CENTRO PER LA FORMAZIONE IN ECONOMIA E POLITICA DELLO SVILUPPO RURALE
UNIVERSITÀ DI NAPOLI FEDERICO II - DIPARTIMENTO DI SCIENZE STATISTICHE
UNIVERSITÀ DI SALERNO - DIPARTIMENTO DI SCIENZE ECONOMICHE E STATISTICHE

Quaderni di STATISTICA

VOLUME 9 - 2007

LIGUORI EDITORE

Volume 9, anno 2007

ISSN 1594-3739 (edizione a stampa)

Registrazione al n. 5264 del 6/12/2001 presso il Tribunale di Napoli
ISBN-13 **978 - 88 - 207 - 4209 - 6**

Direttore responsabile: Gennaro Piccolo

Prima edizione italiana Dicembre 2007
Finito di stampare in Italia nel mese di Dicembre 2007 da OGL - Napoli

Il C.F.E.P.S.R. si avvale per la stampa dei Quaderni di Statistica del contributo dell'Istituto Banco di Napoli - Fondazione.

La carta utilizzata per la stampa di questo volume è inalterabile, priva di acidi, a PH neutro, conforme alle norme UNI EN Iso 9706 ×, realizzata con materie prime fibrose vergini provenienti da piantagioni rinnovabili e prodotti ausiliari assolutamente naturali, non inquinanti e totalmente biodegradabili.

Indice

Quaderni di Statistica
Vol. 9, 2007

A performance indicator for multivariate data

Rosa Arboretti Giancristofaro
Dipartimento di Matematica, Università di Ferrara
E-mail: rosa.arboretti@unife.it

Stefano Bonnini
Dipartimento di Scienze Statistiche, Università di Padova
E-mail: bonnini@stat.unipd.it

Luigi Salmaso
Dipartimento di Tecnica e Gestione dei Sistemi Industriali, Università di Padova
E-mail: salmaso@gest.unipd.it

Summary: In this paper we propose a composite indicator function for performance analysis. Statistical variables represent the satisfaction levels of users or consumers. When evaluation takes into account more than one aspect the problem is quite complicated and some methodological and practical issues arise: standardization, multivariate structure of data, accuracy of partial indicators, distance with respect to target (highest satisfaction level), stratification in presence of confounding factors. The methodological solution here proposed overcomes these problems and gives an easy and useful instrument for performance analysis.

Keywords: Nonparametric Combination, Ordered Categorical Variable, Performance Evaluation.

1. Introduction

Performance analysis is a fundamental activity for every organization. When a public or private company tries to improve efficiency and effectiveness, it has to cope with the problem of measuring processes and results and obtaining global evaluations starting from a multiplicity of partial aspects.

Indicators used to compare the performances of different units should be uniformly calculated with respect to the units both from the methodological point of view (Merlini, 2001; Vitali and Merlini, 1999) and also with respect to the definition of questions and items when data are collected through a questionnaire (i.e. satisfaction level about a service). Along with performance indicators it is often useful to consider targets that represent goal-results to which the performances should be compared. For example, in order to evaluate effectiveness of political decisions in labour market it is useful to establish a target for the unemployment rate and to evaluate the performance "calculating" the "degree of achievement" of the established target. In order to be appealing, targets should be achievable and defined in a rational way. Her Majesty's Treasury *et al.* (2001) proposed the definition of SMART target, where SMART means "clever" but it is also the acronym of "Specific, Measurable, Achievable, Relevant and Timed" which are fundamental "properties" for a good target.

Usually performance evaluation can be obtained from a synthesis of different indexes. For example, the Italian National Committee for University System Evaluation (CNVSU) proposed a set of indicators to evaluate Italian Universities and classified those indicators in four categories, using quality control methods generally applied to evaluate industrial processes (M.U.R.S.T, 1998). Some years ago Censis Institute started to elaborate and published the annual quality ranking of Public Italian Universities and Faculties, typical example of "multivariate ranking". The quality evaluation for the Faculties is conducted considering five macro-dimensions, which correspond to five families of indexes. Other typical examples of ranking-based performance evaluations considering several aspects of the same phenomenon are the National Education rankings, elaborated by OECD (Organisation for Economic Co-operation and Development), which compare member countries from the educational point of view, taking into account some aspects of the educational system like the instructional level, public investments in educational system, etc. Censis and OECD do not compare performances with target values, hence the evaluations are effectively less reliable and appealing than they should be.

Furthermore, when stratification methods are applied, in presence of confounding factors, we obtain partial strata results (simple indicators) and global results that sum up all the available information (composite indicators). Obviously, in this synthesis, partial indicators concerning single dimensions or single strata should be weighted according to the their importance.

Another crucial aspect of performance evaluation based on rankings is highlighted by Bird *et al.* (2005). They underline that the rank, i.e. the position of a unit in a ranking, represents a relative datum. In fact, being ranked lowest or highest does not immediately equate with genuinely worst or best performance. Hence, a relative evaluation like this is often not very useful.

In this paper, we propose a composite performance indicator, which is useful when performance evaluations (based on several aspects) measure the satisfaction level of users or evaluators. This indicator allows us to solve some of the main methodological and computational problems above described.

In section 2, we propose a methodological solution to solve the problem of the "relativity of rankings" by keeping the operational usefulness of the indicators and taking into account the targets. This solution is based on the concept of *extreme satisfaction profiles*. In section 3, we describe the method of combination of dependent rankings, which is useful to deal with the multivariate nature of the problem by reducing the dimensionality of the original problem. Section 4 includes some considerations about the consequences of the application of the proposed index on the distribution of data. In section 5 an application related to a survey on the PhD students at University of Ferrara is illustrated. Section 6 is dedicated to the conclusions.

