anuario de
psicología
The UB Journal of Psychology

Volumen 43
Número 3
Diciembre 2013

facultad de psicología
universidad de barcelona

HORSORI
EDITORIAL
ISSN: 0066-5126
ISSN: 1988-5253

Anuario de Psicología/The UB Journal of Psychology
2013, vol. 43, nº 3, 363-379

Determining assessment performance in Applied Statistics with ROC analysis*

Amparo Oliver[1]
Juana-María Vivo[2]
Laura Galiana[1]
Patricia Sancho[1]
[1] *Universitat de València*
[2] *Universidad de Murcia*

Difficulties in learning and teaching statistics have been a subject of interest for researchers and many studies have attempted to describe the factors influencing academic performance in the subject of statistics. The aim of this research is twofold: a) to test for the effects of attendance to statistics sessions, interest in the subject, collective feedback, satisfaction, and the existence of previous experience with applied statistics, in the students' assessment results in applied statistics; and b) to build an "instrument" to predict which students will find problems to pass the subject. Sample consisted of 166 students of statistics in psychology. Logistic regression and ROC curve analysis were used, and class attendance, collective feedback utility and previous experience with statistics emerged as good predictors of statistics achievement.
Keywords: *Performance in statistics, ROC analysis, class attendance, collective feedback, previous experience.*

Determinación de la evaluación del desempeño en Estadística Aplicada con análisis ROC

Las dificultades en la enseñanza y el aprendizaje de la estadística han sido objeto de interés de los investigadores. Muchos estudios han intentado describir los factores que influyen en los resultados de la asignatura estadística.

* *Acknowledgements:* This work was partially supported by Seneca Foundation of the Regional Government of Murcia (Spain) under Grant 11886/PHCS/09 and in the case of Laura Galiana supported by "V Segles/VLC-CAMPUS" pre-doctoral excellent rated program.
Correspondencia: Amparo Oliver. Departament de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de València. Av. Blasco Ibañez, 21, València. C.P. 46010. E-mail: oliver@uv.es

El objetivo de esta investigación es doble: a) evaluar los efectos de asistir presencialmente a las clases de estadística, del interés por la asignatura, del feedback colectivo, de la satisfacción y de la experiencia previa sobre el rendimiento en asignaturas de estadística aplicada; y b) construir un "instrumento" para pronosticar qué estudiantes tendrán problemas para aprobar la asignatura. La muestra consistió en 166 estudiantes de estadística en psicología. Se utilizaron una regresión logística y un análisis de curvas ROC, siendo la asistencia a las clases, el feedback colectivo y la experiencia previa con la asignatura los mejores predictores de un buen rendimiento en estadística.

Palabras clave: *resultados en estadística, análisis ROC, asistencia a clase, feedback colectivo, experiencia previa.*

Introduction

Difficulties in learning and teaching statistics are well known by teachers and students, and have been a subject of interest for many researchers during last years (e.g. Baloglu, 2003; Cherney & Cooney, 2005; Gal & Ginsburg, 1994; Garfield, 1995; Garfield & Ben-Zvi, 2007; Murtonen & Lehtinen, 2003; Shaughnessy, 1992; Smith, 1998). Most researchers have focused on the teaching and learning of statistics in college classes (Garfield & Ben-Zvi, 2007). These difficulties are clearly shown by Psychology students' outcomes, which are not as good as expected (Garfield & Ahlgren, 1988; Lehtinen & Rui, 1995), causing high drop out and failure rates.

Several studies suggest that statistics students learn better if they are engaged in, and are motivated with their learning (e.g., Capshew, 2005; Garfield & Ben-Zvi, 2007), which has an important impact in students' academic results (e. g. Cameron, Pierce, Banko, & Gear, 2005; Dweck, 1986). Class attendance is, in this sense, a good measure of students' engagement. It has been also directly related to positive effect on exam performance (Alvarado & García, 1997; Brocato, 1989; Cheng & Lin, 2008; García, Alvarado, & Jiménez, 2000; Gunn, 1993; Jones, 1984; Rocca, 2003; VanBlerkom, 1992). A recent study carried by Cheng and Lin (2008) operationalized the effect of attending lectures. In this randomized experiment, an improvement on exam performance that fluctuated between 9.4 and 18% was reported on the students who chose to attend lectures.

In the same vein, interest level in the subject has been related to students' success in many subjects (e. g., Ainley & Ainley, 2011; Shen, Chen, & Guan, 2007; Soric & Palekcic, 2009). However, no specific research has been done in applied statistics context. As some authors relate the lack of perceived relevance to disinterest and disengagement (Gal & Ginsburg, 1994), it seems appropriate to point out the relation between perceived statistics relevance and statistics performance (Capshew, 2005).

