From Knowledge to Wisdom

ISSN 1537-1506 CBSN 270B0069

Chinese Business Review

Volume 10, Number 8, August 2011

DAVID PUBLISHING

David Publishing Company
www.davidpublishing.com

Chinese Business Review

Volume 10, Number 8, August 2011 (Serial Number 98)

Contents

Chinese Business Review

Volume 10, Number 8, August 2011 (Serial Number 98)

DAVID PUBLISHING

David Publishing Company
www.davidpublishing.com

Publication Information:
Chinese Business Review is published monthly in hard copy (ISSN1537-1506) and online by David Publishing Company located at 1840 Industrial Drive, Suite 160, Libertyville, Illinois 60048, USA.

Aims and Scope:
Chinese Business Review, a monthly professional academic journal, covers all sorts of researches on Economic Research, Management Theory and Practice, Experts Forum, Macro or Micro Analysis, Economical Studies of Theory and Practice, Finance and Finance Management, Strategic Management, and Human Resource Management, and other latest findings and achievements from experts and scholars all over the world.

Manuscripts and correspondence are invited for publication. You can submit your papers via Web Submission, or E-mail to economists@davidpublishing.com, china4review@hotmail.com. Submission guidelines and Web Submission system are available at http://www.davidpublishing.com.

Editorial Office:
1840 Industrial Drive, Suite 160 Libertyville, Illinois 60048
Tel: 1-847-281-9862
Fax: 1-847-281-9855
E-mail: economists@davidpublishing.com

Abstracted / Indexed in:
Database of EBSCO, Massachusetts, USA
Ulrich's Periodicals Directory
ProQuest/CSA Social Science Collection, Public Affairs Information Service (PAIS), USA
Summon Serials Solutions
Chinese Database of CEPS, Airiti Inc. & OCLC
Chinese Scientific Journals Database, VIP Corporation, Chongqing, P. R. China

Subscription Information:
Print $450 Online $300
Print and Online $560
David Publishing Company
1840 Industrial Drive, Suite 160, Libertyville, Illinois 60048
Tel: 1-847-281-9862 Fax: 1-847-281-9855
E-mail: order@davidpublishing.com

DAVID PUBLISHING

David Publishing Company
www.davidpublishing.com

Chinese Business Review, ISSN 1537-1506
August 2011, Vol. 10, No. 8, 632-639

The Future University Choices of Secondary Schools Students: Statistical Evaluations in Messina Side

Angela Alibrandi
Messina University, Via dei Verdi, Messina, Italy
Massimiliano Giacalone
Calabria University, Via P. Bucci, Rende (CS), Italy

In the last decades, especially since the 1990s, there was a gradual rising of educational levels, due to a growing schooling. This paper aims to analyze the propensity toward university enrolment in the Messina area, by means of appropriate statistical methods. In particular, we compared the students of different secondary school institutes in Messina, with reference to the choice of the future university career and other information about the scholastic profit and the scholastic context. Our comparative analysis has been performed through a non-parametric approach, using the Non Parametric Combination (NPC) test based on permutation test. This methodology was chosen for optimal characteristics of which it is characterized.

Keywords: university registration, permutation test, comparison among schools

Introduction

In recent decades, especially since the 1990s, there was a gradual raising of educational levels, due to a growing schooling. The evolutionary tendencies of the last decades in terms of education have delineated a background characterized by a progressive increase and a tendency of the formative levels toward a more and more strong schooling of the new demographic levers.

In the Italian universities, in the academic year 2008-2009, the Progetto Lauree Scientifiche (PLS) project was born to promote the scientific university faculties that, in the recent past, suffered a worrisome decrement of registrations. Such phenomenon is to be considered as a social phenomenon. The planning of activity directed to stimulate the registration from the students in University faculties cannot put aside from a careful analysis of the same phenomenon and it is essential, in this context, a continuous monitoring of the causes that produced it.

The purpose of this research is to understand the phenomenon of the propensity toward university enrolment in the Messina area, by means of appropriate statistical methods. In particular, we want to compare the students of different secondary school institutes in Messina, with reference to the choice of the future University career and other information about the scholastic profit and the scholastic context.

Angela Alibrandi, Researcher, Department of Economical, Financial, Social, Environmental, Statistical and Territorial Sciences, Messina University.
Massimiliano Giacalone, Adjunct Professor, Faculty of Political Sciences, Calabria University.
Correspondence concerning this article should be addressed to Angela Alibrandi, Via dei Verdi, 75-98122 Messina. E-mail: aalibrandi@unime.it.