2. Performance indicators and estreme satisfaction profiles

Let us consider a set of k informative variables $Y_1, Y_2 ..., Y_k$, that represent the satisfaction level of users or evaluators concerning k different aspects of quality. Let us suppose that s persons give a quality (or performance) evaluation about those k different aspects. Each variable Y_i (i=1,…,k) is numeric and it can take m_i distinct positive integer values $v_{i1}, v_{i2}, ..., v_{im_i}$ with $v_{i1} < v_{i2} < ... < v_{im_i}$, $m_i \in \mathrm{N} \setminus \{0\}$. Without loss of generality, let us suppose that greater values correspond to higher satisfaction levels. The proposed method can be applied to informative variables like those described above. This restriction is not too strict since probably the majority of the application cases can be reported to above settings. The methodological solution, we are going to describe, can be applied when quality evaluation is measured by means of ordinal categorical variables or numeric (continuous or discrete) variables, provided that, for each of the k considered aspects, it is possible and reasonable to transform data obtaining a numeric response variable Y_i $(i = 1,...,k)$. Clearly, the method is applicable also when the original k variable types (i.e. before the transformation) are mixed (for example some variable are categorical and other ones are numeric). Let us indicate with n, the number of evaluated units (persons, organizations, offices, etc.). The possible types of data are:

1. the *i-th* aspect is evaluated by means of satisfaction judgements, hence the original data are ordered categorical and, in order to obtain Y_i, a set of positive integer values must be attributed to the categories keeping the original order of the modalities;
2. the measure of quality or performance, for the *i-th* aspect, is given by a preference ranking of the n evaluated units, hence Y_i indicates the position in the ranking (i.e. the rank) and it can assume integer values between 1 and n, with n corresponding to the best unit and 1 assigned to the worst unit;
3. quality or performance, considering the *i-th* aspect, is evaluated through a numeric index on a continuous scale, hence it is possible to obtain a ranking of the n units according to (2);
4. evaluation consists in a vote, i.e. an integer number included in a given range, and it is not necessary any data transformation to obtain Y_i.

A weight w_i, such that $0 < w_i \leq 1$, is assigned to each variable Y_i. The set of weights reflect the different degrees of importance of the aspects considered in the analysis, i.e. of the informative variables. Weights are given by experts according to technical or managerial reasons or they come from the results of previous similar surveys.

The aim of the methodological problem we are handling is to detect a global satisfaction index or a global ranking of the *n* units, starting from the *k* informative variables, having *s* observations (evaluations of the *s* evaluators) for each variable, taking into account the dependence structure of the *k* variables. In particular we are focussing to two aspects:

5. the synthesis of the information to reduce the data dimensionality;
6. the criticism of Bird *et al.* (2005), who underline that rank is is a relative datum and extreme positions in ranking do not immediately equate with genuinely worst or best performances.

In order to answer the first question, we propose an extension of the nonparametric combination (NPC) method for combining dependent rankings proposed in Arboretti *et al.* (2005) and described in section 3. The second issue can be faced through the definition of *extreme satisfaction profiles*. In section 2.1 this concept is defined and in Section 2.2 its role in the calculation of the performance indicator is described.

2.1. Extreme satisfaction profiles

The *extreme satisfaction profiles* are theoretical frequency distributions of response variables Y_i, $i = 1,...,k$, a priori defined. Specifically, they are hypothetical distributions corresponding to maximum satisfaction or minimum satisfaction of the *s* evaluators. We distinguish between *strong satisfaction profiles* and *weak satisfaction profiles*.

We define *strong satisfaction profiles* in the following way:

- maximum satisfaction profile is obtained when the s users give the best evaluation for each of the k considered aspects. Hence, let us indicate with f_{ih} the proportion of evaluators who, for the *i-th* aspect, give the *h-th* judgement or score, i.e. the relative frequency of v_{ih} for variable Y_i. Formally:

$$\text{maximum satisfaction} \Rightarrow f_{ih} = \begin{cases} 1 & if \quad h = m_i \\ 0 & if \quad h < m_i \end{cases} \qquad \forall i,\ i=1,...,k \qquad (1)$$

- minimum satisfaction profile is obtained when the s users give the worst evaluation for each of the k considered aspects. Formally, we can write:

$$\text{minimum satisfaction} \Rightarrow f_{ih} = \begin{cases} 1 & if \quad h = 1 \\ 0 & if \quad h > 1 \end{cases} \qquad \forall i,\ i=1,...,k \qquad (2)$$

  where f_{ih} indicates the relative frequency of v_{ih} for variable Y_i.

The *strong satisfaction profiles* correspond to degenerate distributions for each of the k marginal variables Y_i. Practically it is not always possible to observe a performance evaluation like in (1) and (2). For example, with respect to questionnaire-based evaluations of university lectures, it is not much realistic to think that 100% of students may give the best judgement for every aspect or vice versa that 100% of students may give the worst judgement for every aspect. In such a case, even considering the motivational role of performance monitoring, it is convenient to define more realistic satisfaction profiles that represent achievable goals and compare them with the observed performances. To this aim we introduce the *weak satisfaction profile* as follows:

- maximum satisfaction for the *i-th* aspect is obtained when the proportion of subjects who choose the highest judgement is equal to u_i, where $u_i \in (0,1)$. Hence the target percentage of people completely satisfied is less than 100% and can vary according to the considered aspect (for example 70% for Y_1, 65% for Y_2, etc.). In a formal way, global maximum satisfaction is obtained when

$$\begin{cases} f_{ih} = u_i & \text{if} \quad h = m_i \\ \sum_{h=1}^{m_i-1} f_{ih} = 1 - u_i & \text{and} \quad 0 < u_i < 1 \end{cases} \qquad \forall i,\ i=1,\dots,k \qquad (3)$$

- minimum satisfaction, for the *i-th* aspect, is obtained when the proportion of subjects who choose the lowest judgement is equal to l_i, where $l_i \in (0,1)$. Hence, in a formal way, global minimum satisfaction (or maximum dissatisfaction) is obtained when

$$\begin{cases} f_{ih} = l_i & \text{if} \quad h = 1 \\ \sum_{h=2}^{m_i} f_{ih} = 1 - l_i & \text{and} \quad 0 < l_i < 1 \end{cases} \qquad \forall i,\ i=1,\dots,k \qquad (4).$$

We point out that, when u_i and l_i represent realistic achievable targets, they can be fixed observing past experience, for example, by taking the highest percentage of people completely (barely) satisfied observed in the past. When this information is not available we can put $u_i = 1$ and/or $l_i = 1$. If motivational and ambitious targets are needed, they can be fixed by managers and/or organizers in the strategic and business planning.

