysis in Psychology* course in the University of Barcelona's Faculty of Psychology. Out of all the students studied, 18.2% were male, the participants' mean age was 19.8 years old with a standard deviation of 3.4 years and 30.2% of the students registered in these five groups had already taken the course in previous years.

The instruments used were three multiple-choice questionnaires. Specifically, each item had four answer choices, out of which only one was correct. The first questionnaire comprised 30 items, the second questionnaire comprised 28 items, and the third questionnaire comprised 15 items. Each questionnaire corresponded to each of the blocks into which the course's syllabus had been structured. Therefore, the first questionnaire corresponded to aspects basically related to Descriptive Statistics; the second questionnaire corresponded to aspects related to the normal distribution, the sampling distribution, and interval estimation; and finally, the third questionnaire assessed aspects related to bivariate hypothesis tests.

These questionnaires were at the students' disposal one week prior to the in-person evaluation test and they could voluntarily turn them in to the teachers on the day of the test. The first questionnaire was answered by a total of 413 students; the second one, by a total of 390 students; and finally, the third questionnaire was answered by a total of 372 students. The data collection followed the required ethical rules.

DATA ANALYSIS

Data analysis for each of the questionnaires consisted of the study of the characteristic curves of the items based on the item response theory (IRT). The general three-parameter model has been studied in the three questionnaires, given that, since they are performance tests, it is important to bear in mind the possibility of randomly choosing the right answer, and each student's performance and skill. Nonetheless, we should mention that in the case of the second questionnaire, because convergence was not reached in the solution proposed initially, the two-parameter model was adjusted. The parameter estimation method in the three cases has been conditional maximum likelihood estimation, according to the following expressions:

$$\text{Three-parameter model: } p_i(\theta) = c_i + (1 - c_i) \frac{e^{D a_i (\theta - b_i)}}{1 + e^{D a_i (\theta - b_i)}}$$

Where: $p_i(\theta)$ is the probability of choosing the right answer to the item for a (θ) determined value of competence

θ are the values of the measured variable

e is the base of the natural logarithms

D is a constant

a_i is the discrimination index of the i item

b_i is the difficulty index of the i item

c_i is the value of $p_i(\theta)$ for a θ value equal to $-\infty$

RESULTS

Initially, the conditions of applying the IRT were analyzed for each of the questionnaires, thus obtaining favorable results to meet such conditions (homogeneity of the variations and independence of the distributions). Table 1 shows the results of the third questionnaire due to space and to it being the most complex field of the academic syllabus (Inferential Statistics).

Table 3. Parameter summary table of the three-parameter model for the 15 items of the third questionnaire (convergence met at cycle 36)

Item	b_0 (SE)	Factorial (SE)	Parameter a (SE)	Parameter b (SE)	Parameter c (SE)
1	1.83*(0.27)	0.95*(0.28)	3.23*(0.96)	-0.56*(0.21)	0.07*(0.03)
2	-0.56*(0.32)	0.91*(0.30)	2.27*(0.76)	0.25*(0.08)	0.06*(0.03)
3	-1.42*(0.56)	0.95*(0.38)	3.26*(1.30)	0.43*(0.04)	0.06*(0.02)
4	1.65*(0.25)	0.90*(0.30)	2.06*(0.70)	-0.80*(0.36)	0.05*(0.02)
5	0.84*(0.27)	0.91*(0.28)	2.30*(0.70)	-0.36 (0.21)	0.06*(0.03)
6	1.97*(0.24)	0.055*(0.02)	1.83*(0.56)	-1.07*(0.40)	0.87*(0.27)
7	-1.11*(0.26)	0.07*(0.03)	1.15*(0.43)	0.96*(0.23)	0.75*(0.28)
8	-0.91*(0.30)	0.86*(0.29)	1.73*(0.59)	0.52*(0.09)	0.08*(0.03)
9	1.73*(0.39)	0.94*(0.40)	2.79*(1.19)	-0.62 (0.39)	0.05*(0.02)
10	1.02*(0.27)	0.91*(0.28)	2.20*(0.70)	-0.46 (0.25)	0.07*(0.03)
11	0.97*(0.25)	0.90*(0.29)	2.14*(0.70)	-0.45 (0.25)	0.06*(0.03)
12	-0.67*(0.31)	0.90*(0.30)	2.07*(0.69)	0.32*(0.07)	0.06*(0.03)
13	-2.20*(0.51)	0.90*(0.37)	2.06*(0.86)	1.06*(0.24)	0.06*(0.03)
14	1.77*(0.27)	0.92*(0.31)	2.38*(0.80)	-0.74*(0.33)	0.06*(0.03)
15	1.47*(0.28)	0.96*(0.26)	3.45*(0.96)	-0.42*(0.17)	0.08*(0.03)

* indicates all the statistically significant parameters with a signification level below .05. SE: standard error of the estimated parameter.