The paper is organized as follows: in section 2, the dataset is described; in section 3, the NPC test procedure is illustrated; in section 4, the results of NPC test are shown; in section 5, the logistic regression model is described (since such model is the most adequate one to evaluate the dependency relationship); in section 6, the results of the logistic model are shown; finally in section 7, some concluding remarks are discussed.

The Data

The sample is formed by students of those secondary school institutes that gave their adhesion to the PLS project in the academic year of 2008-2009.

Within the PLS project, the faculty of statistical sciences of Messina University started a laboratory-activity of "statistics in class" from February to May 2009, thus, students completed a questionnaire, primarily oriented to the collection of information about the possible choice of the University career (Soliani, 2004).

The students who participated in the PLS activity, compiling the questionnaire, were 1,180 of which 509 were of male (43.1%) and 671 were female (56.9%) (see Figure 1), with a mean age of 19.2. They were divided by attended school, as reported in Table 1 and Figure 2.

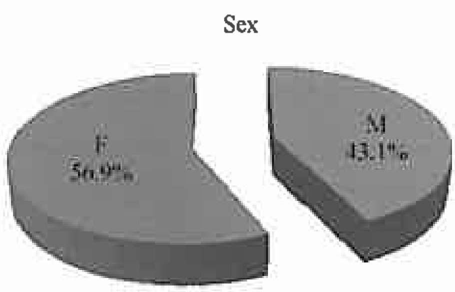

Figure 1. Classification of students according to the sex.

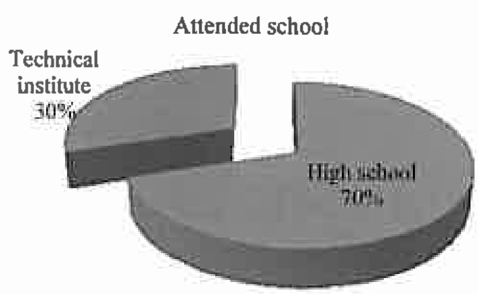

Figure 2. Classification of students according to the attended secondary school.

Table 1

Classification of Students According to the Attended Secondary Schools

High schools	No.	Percent (%)	Technical institutes	No.	Percent (%)
Classical	233	28.17	Commercial	117	14.15
Scientific	433	52.36	Industrial	130	15.72
Linguistic	30	3.63	Surveyors	58	7.01
Art	7	0.85	Nautic	10	1.21
Social sciences	86	10.40	Agrarian	33	3.99
Educational	38	4.59	Professional	5	0.60

The collected data refer to the section "Personal details", "Schooling", "University choice", "Family" and "Leisure". Information relating to the section "University choice" assumes a great relevance because it contains the propensity at university enrolment and the indication of a possible choice. The examined variables are the following: age, gender, type of institution, school career (promotion, promotion with debits, sometimes rejected), preferred disciplinary matters, relationship with math and the math teacher, grade in mathematics (written and oral), university enrolment propensity and any faculty choice (scientific or humanistic), number of family members, father and mother qualifications, eventual University brothers, friends' influences , time spent in studying.

In detail, the likely entry is motivated by a desire to acquire new knowledge and the ability to find a more rewarding job (Checchi, Leonardi, & Fiorio, 2007). This result is consistent with previous reports in the literature (Checchi, Bratti, & De Blasio, 2008), the choice of those who think they do not subscribe, however, is mainly linked to the desire to immediately enter the world of work.

On the basis of different scholastic origins of the interviewed students (high school or technical institutes), we want to perform a statistical comparison, aimed to investigate the existence of significant differences between students attending high school and students attending technical institutes, in reference to each detected variable and, above all, in relation to university enrolment propensity.

The NPC Methodology

Our comparative analysis has been performed through a non-parametric approach, using the Non Parametric Combination (NPC) test based on a permutation test (Pesarin, 2001; Arboretti & Corain, 2009; Corain & Salmaso, 2007) chosen for optimal characteristics of which it is characterized:

- It doesn't request normality and homoschedasticity assumption;
- It draws any type of variable;
- It assumes a good behaviour also in presence of lacking data;
- It is powerful in presence of low sampling size;
- It resolves multivariate problems without the necessity to specify the structure of dependence among variables;
- It allows stratified analyses;
- It resolves problems in which the number of observations is smaller than that of variables.