2.2. Score transformation

In order to embody the *extreme satisfaction profiles* in the analysis, we need a data transformation on values v_{ih}, $h = 1,\dots,m_i$, $i = 1,\dots,k$. Let us suppose that, for response variable Y_i, the last t_i modalities correspond to satisfaction judgements and the other $m_i - t_i$ correspond to dissatisfaction judgements. The aim consists in the transformation of the satisfaction levels according to the distribution of Y_i, $i = 1,...,k$, respect to the set of *s* evaluators, so that scores or ranks v_{ih} reflect the "real" satisfaction level of the set of evaluators. The satisfaction level grows with the proportion of people that give a positive judgement and obviously it decreases with every increase of the proportion of people who give a negative judgement. Often, to compare two or more units (university courses or faculties, offices, organizations, etc.), an index like mean, median, or similar is calculated for each unit and used for the comparison. The idea of our approach consists in making each evaluation score (vote, rank or numeric transformation of a category) less arbitrary, taking into account the proportion of people who choose that score. A score corresponding to a satisfaction judgement should be increased as much as the number of evaluators choosing it is high, while for a dissatisfaction judgement the increase should be a decreasing function of the number of people choosing it or it should be decreased proportionally to the number of users choosing it. In this way, it is possible to apply several transformations. We describe two of this transformation:

(a) We separate satisfaction judgements from dissatisfaction judgements and, for every variable Y_i, we transform data according to the following rule (*asymmetrical transformation*):

$$\begin{cases} v'_{ih} = v_{ih} + f_{ih} \cdot 0{,}5 & \text{if} \quad m_i - t_i + 1 \le h \le m_i \\ v'_{ih} = v_{ih} + (1 - f_{ih}) \cdot 0{,}5 & \text{if} \quad 1 \le h \le m_i - t_i \end{cases} \tag{5}$$

Hence, scores correspondent to positive judgements are increased proportionally to relative frequencies while the increase of scores correspondent to negative judgements is negatively correlated to relative frequencies. This is equivalent to assign additive degrees of importance to the original values. These degrees of importance are function of f_{ih} frequencies. The maximum of the increase for a single modality is equal to 0,5. For example if $v_{ih} = h$, *h=1,2,3,4*, and the first two values correspond to dissatisfaction while the other two values correspond to satisfaction, then the lowest score becomes 1,5 if nobody chooses the worst judgement (i.e. f_{i1} = 0) and this does not change if all evaluators give the worst judgement (i.e. f_{i1} = 1). If a proportion, equal to f_{i1}, of evaluators, strictly included in (0,1), declare the lowest satisfaction level, then the score is transformed in a value greater than 1 and less than 1,5, as high as f_{i1} is little. Similarly, value 4 corresponds to maximum satisfaction and it becomes 4,5 if all the evaluators choose that judgement or a value between 4 and 4,5 if the proportion of evaluators who are completely satisfied is less than 1. The transformed value grows with f_{i4}.

(b) Similarly to point (a) we separate satisfaction judgements from dissatisfaction judgements and, for every variable Y_i, we transform data according to the following rule (*symmetrical transformation*):

$$\begin{cases} v'_{ih} = v_{ih} + f_{ih} \cdot 0{,}5 & \text{if} \quad m_i - t_i + 1 \le h \le m_i \\ v'_{ih} = v_{ih} - f_{ih} \cdot 0{,}5 & \text{if} \quad 1 \le h \le m_i - t_i \end{cases} \tag{6}$$

In this case, scores correspondent to satisfaction levels are increased as in point (a) while scores corresponding to dissatisfaction judgements are similarly decreased proportionally to relative frequencies. In the previous example, score 1 becomes 0,5 if all the evaluators choose the maximum dissatisfaction judgement; it remains unchanged if no evaluator chooses this judgement; it decreases to a value between 0,5 and 1 otherwise, according to f_{i1}. The idea is to “reinforce” the highest scores, which represent satisfaction, and to lower the scores representing dissatisfaction. The magnitude of the score variation directly depends on the relative frequencies f_{ih}.

The second transformation rule, as described in section 4, causes an increase in score variability because it accentuates the distance between satisfaction and dissatisfaction scores.

Applying transformation (5) or (6) on Y_i, z_{ji} indicates the observed value of Z_i on the *j-th* unit with $j=1,2,...,n$, where Z_i is the transformed variable. In order to facilitate interpretation of indexes and to incorporate the extreme satisfaction profiles, a further transformation of

z_{ji} can be applied to obtain λ_{ji} scores included in (0,1). The following properties must be satisfied by λ_{ji} scores:

i. the relation between λ_{ji} and z_{ji} must be monotonic increasing, λ_{ji} tends to 0 in presence of minimum satisfaction and it tends to 1 in presence of maximum satisfaction;
ii. λ_{ji} must assume values greater than 0 and less than 1.

The second property avoids computational problems like null denominators, logarithm arguments equal to zero etc., when the synthesis with NPC method is applied.

There are different ways to transform z_{ji} into λ_{ji}. Here we propose three of them which are useful to compare the absolute performance with the maximum satisfaction profile, with the minimum satisfaction profile or with both of them. Let us indicate with z_{imax} the maximum observed value for Z_i according to the extreme satisfaction profile with $i=1,2,...,k$, then the proposed transformations are:

a. comparison with the highest satisfaction level:

$$\lambda_{ji} = \frac{z_{ji} + 0,5}{z_{i\max} + 1} \tag{7}$$

b. comparison with the lowest satisfaction level:

$$\lambda_{ji} = 1 - \frac{z_{i\min} + 0,5}{z_{ji} + 1} = \frac{z_{ji} - z_{i\min} + 0,5}{z_{ji} + 1} \tag{8}$$

c. comparison with the highest and the lowest satisfaction level:

$$\lambda_{ji} = \frac{z_{ji} - z_{i\min} + 0,5}{z_{i\max} - z_{i\min} + 1}. \tag{9}$$

If we put $v_{ih}=h$, with h=1,2,3,4 and $u_i = l_i = 1$, according to the combination of the transformations, we obtain the following indicators (Table 1). If we put u_i = 0,7 then z_{imax} is equal to 4+0,7·0,5 = 4,35 with obvious consequences about the values of λ_{ji}.