Feedback information also has been described as a factor improving students' understanding and results in statistics (e.g. Garfield, 1994, 1995; Krause, Stark & Mandl, 2009). Being aware of and confronting the misconceptions and errors enhance the process of learning statistics (Garfield & Ben-Zvi, 2007; Wild, Triggs, & Plannkuch, 1997), by giving students the chance to express ideas and test them, and helping to be active, learning by doing (Conners, Mccown, & Roskos-Ewoldsen, 1998). However, most of research in this area is merely descriptive. An important study in this sense is the laboratory experiment conducted by Krause *et al.* (2009), in which authors tested the effects of feedback on a statistics e-learning environment, both in individual and cooperative learning. In this research, it was found that feedback intervention clearly improved learning outcomes. On the contrary, no collaborative learning effect was found.

Many studies, nevertheless, have provided some evidence about the positive impact of collaborative learning on students' satisfaction and statistics results (e.g., Chick & Watson, 2002; Delucchi, 2006; Garfield, 1993, 1995; Guàrdia-Olmos *et al.*, 2008; Keeler & Steinhorst, 1995; Perkins & Saris, 2001; Smith, 1998). These studies support that working cooperatively in groups and learning to argue convincingly help students to learn better and, thus, improve their outcomes in statistics. For example, Perkins and Saris (2001) found several benefits of a collective statistics activity, like a useful tool for understanding the statistical procedure, in a post-test study. Guàrdia-Olmos *et al.* (2008) carried out a study of students' level of satisfaction with both collaborative learning and collaborative learning effects on the improvement of the academic performance. Whereas the satisfaction with this strategy was higher among the collaborative work students than among the individual work students, there were no effects on students' performance.

Satisfaction with the subject development is another important factor related to students' learning and performance (Aitken, 1982; Thomas, 2000). In the area of statistics, the study developed by Guàrdia *et al.* (2006) highlighted the role of satisfaction in the students' academic performance in a statistics course in Psychology. These authors proposed two structural models, being students' satisfaction the factor with greatest weight in the prediction of the mark obtained in one of them. Thus, satisfaction emerges as an important predictor of academic performance in statistics.

Finally, several studies have explored factors related to students' attitudes and previous experience in statistics classes (e. g. Cherney & Cooney, 2005; Dempster & McCorry, 2009; Elmore & Vasu, 1986; Mills, 2004; Onwuegbuzie, 2000, 2004; Schutz, Drogosz, White, & DiStefano, 1998; Suanpang, Petocz, & Kalceff, 2004). However, no researchers have attempt, as far as we know, to describe the effect of taking again the statistics course on the attitudes towards the subject. In this sense, it seems appropriate to operationalized previous experience with statistics in different groups:

– Those who have not previous experience with statistics.

– Those who have previous experience with statistics and failed, without any failure in other subjects.

– Those who have repeated experience with statistics and also had problems in other subjects.

As it has been shown, many researchers have studied factors influencing students' statistics performance. However, difficulties in operationalizing these variables and the virtual absence of attempts to build models for predicting this performance support the need and appropriateness of the present study. Thus, the aim of this research is twofold:

1. To test for the effects of attendance to statistics sessions, interest in the subject, collective feedback, satisfaction, and the existence of previous experience with applied statistics, in the students' assessment results.
2. To build an "instrument" to predict which students will find problems to pass the subject.

Method

Participants and procedure

The sample consisted of 166 Psychology students taking the subject of Statistics. 78.9% were female students. The participants were enrolled in three different groups:

1. Group 1 –first enrollment in statistics– was formed by 48 students, with an average age of 19.62 years (*SD* = 3.036). All of them were attending to the subject for the first time.
2. Group 2 –second or further enrollment in statistics, with problems in other subjects– was formed by 61 students. The mean age was 22.36 years (*SD* = 6.382). Every student had taken the subject before at least for twice, and had problems in other subjects.
3. Group 3 –second enrollment in statistics, without problems in other subjects– consisted of 57 subjects, with an average age of 19.92 years (*SD* = 3.499). All the students were repeating the subject for their first time, without fails in other subjects.

Sample was composed by natural groups. Groups were created by university management criteria: the registration system favors the natural formation of the three groups alleged. Students were eligible for each of these groups according to whether they were new students (1); if it was second tuition and between their previous qualifications they failed in more subjects (2) or just failed statistics (3). They voluntarily complete the questionnaires during an ordinary statistics session.

The three groups were taught by the same teacher, and follow the same teaching structure of the subject. The course consisted of 30 lessons of 2 hours of duration each of them, with a total of 60 hours. The assessment was obtained by the sum of the result in the final exam (with 50% of weight in the final assessment), the results in two exams that took place along the course (with 20% of weight each of them), and the result in a work done by pairs (with the resting 10% of weight). Three of the 30 lessons were dedicated to the collective feedback, always after the exams that accounted for the 40% of the assessment and took place along the course, and after the correction of the work did by pairs.