The NPC test allows the resolution of a lot of complex problems related to verification of multidimensional hypotheses. We consider that two or more (k) variables are observed on a set of n statistical units, that are grouped in two or more C groups defined by a classification criterion (a treatment). Purpose of this analysis is to verify if there are statistically significant differences among the multivariate profiles of answer-variables of the C compared groups.

We suppose that an appropriate k-dimensional P distribution (with $P_j \in F, j = 1, ..., C$) exists; the null hypothesis establishes the equality in distribution of the k-dimensional distribution among the C groups, against the alternative one, where at least one strict inequality is satisfied as:

$$H_0 : [X_1 \overset{d}{=} ... \overset{d}{=} X_C]$$

which can be also expressed as:

$$H_0 : [\bigcap_{i=1}^{k} X_{1i} \overset{d}{=} \ldots \overset{d}{=} X_{ci}] = [\bigcap_{i=1}^{k} H_{0i}]$$

against:

$$H_1 : \bigcup_{i=1}^{k} [X_{1i} \overset{d}{\neq} \ldots \overset{d}{\neq} X_{Ci}] = [\bigcup_{i=1}^{k} H_{1i}]$$

By means of mentioned procedure, it is preliminarily possible to define a set of k one-dimensional permutation tests, denominated partial test, through which the marginal contribution of every answer-variable can be examined in the comparison among groups.

The partial tests are non-parametrically combined through Conditional Monte Carlo (CMC) procedure in combined tests, using an opportune combination function (generally Fisher, Tippett or Liptak), and these tests globally verify the existence of differences among the multivariate distributions of the groups.

If the analysis is stratified, it's possible to determine an only test that combines the gotten tests by every stratification, and it allows drawing evaluations on the possible differences among the groups in relationship to all examined variables and all strata.

If the p-value is significant at a prefixed significance level, the null hypothesis has to be rejected. This procedure allows verifying directional hypotheses, solving the problems of restricted alternative hypotheses. Figure 3 shows the two-phase algorithm, which illustrates the NPC procedure.