CONCLUSIONS

Based on the analysis of the items in the three questionnaires, it can be concluded that the questionnaire that presents the best characteristic curves of the items is the third one, given that, for the 15 items, the discrimination indices present an adequate value, that is to say, the slope of the tangent of the item's characteristic curve is very steep. However, as far as item difficulty is concerned, there seem to be no differences from the other two. The items analyzed present an adequate measure, as shown by their characteristic curves as well as in the total estimated parameters in each case. Moreover, as was shown in a previous work (Guàrdia, Però, Freixa, Turbany & Gordóvil, 2009c), the fact of passing the course and the fact of turning in the three questionnaires are related. On the other hand, this paper also shows that the students with the best grade in the course are also the ones who present the highest score in the questionnaires analyzed, which would support an adequate predictive criterion validity of the three questionnaires analyzed in the present work.

REFERENCES

- Boyle, C. R. (1999). A problem-based learning approach to teaching biostatistics. *Journal of Statistics Education*, 7(1).
- Castejón, J. L., Navas, L., & Sampascual, G. (1996). Un modelo estructural del rendimiento académico en matemáticas en la educación secundaria. *Revista de Psicología General y Aplicada*, 49(1), 27-43.
- Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: towards an assessment framework. *Journal of Statistics Education*, 2(2).
- García, M.V., Alvarado, J.M., & Jiménez, A. (2000). La predicción del rendimiento académico: regresión lineal versus regresión logística. *Psicothema*, 12(supl. 2), 248-252.
- García, L., & Fumero, A. (1998). Personalidad y rendimiento académico en estudiantes universitarios: un estudio predictivo en tres cursos académicos. *Análisis y Modificación de Conducta*, 24(93), 65-77.
- Gardner, P., & Hudson, I. (1999). University students' ability to apply statistical procedures. *Journal of Statistics Education*, 7(1).
- Garfield, J. B. (1994). Beyond testing and grading: using assessment to improve student learning. *Journal of Statistics Education*, 2(1).

- Garrido, I., & Rojo, C. (1996). Motivación, cognición y rendimiento. *Revista de Psicología General y Aplicada*, 49(1), 5-12.
- Guàrdia, J., Freixa, M., Però, M., Turbany, J., Cosculluela, A., Barrios, M., et al. (2006). Factors related to the academic performance of students in the statistics course in Psychology. *Quality & Quantity*, 40(4), 461-474.
- Guàrdia, J., Freixa, M., Turbany, J., Cosculluela, A., Però, M., Barrios, M., et al. (2002). Estudio sobre los factores que inciden en el rendimiento académico en la asignatura análisis de datos en Psicología. *Revista de Metodología de las Ciencias del Comportamiento*, special volume, 275-278.
- Guàrdia, J., Freixa, M., Turbany, J., & Però, M. (2008). Collaborative learning in the teaching of statistics in psychology: an alternative to traditional teaching. *Revista de Formación e Innovación Educativa Universitaria*, 1(4), 96-106.
- Guàrdia, J., Però, M., Freixa, M., Turbany, J., Barrios, M., & Gordóvil, A. (2009a). *Elaboración de material para el aprendizaje autónomo del alumno en Análisis de Datos en Psicología impartida en la Universidad de Barcelona*. Communication presented at the I Congreso de Docencia Universitaria held in Vigo (Spain) from 2 to 4 of July.
- Guàrdia, J., Però, M., Freixa, M., Turbany, J., & Gordóvil, A. (in press, 2009b). Analysis of the Implications of Collaborative Work in Terms of Performance and Satisfaction among Students Registered in the *Data Analysis in Psychology* Course. *Quality & Quantity*.
- Guàrdia, J.; Però, M.; Freixa, M.; Turbany, J., & Gordóvil, A. (2009c). Análisis de cuestionarios de autoevaluación en la asignatura de estadística en psicología. Una aproximación al estudio de las curvas características de los ítems. *Revista Avances en Medición*. Accepted for publication.
- Guillén, M., Alea, V., Muñoz, C., Soldevilla, C., Torrelles, E., Viladomiu, N., et al. (2001). *Estadística Descriptiva Básica*. Barcelona: Edicions UB Colección: Team, 8 (CD-Rom).
- Huberty, C.J. (2000). Assessment of student performance in statistics. *Teaching Statistics*, 22(2), 44-48.
- Pérez-Sánchez, A. M., and Castejón, J. L. (1996). Factores psicosociales en la predicción del rendimiento académico. *Revista de Psicología de la Educación*, 19, 31-51.
- Peró, M., Turbany, J., Freixa, M., Guàrdia, J., Barrios, M., Ferrer, R., et al. (2004). Evaluación de la implementación de un CD-Rom de estadística descriptiva en la asignatura de Análisis de Datos en Psicología. *Metodología de las Ciencias del Comportamiento*, Volumen especial, 485-490.
- Roces, C., González, M.C., & Touron, J. (1997). Expectativas de aprendizaje y de rendimiento de los alumnos universitarios. *Revista de Psicología de la Educación*, 22, 99-123.
- Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education*, 6(3).