Instruments

Beside socio-demographic variables, students' grade point average, the time dedicated to study the subject, statistics anxiety, and perception of statistics use, the instrument built to determine probabilities of statistics success asked for information on:

– Class attendance (*CA*), comprising the item: "Percentage of statistics theory sessions you have attended to". The possible answers went from 0 to 100%.

– Interest (*I*): "Interest about Applied Statistics". One indicator from 1 (no interest at all) to 10 (total interest).

– Collective feedback utility (*CFU*). Measured by a single indicator from 1 (no interest at all) to 10 (total interest).

– Satisfaction (*S*): "Satisfaction with the subject development". It was measured by one indicator from 1 (no satisfied at all) to 10 (totally satisfied).3

– Previous experience with statistics (*PES*). Depending on the group the students belong to, students had no university statistics experience; one or more university previous experience in statistics plus other failed subjects; or just one university year previous experience in statistics (second tuition).

The final instrument, thus, had five items, each of them for the measurement of a variable. For more details, see Annex 1.

Statistical Analyses

Statistical analyses were performed on SPSS 19. The logistic regression for dichotomous endogenous variables was used.

In this context, there exists a trivial categorization procedure for the continuous variable *students' outcomes in the subject* (*OS*), and so then, each student is categorized as follows:

– "Success": if student passed the subject.
– "Failure": if student did not passed the subject.

Such statistical prediction rule (*SPR*) will allow us to determine which "predictor factors" (*CA*, *I*, *CFU*, *S* and *PES*) are relevant to success, as well as to evaluate and to express the evidence as an estimation of the probability that the condition of success is present (predictive probability).

Both calibration and discrimination capacity are provided by the logistic regression analysis. The ability to classify correctly of the logistic regression model is measured using the arbitrary choice of the cut-off value probability, "threshold decision", which determines whether the prediction of a student would be of success or not. Further, it can be evaluated by the receiver operating characteristic (ROC) curves.

The prediction of the students' success or failure in applied statistics is based on whether the predicted probability is higher or lower than a specified cut-off probability or decision threshold. Thus, the concordance between predictions and actual observations is summarized using a 2x2 classification table for each possible cut-off point, as it is shown in table 1.

TABLE 1. THE ROC CONFUSION MATRIX.

	Success	*Failure*	
Predicted positive	True-Positive (*a*)	False-Positive (*c*)	$a + c$
Predicted negative	False-Negative (*b*)	True-Negative (*d*)	$b + d$
	$a + b$	$c + d$	$a + b + c + d$

Note: The confusion matrix juxtaposes the observed classifications for the student's academic performance: success or failure (columns) with the predicted classifications of the model (rows). Each of the values *a*, *b*, *c*, and *d* represents numbers of observations, so that their summation, $a + b + c + d$, is equal to the sample size. The classifications that lie along the major diagonal of the table are the correct classifications, that is, the true positives (*a*) and the true negatives (*d*). The other fields signify incorrect classifications, that is, false negatives (*b*) and false positives (*c*).

Table 1 summarizes the correct and erroneous classifications, obtained by the model from several cut-off points, which are denoted as follows:

– True-Positive (*TP*): when the model predicts as success a student who passes the subject, i.e., the classification is correct.

– False-Positive (*FP*): when the model predicts as success a student who does not pass the subject, i.e., the classification is erroneous.

– False-Negative (*FN*): when the model predicts as failure a student who passes the subject, i.e., the classification is erroneous.

– True-Negative (*TN*): when the model predicts as failure a student who does not pass the subject, i.e., the classification is correct.

Their frequencies are denoted by *a*, *b*, *c*, and *d*, respectively. From the ROC confusion matrix a few performance measures can be derived such as sensitivity and specificity, which are two conditional proportions that allow us to analyse this classification through the cut-off points:

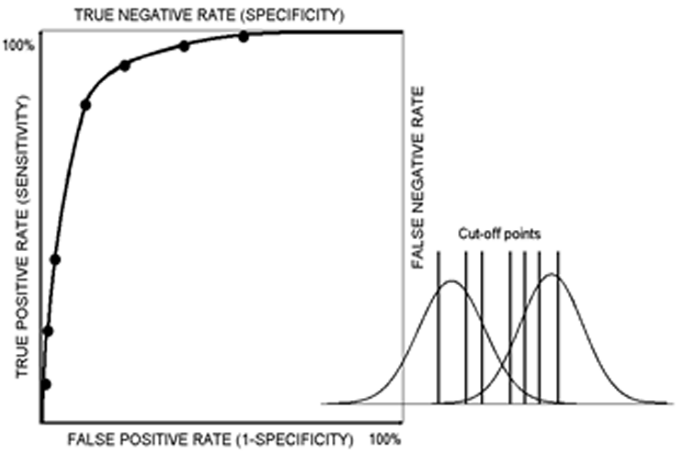
– *Sensitivity* $= a / (a + b)$ is the proportion of the students classified as success with respect to the true successes, i.e., positive prediction subject to "success", also called True Positive Rate (*TPR*).