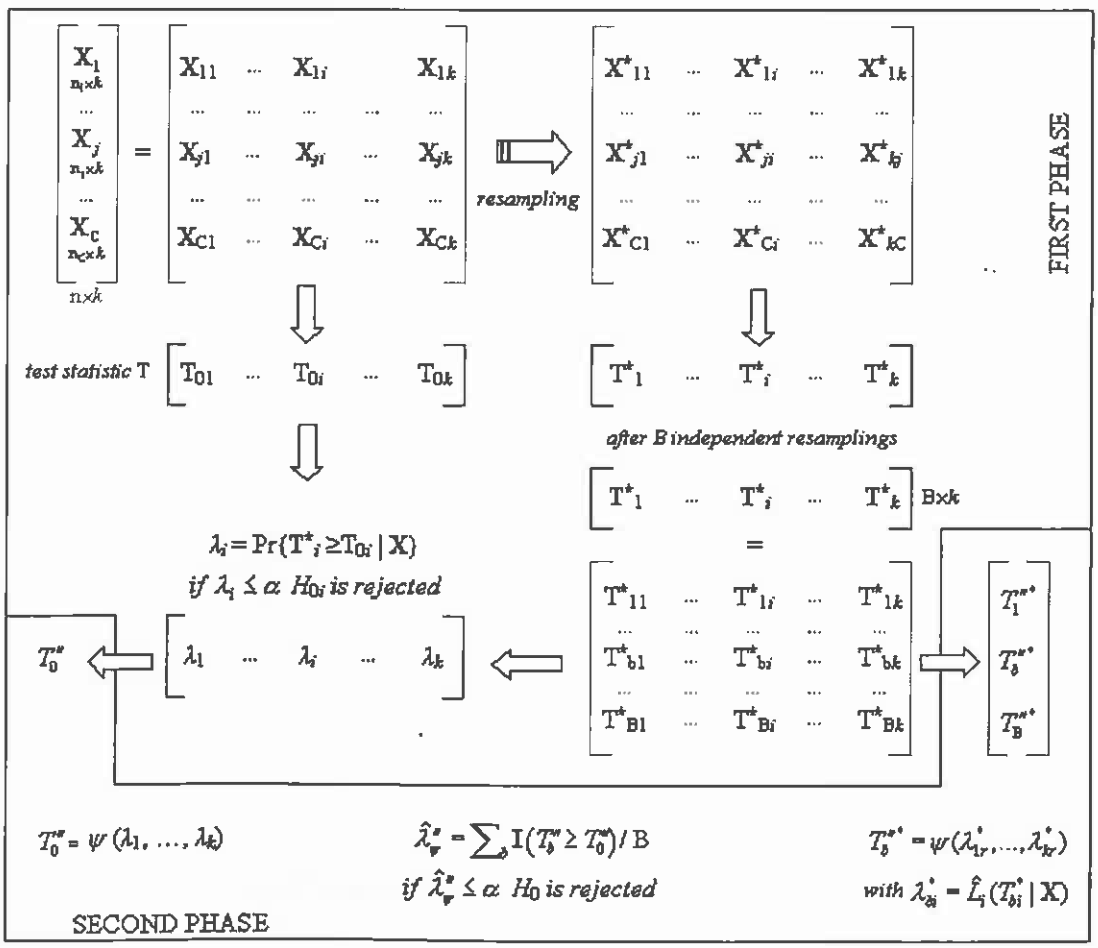

Figure 3. Two phases NPC algorithm.

Results of NPC Application

By means of the above mentioned permutation solution, we want to investigate the existence of possible significant differences between the high school institutes and the technical institutes for the measured variables.

The null hypothesis that postulates the indifference among the variables distributions, and the alternative hypothesis can be expressed through the relationships:

$$H_0 : \left\{ Age_{High\ School} \overset{d}{=} Age_{Tecnhical\ Institutes} \right\} \cap \ldots \cap \left\{ Time\ for\ study_{High\ School} \overset{d}{=} Time\ for\ study_{Tecnhical\ Institutes} \right\}$$

Against:

$$H_1 : \left\{ Age_{High\ School} \overset{d}{\neq} Age_{Tecnhical\ Institutes} \right\} \cup \ldots \cup \left\{ Time\ for\ study_{High\ School} \overset{d}{\neq} Time\ for\ study_{Tecnhical\ Institutes} \right\}$$

The comparison among the two kind of secondary schools, performed by the aforesaid methodology, furnished the results (partial and combined p-values and directionality of comparisons) shown in Table 2.

Table 2

Results of Comparison Performed by NPC Test Between High School and Technical Institutes

Compared variables (high school vs. technical institutes)	p-value	Directionalitys
Age	0.620	
Sex	0.437	
Promotion	0.237	
Mean marks	0.043	<
University enrolment	0.000	>
Friends' influences	0.012	>
Father studies	0.027	>
Mother studies	0.000	>
Eventual University brothers	0.002	>
Time for study	0.001	>
	↓	
Combined	0.037	

Examining the NPC test results, we found that high school students and technical institutes students have a different profile, expressed by the different variables, such as shown by combined p-value. In particular, analyzing the partial p-value, we can observe that high school students seem to have a more elevated significant propensity towards the university enrolment compared to the technical institutes, at the significance α level of 0.05. The mean mark in technical institutes is significantly higher than the high school, moreover, there is a directionality in favour of high schools with reference to the time for study, the friends' influence and the familiar context, because the higher parents' qualification corresponds to a major propensity to begin the university studies.

Another interesting result is related to the presence, in the family, of brothers that attend the university: high school students have, with a probability higher than technical institutes students, brothers who attend the university.

These results could be seen as an evidence of the positive influence of family background, that can represent a motivating factor in choosing. Finally, there is no significant differences between the two types of secondary schools in relation to the other examined variables.

The Logistic Regression Model

Methodological Aspects of Logistic Regression Model

Logistic regression model is a special case of generalized linear model with link function as the logit function (Corbetta, Gasperoni, & Pisati, 2001; Piccolo, 2010). This is a regression model applied in cases where the dependent variable y is precisely the type attributable to the dichotomous values of 0 and 1. The probability that the variable y took the value 1 is a function of the regressors, $P(Y = 1 \mid X = x)$. In particular, we have:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)}$$

For the interpretation of the parameters, we have recourse to odds, i.e., the ratio of the probability that $Y = 1$ and $Y = 0$, conditional on the value assumed by the regressors:

$$odds(x) = \frac{P(Y = 1|X = x)}{P(Y = 0|X = x)} = \frac{P(Y = 1|X = x)}{1 - P(Y = 1|X = x)} = \frac{\pi(x)}{1 - \pi(x)}$$

The logit, i.e., the logarithm dell'odds, is a linear function of the parameters:

$$\log it(x) = \ln[odds(x)] = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

If the parameter β_j is positive, an increase in the explanatory variable x_j induces an increase in the probability that the dependent variable takes the value 1.