Table 1. Performance indicators for different score tranformations

Transformation $y \to z$	$z_{i\,min}$	$z_{i\,max}$	transformation $z \to \lambda$		
			a	b	c
(a) asymmetrical	1	4,5	$\lambda_{ji} = \frac{z_{ji}+0{,}5}{5{,}5}$	$\lambda_{ji} = \frac{z_{ji}-0{,}5}{z_{ji}+1}$	$\lambda_{ji} = \frac{z_{ji}-0{,}5}{4{,}5}$
(b) symmetrical	0,5	4,5	$\lambda_{ji} = \frac{z_{ji}+0{,}5}{5{,}5}$	$\lambda_{ji} = \frac{z_{ji}}{z_{ji}+1}$	$\lambda_{ji} = \frac{z_{ji}}{5}$

From the point of view of interpretation, transformation (a), described by (7), consists in a normalization of values respect to z_{imax}, hence, it allows to have values between 0 and 1 measuring the "degree of closeness" to the best performance instead of the original absolute z_{ji} values; transformation (b), described by (8), similarly allows us to evaluate the "degree of distance" from the worst performance; finally transformation (c), described by (9) compare the difference between the observed performance and the worst performance with the highest attainable value (range).

Transformations of z_{ji} values into λ_{ji} scores, make the data interpretation easier as described, preserve the order with respect to the units and make the performance comparable with respect to the variables by eliminating the scale effect. We observe that transformation (a) and (c) are linear and induce a variability reduction and a location effect on data distribution. Transformation (b) modifies the distribution shape and change the original correlation structure between variables. In particular, it emphasizes little differences between the worst performances and it is sensitive to little changes of z_{imin}. Furthermore, a change in the worst performance has a greater impact on the lowest scores than on the highest ones. Hence, outliers produce asymmetric effects on the distribution so that, in our opinion, such transformation is less preferable than the others.

3. Synthesis of partial indicators through nonparametric combination

In order to make a synthesis of the information provided by the k Y_i variables through the λ_{ji} scores, $j=1,2,...,n,$ $i=1,2,...,k$, we propose the application of the NPC method (Arboretti *et al.* 2005), that allows to face the multivariate nature of data using a combining function to reduce data dimensionality. Section 3.1 is dedicated to the description of the method and Section 3.2 contains an extension of the method.

3.1. The NPC method

The NPC method (Arboretti *et al.* 2005) is based on the detection of a combining function $\psi : \Re^{2k} \to \Re$, whose arguments are the λ_{ji} scores and the w_i weighs $=1,2,...,n,$ $i=1,2,...,k$, that allows to reduce data dimensionality, implicitly considering the dependence structure of marginal variables $Y_1, Y_2 ..., Y_k$, without modelling it.

In order to calculate the synthesis score the following formula is applied:

$$T_j = \psi(\lambda_{j1},...,\lambda_{jk}; w_1,...,w_k).$$

Combining function ψ must satisfy the following minimal and easy to check properties (Lago and Pesarin, 2000):

i. ψ must be continuous in all its $2k$ arguments, i.e. little changes in some subset of arguments imply little changes in ψ;

ii. ψ must be non decreasing function of all the λ_{ji} scores, i.e. $\psi(\ldots,\lambda_{ji},\ldots;w_1,\ldots,w_k) \geq \psi(\ldots,\lambda'_{ji},\ldots;w_1,\ldots,w_k)$ when $1 > \lambda_{ji} > \lambda'_{ji} > 0$ for every $i \in \{1,2,\ldots,k\}$;

iii. ψ must be symmetric respect to argument permutations, i.e. $\psi(\lambda_{ju_1},\ldots,\lambda_{ju_k};w_{u_1},\ldots,w_{u_k}) = \psi(\lambda_1,\ldots,\lambda_k;w_1,\ldots,w_k)$ where $u_1, u_2,\ldots,u_k$ is a permutation of 1,2,...,k.

The described properties are satisfied by several combining functions. Some of them are:

a. Fisher's combining function: $T_F = -\sum_{i=1}^{k} w_i \cdot \log(1-\lambda_i)$;

b. Liptak combining function: $T_L = \sum_{i=1}^{k} w_i \cdot \Phi^{-1}(\lambda_i)$, where Φ represents the normal cumulative distribution function;

c. Logistic combining function: $T_{Log} = \sum_{i=1}^{k} w_i \cdot \log[\lambda_i/(1-\lambda_i)]$;

d. Tippett combining function: $T_T = \max_i(w_i\lambda_i)$;

e. Additive combining function: $T_A = \sum_{i=1}^{k} w_i \cdot \lambda_i$.

Finally, the following transformation which normalizes the global index in the interval [0,1], is useful to facilitate data interpretation:

$$S_j = \frac{T_j - T_{\min}}{T_{\max} - T_{\min}}, j=1,2,\ldots,n,$$

where

$T_{\min} = \psi(\lambda_{1\min},\ldots,\lambda_{k\min};w_1,\ldots,w_k)$, $T_{\max} = \psi(\lambda_{1\max},\ldots,\lambda_{k\max};w_1,\ldots,w_k)$ and $\lambda_{i\min}$, $\lambda_{i\max}$, i=1,...,k, are obtained applying the extreme satisfaction profiles to (7), (8) e (9), i.e. calculating λ_{ji} with $z_{i\min}$ and $z_{i\max}$ respectively, instead of using z_{ji}.

3.2. An extension of NPC

The nonparametric nature of the NPC method, i.e. the fact that it is not necessary to describe in a formal way the dependence among the *k* informative variables, is one of its most important features. In the methodological solutions proposed in the literature, the calculation of composite indicators are often obtained through nonlinear transformations of the informative variables.

An example is given by the "quality of life ranking" of Italian provinces calculated by "Il Sole 24 ore" (Cadeo, 2003). Some of the marginal variables, on which the *Quality of Life* index is based, are transformed in a nonlinear way and consequently it is not possible to express in analytic way their relation with the synthetic index. Hence the use of an additive combining function like the arithmetic mean, according to the proposal of "Il Sole 24 ore" is generally not a good solution as stated in Aiello and Attanasio, 2004.