– *Specificity* $= d / (c + d)$ is the proportion of the students classified as failure with respect to the true failures, i.e., negative prediction subject to "failure", also called True Negative Rate (*TNR*).

A good academic performance instrument should have both high sensitivity (for that the classifications in success to be useful) and high specificity (for that the classifications in failure to be useful). The sensitivity, specificity and theirs complementary (false positive rate and false negative rate, respectively) will vary as the decision threshold is changed. ROC curve is the plot of (1-specificity, sensitivity) with respect to the possible cut-off points, which represents the trade-off between sensitivity and specificity. So that, a ROC curve shows the trade-off between sensitivity and specificity across a series of cut-off points, see figure 1, next page. The optimal cut-off point is generally chosen in order to minimize some function of the erroneous predictions.

Although ROC curves themselves are useful in assessing the discriminatory ability of a logistic regression model, it is common to summarize the information of the ROC curve into a single global index. The area under the ROC curve, *AUC*, is the most widely used index due to it meets the requirements of an unbiased discrimination index (Metz, 1986; Fielding & Bell, 1997; Vivo & Franco, 2008).

The value of the AUC will always be between .50 (not apparent accuracy) and 1.00 (perfect accuracy), as the ROC curve moves towards the left and top boundaries of unit square. In particular, a random guessing will produce the diagonal line between (0,0) and (1,1), which has an area of .50. Indeed, this area is the probability that given two students, one who will pass the subject and the other who will fail, the model will assign a higher probability to the student with success (Handley & McNeil, 1982). Other interpretations have been given by different authors: the average sensitivity for all values of specificity or the average specificity for all values of sensitivity. For instance, a summarized review of the usual tools of the ROC analysis can be found in Franco and Vivo (2007) and the references therein.



Note: A ROC curve is a plot of the sensitivity (true positive rate) against the false positive rate (1-specificity) for a range of the cut-off points. Thus, if the cut-off point changes throughout this range then its associated sensitivity and specificity also change and in opposite directions from each other, such as shown in figure 1.

Figure 1. The ROC curve and seven cut-off points.

Results

Descriptive statistics of the variables under study are offered in table 2 (see next page).

In order to determine which predictor factors are associated with the success, was performed a logistic regression analysis following Hosmer and Lemeshow's criteria (Hosmer & Lemeshow, 2000). In a first stage, it was considered the variables whose odds ratios showed a statistical significance $p < .25$. Class attendance, collective feedback utility, and previous experience with statistics were considered as predictors by using the forward Wald stepwise method for variables selection with entry and removal probabilities .05 and .15, respectively.

In this stepwise selection procedure (IBM, 2011), the variables are tested for entry into the model one by one, with entry testing based on the significance of the score statistic and removal testing based on the probability of the Wald statistic. Thus, the variable with the smallest significance less than specified entry probability is entered into the model. And after each entry, the variable with the largest probability greater than the specified removal probability is removed, and the model re-estimated. Then, variables in the model are then evaluated again for removal and if no more variables satisfy the removal criterion, covariates that are not in the model are evaluated for entry. Unlike non-stepwise methods, the model building stops if no more variables find entry or removal criteria or if the current model is the same as a previous model. Subsequently possible interactions between the final variables were analyzed, without finding any.

TABLE 2. DESCRIPTIVE STATISTICS OF THE VARIABLES UNDER STUDY.

	Mean	*SD*	*Maximum score*	*Minimum score*
Class attendance (CA)	85.28	19.96	2	100
Interest (I)	6.84	1.80	0	10
Collective feedback utility (CFU)	8.51	1.45	3	10
Satisfaction (S)	8.48	5.97	3	10
	Percentage of students in each group			
	Group 1	*Group 2*	*Group 3*	
Percentage of students in each group (previous experience in statistics)	26.8%	36.7%	34.3%	

Results of the logistic regression model are shown in table 3. Group 3 was designed as the baseline for the categorical explanatory variable *PES*.

The fitted model is most easily interpreted by considering the odds ratios corresponding to the parameters.

$$Odds\,(success) = \exp\,(-3.362+0.033\cdot CA+0.377\cdot CFU-2.213\cdot PES\,(1)-2.851\cdot PES(2))$$

Test of the full model against a constant only model was statistically significant, indicating that the predictors as a set reliably distinguished between students' success or failure in applied statistics ($\chi^2 = 45.0626$, $p < .000$ with $df = 4$). Nagelkerke's R^2 was .367. The overall correct prediction was 73.5% for the cut-off point .50, in particular 48.1% for failure and 87.4% for success.