Denoting with $y_1, y_2, \ldots, y_n$, the observed values of the dependent variable and with x_i vectors containing the corresponding values of the explanatory variables, we yield the following likelihood function $L(\beta)$ and log-likelihood function $l(\beta)$:

$$L(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i}$$

$$l(\beta) = \sum_{i=1}^{n} \{y_i \ln[\pi(x_i)] + (1 - y_i)\ln[1 - \pi(x_i)]\}$$

The Results of the Logistic Regression Model Estimation

In order to obtain a statistical model able to express and test the dependence of the "propensity to university enrolment", given the nature of the variables, we considered appropriate to use the logistic regression model (Kleimbaum, 2002). In particular, we estimated two statistical models: The first aims at formalizing the dependence of the propensity for university enrolment by variables related to the student, as educational background and influence from family and peers; The second aims at identifying factors related to the choice of the university faculties, only for students planning to enrolment.

In order to choose the most appropriate model, we estimated several logistic models, the models that have provided the best results are reported in Table 3 and Table 4.

The model furnished a Log-Likelihood value of -222.775, a Hosmer-Lemeshow test of 5.955 (p-value = 0.652), a G test value of 211.715 (with degrees of freedom = 8 and p-value = 0.000) and a Goodman-Kruskal Gamma = 0.72.

Examining the results reported in Table 1, we can see that there is a strong propensity university enrolment for students attending high schools compared to those attending technical colleges (1 = high school; 0 = technical school); moreover, students who are usually promoted are most likely to go to university (1 = usually promoted, 0 = sometimes rejected). The gender of students is significant and coefficient positive sign indicates that male students shows a higher propensity to enrolment than female ($M = 1$, $F = 0$). The mothers' level of education appears to be a significant factor, with a positive sign coefficient, indicating that the sons of more educated parents are most likely to go to university.

Table 3

Logistic Regression Model for Propensity University Enrolment

Predictor	Coeff.	ES	Z	P	OR	Inf.I.C.	Sup.I.C.
Constant	-3.269	0.677	-4.82	0.000	-	-	-
Sex	0.869	0.374	2.33	0.020	2.39	1.15	4.97
Type of school	2.415	0.280	8.60	0.000	11.19	6.45	19.42
Promotion	1.151	0.244	4.71	0.000	3.16	1.96	5.11
Peer influence	-0.226	0.154	-1.46	0.143	0.80	0.59	1.08
Father educ.	-0.151	0.153	-0.98	0.325	0.86	0.64	1.16
Mother educ.	0.782	0.180	4.35	0.000	2.19	1.54	3.11
Univ.brothers.	0.303	0.259	1.17	0.242	1.35	0.82	2.25
Study time	0.072	0.082	0.88	0.380	0.08	0.91	1.27

The model furnished a Log-Likelihood value of -130.992, a Hosmer-Lemeshow test of 4.163 (p-value = 0.842), a G test value of 206.001 (with degree of freedom = 7 and p-value = 0.000) and a Goodman-Kruskal Gamma = 0.81.

Table 4

Logistic Regression Model for Propensity to Enrolment in Scientific Faculties

Predictor	Coeff.	ES	Z	P	OR	Inf.I.C.	Sup.I.C.
Constant	-2.618	0.942	-2.78	0.005	-	-	-
Sex	1.415	0.445	3.18	0.001	4.12	1.72	9.86
Preferred field	3.183	0.408	7.80	0.000	24.14	10.85	53.71
Maths rel.	1.001	0.176	3.08	0.002	1.73	1.22	2.44
Teach. maths rel.	-0.087	0.162	-0.54	0.591	0.92	0.67	1.26
Univ. brothers	0.282	0.337	0.83	0.404	1.33	0.68	2.57
Study time	-0.350	0.122	-2.87	0.004	0.70	0.55	0.90
Peer influence	0.139	0.214	0.65	0.516	1.15	0.75	1.75

From the results shown in Table 4, there is a significant propensity to enrolment in scientific faculties rather than humanities faculties (mostly composed by male students). A significant variable is the "relationship with mathematics" and the coefficient positive sign indicates a higher propensity to enrolment in scientific skills for students attracted by this discipline. The relationship with the maths teacher seems to be not significant, even the hours of study appear to be a statistically significant factor.