Moreover, a linear transformation preserves the original shape of distribution with negative consequences described in Pagnotta (2003), when data are asymmetric. In these cases, this is usually the rank transformation. Attanasio and Capursi (1997) and Terzi and Moroni (2004) deal with the consequences of different (linear and nonlinear) transformations in the calculation of a composite indicator. The importance and the difficulties connected to the construction of a multivariate ordering of a set of *n* units in performance analysis problems are highlighted in D'Esposito and Ragozini (2004).

These works allow us to emphasize the good properties of our method, that makes it possible both using linear or nonlinear transformations, and applying different kinds of combining functions. In order to underline the generality and flexibility of our method, being inspired by the rule proposed by David Firth in Fayers and Hand (2002), we introduce a more general formula for the combining function ψ. By using this formulation, it is possible to derive several different combining functions satisfying properties described in section 3.1. In order to calculate a composite index q starting from a set of partial indicators x_i, taking values in (0,1), and weighs β_i, Blalock (1982) suggests the use of the following rule:

$$\log(1-q) = \sum_i \beta_i \log(1-x_i).$$

This is equivalent to:

$$q = 1 - \prod_i (1-x_i)^{\beta_i}.$$

A generalization of it (Fayers and Hand, 2002) is given by:

$$\log\{1 - q[1 - \exp(-\gamma)]\} = \sum_i \beta_i \log\{1 - x_i[1 - \exp(-\gamma)]\},$$

where γ is a non negative parameter.

Similarly, we propose a general rule for the combining function of the NPC method:

$$T = -[1-\exp(-\gamma)]^{-1} \cdot \sum_{i=1}^{k} w_i \log\{1-\lambda_i[1-\exp(-\gamma)]\}, \tag{10}$$

where γ>0. It is evident that if $\gamma \to \infty$ formula (10) corresponds to Fisher's combining function. Moreover, when $\gamma \to 0$ formula (10) is equivalent to the additive rule. In fact, by replacing $\delta = [1-\exp(-\gamma)]$ in (10), we obtain

$$\lim_{\gamma \to 0}\left\{-[1-\exp(-\gamma)]^{-1} \cdot \sum_{i=1}^{k} w_i \log\{1-\lambda_i[1-\exp(-\gamma)]\}\right\} =$$

$$= \lim_{\delta \to 0}\left\{-\frac{\sum_i w_i \log[1-\lambda_i \delta]}{\delta}\right\} =$$

$$= \sum_i w_i \cdot \lambda_i .$$

When γ moves in $(0,\infty)$, from formula (10) it is possible to obtain different combining functions satisfying the properties described in section 3.1, hence suitable for the application of the NPC method.

4. Some considerations about the distribution of transformed variables

Let us consider the $Y \rightarrow Z$ transformations introduced in section 2.2 and described by (5) and (6). The so called *asymmetric* transformation described by (5) can be written as follows

$$z_{ji} = y_{ji} + \sum_{h=1}^{m_i - t_i} I_h(y_{ji}) \cdot (1 - f_{ih}) \cdot 0{,}5 + \sum_{h=m_i-t_i+1}^{m_i} I_h(y_{ji}) \cdot f_{ih} \cdot 0{,}5$$

where $I_h(y_{ji})$ is equal to 1 if $y_{ji} = v_{ih}$, i.e.

$$I_h(y_{ji}) = \begin{cases} 1 & if \quad y_{ji} = v_{ih} \\ 0 & if \quad y_{ji} \neq v_{ih} \end{cases}$$

It is worth observing the consequences induced on the distribution by the transformation. An evident consequence of asymmetrical transformation is the right shifting of distribution, i.e. the increase of the first moment. Let us indicate the first moment with $M_1(\cdot)$ and consider

$$M_1(Z_i) = \sum_{h=1}^{m_i} v'_{ih} f_{ih} =$$

$$= \sum_{h=1}^{m_i-t_i} [v_{ih} + (1 - f_{ih}) \cdot 0{,}5] \cdot f_{ih} + \sum_{h=m_i-t_i+1}^{m_i} [v_{ih} + f_{ih} \cdot 0{,}5] \cdot f_{ih} =$$

$$= \sum_{h=1}^{m_i} v_{ih} f_{ih} + 0{,}5 \cdot \left[\sum_{h=1}^{m_i-t_i} (1 - f_{ih}) \cdot f_{ih} \cdot + \sum_{h=m_i-t_i+1}^{m_i} f_{ih}^{\;2} \right] =$$

$$= M_1(Y_i) + 0{,}5 \cdot \left[\sum_{h=1}^{m_i-t_i} f_{ih} - \sum_{h=1}^{m_i-t_i} f_{ih}^2 + \sum_{h=m_i-t_i+1}^{m_i} f_{ih}^{\;2} \right].$$

Even if we consider the second moment we observe an increase induced by the transformation. Let us indicate the second moment with $M_2(\cdot)$ and consider

$$M_2(Z_i) = \sum_{h=1}^{m_i} (v'_{ih})^2 f_{ih} =$$

$$= \sum_{h=1}^{m_i - t_i} [v_{ih} + (1 - f_{ih}) \cdot 0{,}5]^2 \cdot f_{ih} + \sum_{h=m_i - t_i + 1}^{m_i} [v_{ih} + f_{ih} \cdot 0{,}5]^2 \cdot f_{ih} =$$

$$= \sum_{h=1}^{m_i - t_i} v_{ih}^2 f_{ih} + 0{,}25 \sum_{h=1}^{m_i - t_i} (1 - f_{ih})^2 f_{ih} + \sum_{h=1}^{m_i - t_i} v_{ih} (1 - f_{ih}) \ f_{ih} +$$

$$+ \sum_{h=m_i - t_i + 1}^{m_i} v_{ih}^2 f_{ih} + 0{,}25 \sum_{h=m_i - t_i + 1}^{m_i} f_{ih}^3 + \sum_{h=m_i - t_i + 1}^{m_i} v_{ih} f_{ih}^2 =$$

$$= M_2(Y_i) + \left[\sum_{h=1}^{m_i - t_i} (0{,}25 + v_{ih}) f_{ih} - \sum_{h=1}^{m_i - t_i} (0{,}5 + v_{ih}) f_{ih}^2 + \right.$$

$$\left. + \sum_{h=m_i - t_i + 1}^{m_i} v_{ih} f_{ih}^2 + 0{,}25 \sum_{h=1}^{m_i} f_{ih}^3 \right]$$