According to the model, class attendance and collective feedback utility are significant in the logistic regression model to predict the probability of success ($p = .015$ and $p = .013$, respectively), and probability is increasing as the level of the *CA* or *CFU* increases, since their odds ratios were 1.033 (95% CI from 1.006 to 1.061) and 1.458 (95% CI from 1.082 to 1.966), respectively. In addition, the *previous experience with statistics (PES)* is significant in the logistic regression model ($p = .003$), and its odds ratios indicates that, fitted for the effects of the other variables (*CA* and *CFU*):

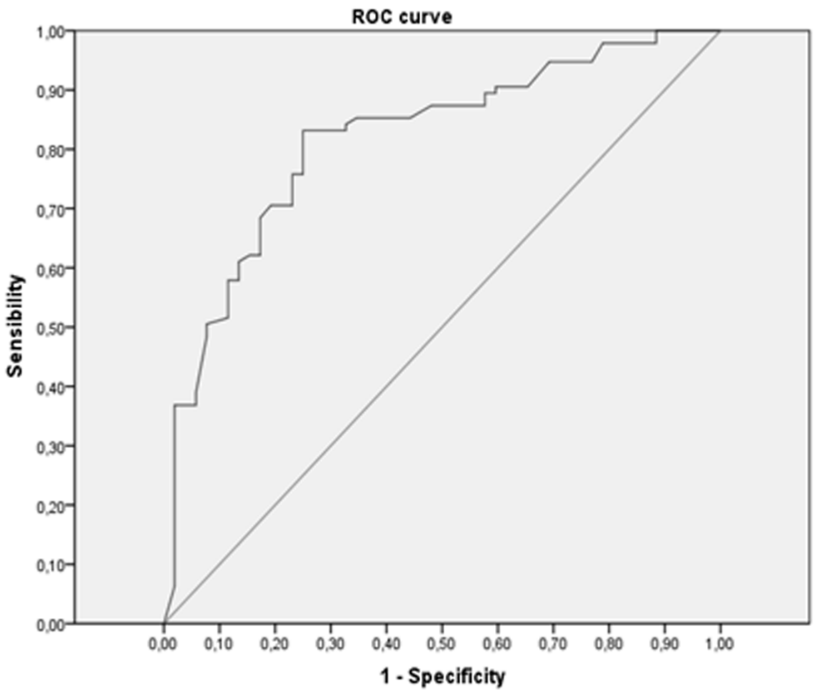
TABLE 3. RESULTS OF THE LOGISTIC REGRESSION ANALYSIS.

		β	$SE\ \beta$	$Wald's\ \chi^2$	df	p	e^{β} *odds ratio*	*95.0% C.I. for* e^{β}		
								Lower		*Upper*
Predictors	*CA*	.033	.013	5.935	1	.015	1.033	1.006		1.061
	CFU	.377	.152	6.132	1	.013	1.458	1.082		1.966
	PES			19.029	2	.000				
	PES (1)	-2.213	.751	8.683	1	.003	.109	.025		.477
	PES (2)	-2.851	.662	18.562	1	.000	.058	.016		.211
	Constant	-3.362	1.431	5.520	1	.019	.035			
								χ^2	df	p
Hosmer-Lemeshow		-	-	-	-	-	-	12.196	8	.143
Overall model evaluation		-	-	-	-	-	-	45.626	4	.000

Notes: *df* = degrees of freedom. SPSS Binary Logistic Regression procedure is available in the Regression option. It was considered the dichotomized OS as the dependent variable with CA, CFU and PES as the independent variables. In addition, PES variable is identified as categorical and its last category is selected as the reference.

– The students of the Group 1 are .109 times more likely to achieve success than the students of the Group 3 (95% CI from .025 to .477).

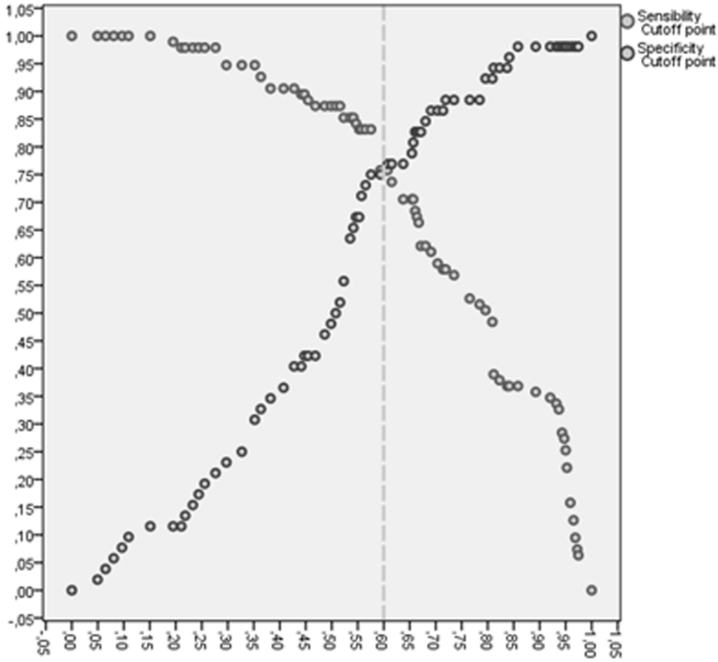
– The students of the Group 2 are .058 times more likely to achieve success than the students of the Group 3 (95% CI from .016 to .211).