For both logistic models, we estimated the log-likelihood test (or G test) for the model significance, the Hosmer-Lowershow test for goodness of fit and, finally, the Goodman-Kruskal Gamma test, in particular, the rather high value of this last test shows a good degree of fit between the modes of the response variable and the expected probabilities.

Final Remarks

Based on our statistical analysis, it was possible to reach some interesting results, such as to explain the propensity of leavers in the province of Messina towards university entry.

The study, performed in the secondary schools of Messina, on the basis of collected data, essentially confirms that the propensity to the University enrolment is related to the kind of the attended school, and therefore largely already determined when a student choose the secondary school. In fact, it appears that a great part of the high school students decides to continue studies, the others have lower rates.

An important result is the significance of the parents instruction and, in particular, of the mothers: the propensity of sons in the university choice increases with mother education level, of course, mothers with higher educational qualifications frequently motivate their children. So, the motivation exercised by the family into choosing the university studies seems to be very strong.

Even if the survey should be replicated on a more representative sample, we seem to be able to say, on the bases of the obtained results, that choosing university is not significantly related to the degree of socialization of the student towards his companions but can be traced, at least in this local context, mainly in the sphere of the individual family.

The influence of the overall school performance reported by the student is quite crucial: the chance to pursue results of university studies positively related to time spent in studying.

With regard to the propensity to enrolment in scientific faculties, a crucial role is played by the mathematical ability and inclination of the student. Probably, the phenomenon of the low rate of students in science, including statistics, is failure to connect to the vocation of the younger generation to a discipline, commonly considered impenetrable and difficult to learn. The attitude to mathematics is considered by many young people a "privilege of few" and, so, they don't spend much time on the study of this discipline.

These reasons explain the disturbing phenomenon of decrease in student enrolments in science, and confirm the need to promote policy interventions that are able to balance the number of registrations between science and humanities.

References

Arboretti, G. R., & Corain, L. (2009). *The NPC ranking method with application to the evaluation of educational services and quality of products*. Multivariate methods and models for evaluating public services, Rimini, May, 25-26.

Brunello, G., & Checchi, D. (2005). School quality and family background in Italy. *Economics of Education Review, 24*(5), 563-577.

Checchi, D., Bratti, M., & De Blasio, G. (2008). Does the expansion of higher education increases equality of educational opportunities? *Evidence from Italy, Labour, 22*, 53-88.

Checchi, D., Leonardi, M., & Fiorio, C. (2007). *Intergenerational persistence in educational attainment in Italy*. Mimeo.

Corain, L., & Salmaso, L. (2007). A nonparametric method for defining a global preference ranking of industrial products. *Journal of Applied Statistics, 34*(2), 203-216.

Corbetta, P., Gasperoni, G., & Pisati, M. (2001). *Statistica per la ricerca sociale*. Il Mulino, Bologna.

Kleinbaum, D., & Klein, M. (2002). *Logistic regression: A self-learning text* (2nd ed.). New York, N.Y.: Springer Verlag.

Pesarin, F. (2001). *Multivariate permutation test, with application in Biostatistic*. Chichester: John Wiley & Sons..

Pesarin, F., & Lago, A. (1998). *Non parametric combination of dependent rankings with application to the quality assessment of industrial products*. Metron.

Pesarin, F., & Salmaso, L. (2006). Permutation tests for univariate and multivariate ordered categorical data. *Austrian Journal of Statistics, 35*(2-3), 315-324.

Piccolo, D. (2010). *Statistica* (3rd ed.). Il Mulino, Bologna.

Soliani, L. (2004). *Manuale di statistica per la ricerca e la professione. statistica univariata e bivariata parametrica e non parametrica nelle discipline ambientali e biologiche*. Uninova.

Chinese Business Review
Volume 10, Number 8, August 2011
David Publishing Company
1840 Industrial Drive, Suite 160, Libertyville, IL 60048
Tel: 1-847-281-9862; Fax: 1-847-281-9855
http://www.davidpublishing.com
economists@davidpublishing.com

ISSN 1537-1506