Concerning the so called *symmetric transformation,* described by (6), we have

$$z_{ji} = y_{ji} - \sum_{h=1}^{m_i - t_i} I_h(y_{ji}) \cdot f_{ih} \cdot 0{,}5 + \sum_{h=m_i - t_i + 1}^{m_i} I_h(y_{ji}) \cdot f_{ih} \cdot 0{,}5 .$$

In this case, the effect of the transformation on the first and second moment depend on the distribution of frequencies. When frequencies are concentrated on the highest scores we have an increase of the moment values, otherwise we have a decrease of the moment values. For the first moment we have

$$M_1(Z_i) = \sum_{h=1}^{m_i} v'_{ih}\, f_{ih} =$$

$$= \sum_{h=1}^{m_i - t_i} [v_{ih} - f_{ih} \cdot 0{,}5] \cdot f_{ih} + \sum_{h=m_i-t_i+1}^{m_i} [v_{ih} + f_{ih} \cdot 0{,}5] \cdot f_{ih} =$$

$$= \sum_{h=1}^{m_i} v_{ih} f_{ih} + 0{,}5 \cdot \left[\sum_{h=m_i-t_i+1}^{m_i} f_{ih}{}^2 - \sum_{h=1}^{m_i-t_i} \left(f_{ih}^2 \right) \right] =$$

$$= M_1(Y_i) + 0{,}5 \cdot \left[\sum_{h=m_i-t_i+1}^{m_i} f_{ih}{}^2 - \sum_{h=1}^{m_i-t_i} f_{ih}{}^2 \right].$$

For the second moment we have

$$M_2(Z_i) = \sum_{h=1}^{m_i} (v'_{ih})^2 f_{ih} =$$

$$= \sum_{h=1}^{m_i - t_i} [v_{ih} - f_{ih} \cdot 0{,}5]^2 \cdot f_{ih} + \sum_{h=m_i-t_i+1}^{m_i} [v_{ih} + f_{ih} \cdot 0{,}5]^2 \cdot f_{ih} =$$

$$= \sum_{h=1}^{m_i - t_i} \left[v_{ih}^2 f_{ih} + 0{,}25 f_{ih}^3 - v_{ih} f_{ih}^2 \right] + \sum_{h=m_i-t_i+1}^{m_i} \left[v_{ih}^2 f_{ih} + 0{,}25 f_{ih}^3 + v_{ih} f_{ih}^2 \right] =$$

$$= M_2(Y_i) + 0{,}25 \cdot \sum_{h=1}^{m_i} f_{ih}^3 + \sum_{h=m_i-t_i+1}^{m_i} v_{ih} f_{ih}{}^2 - \sum_{h=1}^{m_i-t_i} v_{ih} f_{ih}{}^2.$$

In Figure 1 and Figure 2 it is shown the graphical evaluation of the two types of $Y \rightarrow Z$ transformation about location, variability and shape of data distribution. The considered example refers to an ordered categorical variable with six modalities. We have supposed three possible probability configurations.

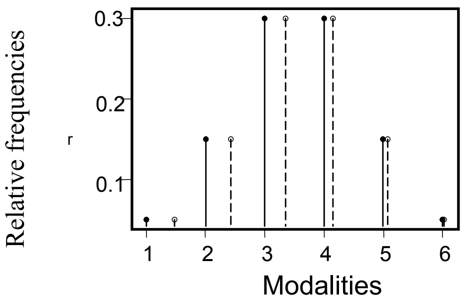
The asymmetric transformation does never imply a decrease of scores, hence the distribution moves on the right and the mean value grows; data variability tends to decrease. If the distribution is not symmetric, the asymmetry of scores decreases.

The symmetric transformation does not take changes of the first moment if the original distribution is symmetric too.

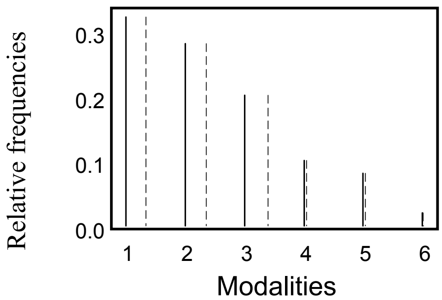
In presence of distributional asymmetry, if frequencies are concentrated on high values, the arithmetic mean increases, otherwise it decreases. The consequences on variability and symmetry of distribution are opposite with respect to the other type of transformation: both data dispersion and asymmetry (where present) increase[1].

[1] In Figure 1 and 2 we use the following notation: MR (or VR) indicates the relative reduction of mean (or std.dev.) of Z respect to Y; Asy(.) indicates the difference between mean and median.

SQ(Y) = 1.20
SQ(Z) = 1.07
V.R. = 0.11
$M_1(Y)$ = 3.50
$M_1(Z)$ = 3.75
M.R. = -0.07
Asy(Y) = 0.00
Asy(Z) = 0.00



SQ(Y) = 1.34
SQ(Z) = 1.25
V.R. = 0.07
$M_1(Y)$ = 2.40
$M_1(Z)$ = 2.70
M.R. = -0.12
Asy(Y) = 0.40
Asy(Z) = 0.33



SQ(Y) = 1.34
SQ(Z) = 1.25
V.R. = 0.07
$M_1(Y)$ = 4.60
$M_1(Z)$ = 4.80
M.R. = -0.04
Asy(Y) = -0.40
Asy(Z) = -0.34

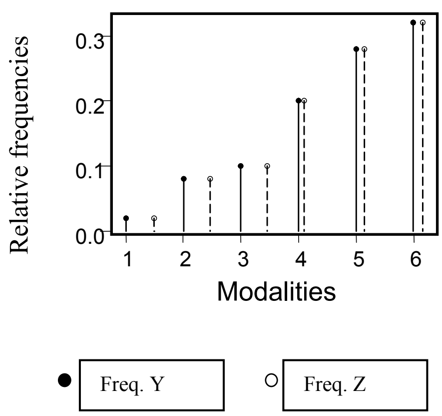


Figure 1. Distributions of an ordered categorical variable with six modalities and asymmetrical transformation