Note: Discrimination capacity of the distribution model developed for the academic performance.

Figure 2. Area under the ROC curve.

In the ROC curve, the proportion of student correctly predicted to be success (sensitivity) and the proportion of the students incorrectly predicted to be success (1 minus specificity) are plotted against each other (figure 2). AUC was .818 ± .036 (95%CI from .747 to .889). This area shows high accuracy (Hosmer and Lemeshow, 2000). In addition, the optimal cut-off point of the ROC curve was .60, since the sensitivity (75.8%) and specificity (75.8%) are crossed between the cut-off points .593 and .606, i.e., the 75.8% of the overall students were correctly classified. This corresponds to minimal false negative and false positive results, as shown in figure 3 (see next page).



Note: A usual criterion for choosing the optimal cut-off point is given by the cut-off point for which sensitivity equals specificity. Figure 3 displays both sensitivity and specificity against different cut-off points for the sample data set. Thus, the optimal cut-off point, which can be defined as the crossing point of the two curves, was found to be .60.

Figure 3. The optimal cut-off point of the ROC curve.

Discussion

The aim of the current research was twofold. As regards the first objective, some variables have been proved to be predictors of students' outcomes in applied statistics, whereas some others found significantly relevant in previous investigations did not emerged as significant in the current research.

On one hand, interest level in the subject and satisfaction with the subject development were not significant predictors of the students' results. The relevance of students' interest level in the subject was found in different research fields (e.g., Ainley & Ainley, 2011; Shen *et al*., 2007; Soric & Palekcic, 2009). In this study, however, a non-significant effect was found for the students' interest in the subject. Satisfaction with the subject development was not found significant for predicting performance in statistics, neither. This result contradicts those found by Guàrdia *et al*. (2006), in which satisfaction with the subject emerged as the factor with greatest weight in the prediction of students' marks. Because of sample similarities, differences found may be due to the disparities on the model built.

On the other hand, class attendance, collective feedback utility, and previous experience with applied statistics have significantly predicted students' performance. In the model built for the current research, the odds-ratio of 1.033 for *CA* means that for every 1-unit increase in the dependent variable, the odds of being in the success student's group increases by 3.3%, similarly to the effect found for this variable by Cheng and Lin (2008). This result supports the positive relation between class attendance and exam performance existing in previous literature (Alvarado & García, 1997; Brocato, 1989; Cheng & Lin, 2008; García *et al.,* 2000; Gunn, 1993; Jones, 1984; Rocca, 2003; VanBlerkom, 1992), extending it to the applied statistics specific area. Utility of collaborative feedback has also emerged as a statistically significant predictor for students' performance, which odds-ratio of 1.458 means that for every 1-unit increase in that independent variable, the odds ratio of being in the success students' group increases by 45.8%. This predicting capacity is in line with similar studies carried on by several authors (e. g. Chick & Watson, 2002; Guàrdia-Olmos *et al.,* 2008; Krause *et al.*, 2009; Perkins & Saris, 2001), and confirms the hypothesis that feedback activities enhance students' statistics results. Finally, attitudes and previous experience with statistics has also arisen as a crucial variable for predicting not only performance, but probabilities of success or failure, too. The odds-ratio of .109 for the Group 1 of the *PES* indicates that a student of the Group 3 is 9.17 times more likely to achieve success than other of the Group 1, i.e., the odds of being in the success students' group decreases by 89.1% for every change from the Group 3 to Group 1. Likewise, the odds-ratio of .058 for the Group 2 of the *PES* indicates that a student of the Group 3 is 17.24 times more likely to achieve success than other of the Group 2, i.e., the odds of being in the success students' group decreases by 94.2% for every change from the Group 3 to Group 2. Results showed that students' scores equal or lower to .60 are clear sign of problems in applied statistics, with an increased probability to fail.

Current research has extended previous literature, with two main strengths: the operationalization of different predictors of performance in applied statistics, in order to test for their effect on this construct; and the construction of an instrument for detecting students with higher probabilities to fail the subject. However, the study has some limitations: feedback and collective learning have been assessed together with the variable collective feedback utility, and so, no differentiations between these two effects could be done; and previous attitudes have not been tested, by their own, but have been taken over by group variable. Another point to take into account is the assessment of previous experience in statistics. Whereas in Group 1 it is clear that students have no previous experience, and in Group 3 that they have had only one previous experience, Group two combine both people with their second experience in statistics and people with three or more previous tuitions. Finally, there is also the non-randomized sample procedure. As groups were naturally given, this makes more difficult the representativeness of the results of the current research. Also, this can be the reason of the non-significant effects of some variables, which may be, with other sampling procedures, could become significant.

Thus, future research should take into account these limitations, combining more accurate definition and reliability study of predictors, and more research with reliable indicators of these variables is needed. Moreover, a study on the instrument truthfully capacity to detect students' with higher failure probabilities and its implications for classroom interventions would be an interesting research arena.

REFERENCES

Ainley, M., & Ainley, J. (2011). Student engagement with science in early adolescence: The contribution of enjoyment to students' continuing interest in learning about science. *Contemporary Educational Psychology, 36*(1), 4-12.

Aitken, N.D. (1982). College student performance, satisfaction and retention: Specification and estimation of a structural model. *The Journal of Higher Education, 53*(1), 32-50.

Alvarado, J., & Jiménez, M.V. (1997). Incidencia de la asistencia a clase, del trabajo efectivo y de factores individuales sobre el rendimiento académico. *V Congreso de Metodología*, Sevilla.

Baloglu, M. (2003). Individual differences in statistics anxiety among college students. *Personality and Individual Differences, 34*, 855-865.

Brocato, J. (1989). How much does coming to class matter? Some evidence of class attendance and grade performance. *Educational Research Quarterly, 13*(3), 2-6.

Cameron, J., Pierce, W.D., Banko, K.M., & Gear, A. (2005). Achievement-based rewards and intrinsic motivation: A test of cognitive mediators. *Journal of Educational Psychology, 97*(4), 641-655.

Capshew, T.F. (2005). Motivating social work students in statistics courses. *Social Work Education, 24*(8), 857-868.

Chen, J., & Lin, T-F. (2008). Class attendance and exam performance: A randomized experiment. *Journal of Economic Education, 39*(3), 213-227.

Cherney, I.D., & Cooney, R.R. (2005). Predicting student performance in a statistics course using the mathematics and statistics perception scale (MPSP). *Transactions of the Nebraska Academy of Sciences, 30*, retrieved on 1st of November 2011 from http://www.icherney.com/Research/Publi cations/MPSP.pdf.

Chick, H.L., & Watson, J.M. (2002). Collaborative influences on emergent statistical thinking – a case study. *Journal of Mathematical Behavior, 21*, 371-400.

Conners, F.A., Mccown, S.M., & Roskos-Ewoldsen, B. (1998). Unique challenges in teaching undergraduate statistics. *Teaching of Psychology, 25*, 40-42.

Delucchi, M. (2006). The efficacy of collaborative learning groups in an undergraduate statistics course. *College Teaching, 54*(2), 244-248.

Dempster, M., & McCorry, N.K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education, 17*(2), retrieved on 1st of November 2011 from www.amstat.org/publications/jse/v 17n2/dempster.html.

Dweck, C.K. (1986). Motivational processes affecting learning. *American Psychologist, 41*(10), 1040-1048.

Elmore, P. B., & Vasu, E. S. (1986). Model of statistics achievement using spatial ability, feminist attitudes and mathematics-related variables as predictors. *Educational and Psychological Measurement, 46*(1), 215-222.

Fielding, A.H., & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv., 24*, 221-224.

Franco, M, & Vivo, J.M. (2007). Análisis de curvas ROC: principios básicos y aplicaciones. Madrid: La Muralla.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education, 2*(2), retrieved on 1st of November 2011 from http://www.amstat.org/publications/jse/v2n2/gal.html.

Gal, I., & Ograjensek, I. (2010). Qualitative research in the service of understanding learners and users of statistics. *International Statistical Review, 78*(2), 287-296.

García, M.V., Alvarado, J.M., & Jiménez, A. (2000). La predicción del rendimiento académico: Regresión lineal versus relación logística. *Psicothema, 12*(2), 248-252.

Garfield, J. (1993). Teaching statistics using small-group cooperative learning. *Journal of Statistics Education, 1*(1), retrieved on 1st of November 2011 from http://www.amstat.org/publications/jse/v1n1/garfield.html.

Garfield, J. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education, 2*(1), retrieved on 1st of November 2011 from http://www.amstat.org/publications/jse/v2n1/garfield.html.

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

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*, 44-63.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372-396.

Guàrdia, J., Freixa, M., Peró, M., Turbany, J., Cosculluela, A., Barrios, M., & Rifà, X. (2006). Factors related to the academic performance of students in the statistics course in Psychology. *Quality & Quantity, 40*, 661-674.