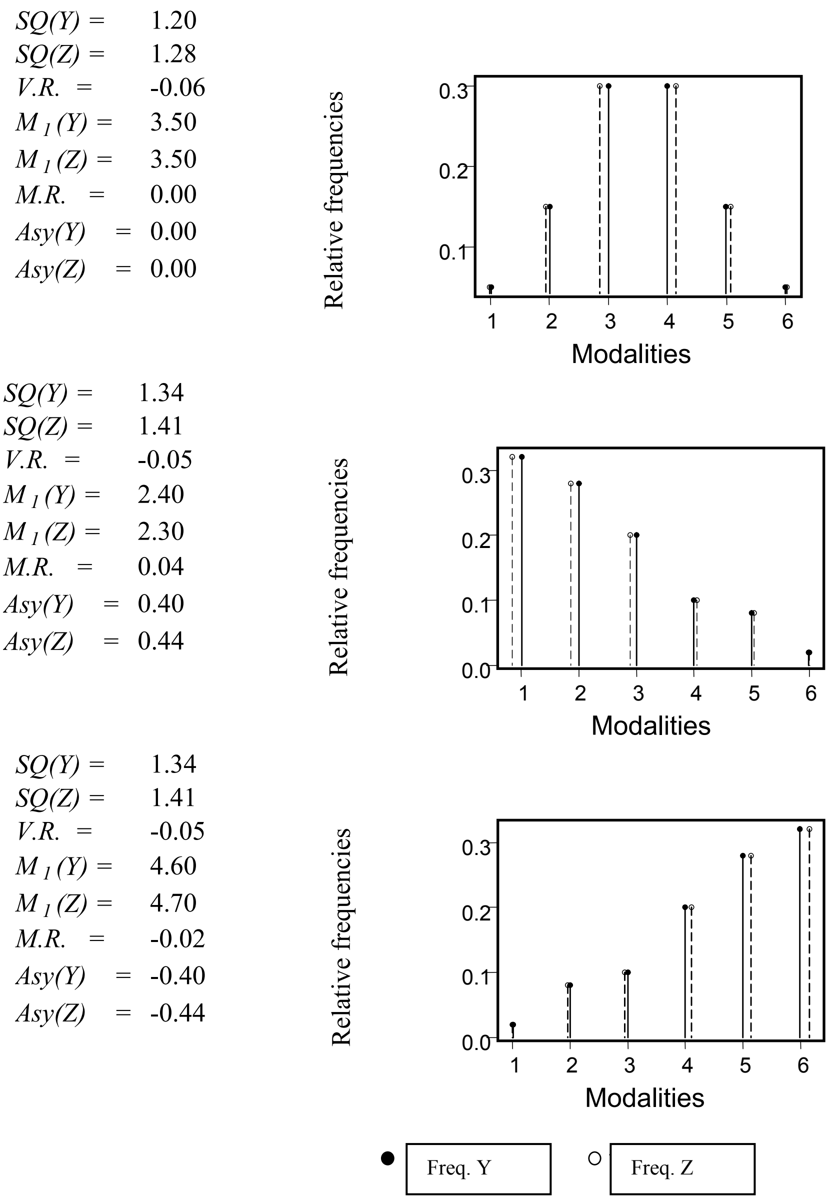


Figure 2. Distributions of an ordered categorical variable with six modalities and symmetrical transformation

5. Post-Doc survey at University of Ferrara

In 2004 a CATI survey on a sample of Post-Docs of University of Ferrara was carried out (http://www.unife.it/comstat/documenti). A sample of 120 Post-Docs was randomly extracted from a population of 288 persons and interviewed with the CATI method. We show some results of the survey separately for each PhD area. We distinguish three areas: Economic-Legal area (EL), Medical-Biological area (MB) and Scientific-Technological area (ST). Post-Docs was asked to indicate the satisfaction level about some aspects of the PhD program. We classify these aspects into two categories:

1. *Relation between teaching and job*, which is composed by three variables:
   a. coherence between teaching and employment;
   b. use at work of the abilities acquired during the PhD studies;
   c. pertinence of training received during PhD studies with respect to the tasks assigned at work.
2. *Job opportunities*, which is composed by three variables:
   a. Academic job opportunities;
   b. Labour market opportunities;
   c. Openess towards scientific community.

The interviewers asked the Post-Docs to express their judgements about each of the considered aspects, choosing between four possible ordered categories (unsatisfied, not very satisfied, quite satisfied, very satisfied). Every ordered category was associated to an integer increasing score ranged from 1 to 4. Using the asymmetric score transformation, the strong satisfaction profile and Fisher's combining function, we calculated a global satisfaction index for each of the two categories. Graphs in Figure 3 show the global satisfaction index distributions for the three groups.