Guàrdia-Olmos, J., Freixa-Blanxart, M., Turbany-Oset, J., & Peró-Cebollero, 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.

Gunn, P. (1993). A correlation between attendance and grades in a first-year psychology course. *Canadian Psychology, 34*(2), 201-202.

Hanley, J.A., & McNeil, B.J. (1982). The meaning and use of the area under receiver operating characteristic (ROC) curve. *Radiology*, *134*(1), 29-36.

Hosmer, D.W., & Lemeshow, S. (2000). Applied Logistic Regression, 2nd ed., New York, Chichester, Wiley.

IBM. IBM SPSS Regression 20. (2011). 2–8. http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Regression.pdf.

Jones, H. (1984). Interaction of absences and grades in a college course. *Journal of Psychology, 116*(1), 133-136.

Keeler, C.M., & Steinhorst, R.K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education, 3*(2), retrieved on 1st of November 2011 from http://www.amstat.org/publications/jse/v3n2/keeler.html.

Krause, U., Stark, R., & Mandl, H. (2009). The effects of cooperative learning and feedback on e-learning statistics. *Learning and Instruction, 19*, 158-170.

Lehtinen, E., & Riu, E. (1995). Computer-supported complex learning: An environment for learning experimental methods and statistical inference. *Machine-Mediated Learning, 5*(3&4), 149-175.

Metz, C.E. (1986). ROC methodology in radiologic imaging. *Invest. Med., 8*, 720-733.

Mills, J.D. (2004). Students' attitudes toward statistics: Implications for the future. *College Student Journal, 38*(3), 349-361.

Murtonen, M., & Lehtinen, E. (2003). Difficulties experienced by education and sociology students in quantitative method courses. *Studies in Higher Education, 28*(2), 171-185.

Onwuegbuzie, A.J. (2000). Attitudes toward statistics assessments. *Assessment & Evaluation in Higher Education, 25*(4), 321-339.

Onwuegbuzie, A.J. (2004). Academic procrastination and statistics anxiety. *Assessment & Evaluation in Higher Education, 29*(1), 3-19.

Perkins, D.V., & Saris, R.N. (2001). A "Jigsaw Classroom" technique for undergraduate statistics courses. *Teaching of Psychology, 28*(2), 111-113.

Rocca, K. (2003). Student attendance: A comprehensive literature review. *Journal on Excellence in College Teaching, 14*(1), 85-107.

Schutz, P.A., Drogosz, L.M., White, V.E., & DiStefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics course. *Learning and Individual Differences, 10*(4), 291-308.

Shaughnessy, M.J. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan.

Shen, B., Cheng, A., & Guan, J. (2007). Using achievement goals and interest to predict learning in physical education. *Journal of Experimental Education, 75*(2), 89-108.

Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education, 6*(3), retrieved on 1st of November 2011 from http://www.amstat.org/publications/jse/v6n3/smith.html.

Soric, I., & Palekcic, M. (2009). The role of students' interests in self-regulated learning: The relationship between students' interests, learning strategies and causal attributions. *European Journal of Psychology of Education, 24*(4), 545-565.

Suanpang, P., Petocz, P., & Kalceff, W. (2004). Student attitudes to learning business statistics: Comparison of online and traditional methods. *Journal of Educational Technology & Society, 7*(3), 9-20.

Thomas, S.L. (2000). Ties that bind: A social network approach to understanding student integration and persistence. *The Journal of Higher Education, 71*(5), 591-615.

VanBlerkom, L. (1992). Class attendance in an undergraduated course. *Journal of Psychology, 126*(5), 487-494.

Vivo, J.M., & Franco, M. (2008). How does the accuracy of academic success predictors assess? ROC analysis applied to university entrance factors. *International Journal of Mathematical Education in Science and Technology, 39*(3), 325-340.

Wild, C., Triggs, C., & Plannkuch, M. (1997). Assessment on a budget: Using traditional methods imaginatively. In I. Gal & J.B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 205-220). IOS Press.

ANNEX 1
INSTRUMENT FOR DETERMINING ASSESSMENT PERFORMANCE IN APPLIED STATISTICS

– Percentage of statistics theory sessions you have attended to: ______

– Indicate your interest about Applied Statistics, from 1 (no interest at all) to 10 (total interest): ______

– Indicate the utility of the collective feedback utility, from 1 (no interest at all) to 10 (total interest): ______

– Indicate your satisfaction with the subject development, from 1 (no satisfied at all) to 10 (totally satisfied): ______

– Indicate the class group you belong to: ______

anuario de psicología

Artículos

P.V.P.: 14€ (I.V.A. incluido)

ISSN 1988-5253
9 771988 525007
ISSN 00665126
9 770066 512007 36105