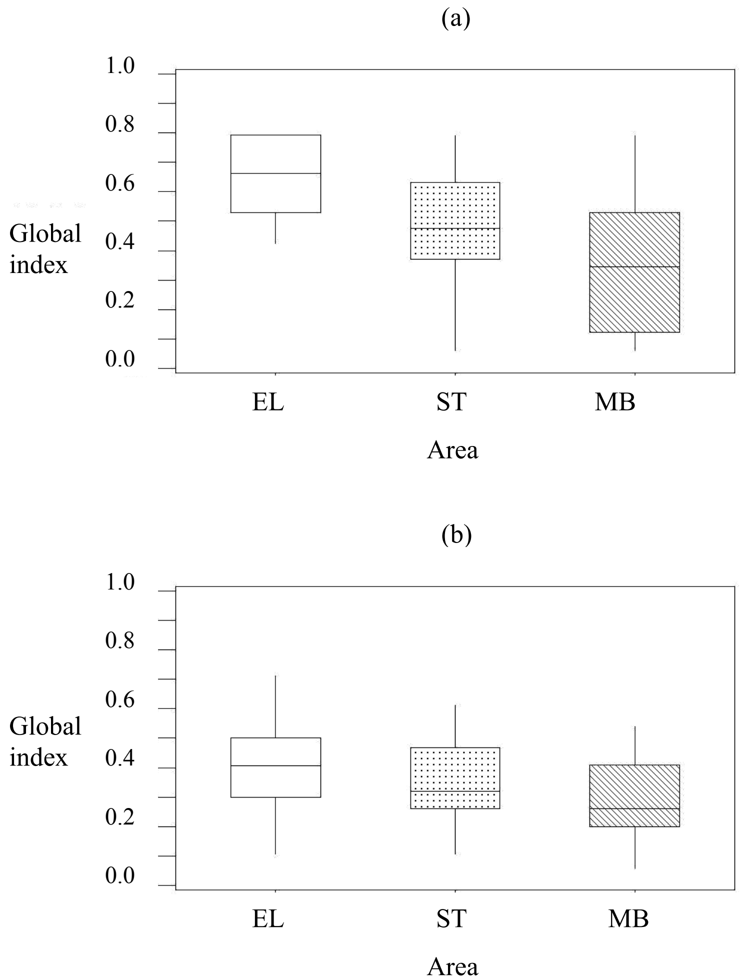


Figure 3. Box-Plot of global satisfaction index for relation between teaching and job (a) and job opportunities (b).

For each possible pairwise area comparison, simultaneous confidence intervals were calculated using Bonferroni correction (Bland and Altman, 1995) for multiplicity. With a 95% confidence level, the confidence intervals for the difference between the mean scores for *teaching-job relation*, i.e. for $\mu_{aEL} - \mu_{aST}$, $\mu_{aEL} - \mu_{aMB}$ e $\mu_{aST} - \mu_{aMB}$ are [0,0624;0,3007], [0,1720;0,4504] e [-0,0303;0,2895] respectively. For *job opportunities* the confidence intervals for $\mu_{aEL} - \mu_{aST}$, $\mu_{aEL} - \mu_{aMB}$ e $\mu_{aST} - \mu_{aMB}$ are [0,0624;0,3007], [0,1720;0,4504] e [-0,0303;0,2895] respectively.

Post-Docs from the Economic-Legal area tend to be less critical than the colleagues coming from the other areas, without intention to give a judgement about the real quality of PhD courses. The proposed extension of the non parametric combination of dependent ranking method allows us to evaluate the distance between the observed satisfaction levels and the levels corresponding to the maximum/minimum satisfaction (according to the extreme satisfaction profiles) which could be target values for the organizers of PhD courses.

6. Conclusions

In this paper we presented a method to construct a composite performance indicator, useful when the performance evaluations (about more than one aspect) measure the satisfaction level of a set of users or evaluators.

The proposed solution allows to overcome some methodological problems recently pointed out in the literature:

- The data standardization respect to units and variables;
- The synthesis without loss of information, in particular taking into account the information about the dependence of variables;
- The need to express the real users' satisfaction level, through absolute evaluations and not relative ones like in the case of ranks;
- The inclusion of the different degrees of importance of each single considered aspects;
- The appealing aim of indicators, taking into account the degree of achievement of fixed goals;

- The necessity to carry out stratified analysis in presence of confounding factors;
- The easy-to-interpret feature of proposed global and partial indeces.

We do not think this method is the definitive solution to the highlighted problems but we think it contains some useful ideas like the use of NPC methodology, the extreme satisfaction profile concept, some score transformations, etc. in order to obtain more useful and reliable indicators under given conditions and in some application problems. A SAS macro performing the above proposed method is reported in the Appendix.

Appendix: ***SAS MACRO NPC Ranking for ordered variables***

```
/*****************************************************************/
/****************** SAS MACRO NPC Ranking for ordered variables ********/
/*****************************************************************//*
/*          SINTAX: %macro graduatoria (dataset,cod,w,k,m,t,list);
```

LIST OF MACRO PARAMETERS:

dataset	=	SAS dataset's name;
cod	=	name of variable identifing statistical units;
w	=	list of weights for the variables (weights must sum to the number *k* of variables);
k	=	number of ordered variables;
m	=	number of ordered variables' values representing ordered discrete scores, $h = 1,...,m, m \in \mathbf{N} \setminus \{0\}$, with the value 1 corresponding to lower satisfaction, *m* to higher satisfaction ;
t	=	number of last values from $h = 1,...,m,$ corresponding to satisfaction's judgements
u	=	relative frequency (expressed by % without decimals, e.g. 60) of subjects with value *m* in the extreme satisfactory profile
l	=	relative frequency (expressed by % without decimals, e.g. 70) of subjects with value 1 in the extreme satisfactory profile
list	=	list of names of ordered variables.

MACRO OUTPUT:

The macro generates the SAS temporary dataset named Fisher containing the variable ‘y_def’ representing the combined ϕ -index varying from 0 to 1.

Example:
%*graduatoria*(dataset=elab,cod=matricola,w=1 1,k=2,m=4,t=2,u=60,l=70,list=v1 v2);

```
*******************************************************************/

%macro graduatoria(dataset,cod,w,k,m,t,u,l,list);

proc freq noprint;
        %do i=1 %to &k;
                tables %scan(&list,&i) /out=freq&i;
        %end;
run;

%do i=1 %to &k;
        proc sort data=&dataset;
                by %scan(&list,&i);
        run;
        proc sort data=freq&i;
                by %scan(&list,&i);
        run;
        data f&i(drop=count percent);
                merge &dataset freq&i;
                by %scan(&list,&i);
                freq&i=percent/100;
        run;
        proc sort;
                by &cod;
        run;
%end;
data f_tot;
        merge %do i=1 %to &k; f&i %end;;
        by &cod;
run;
data lambda;
        set f_tot;
        h=&m-&t;
        array x(&k) &list;
```

```
        array f(&k) %do i=1 %to &k;freq&i %end;;
        array new(&k) z1-z&k;
        do i=1 to &k;
                if x(i)<=h then
                new(i)=((1-f(i))*0.5)+x(i);
                else new(i)=(f(i)*0.5)+x(i);
        end;
        array neww(&k) l1-l&k;
        do i=1 to &k;
                neww(i)=(new(i)-(1+0.5*(1-(&l/100)))+0.5)/((&m+0.5*(&u/100))-
(1+0.5*(1-(&l/100)))+1);
        end;
run;
data fisher;
        set lambda;
        array provv(&k) l1-l&k;
        %do i=1 %to &k;
                ww&i=%scan(&w,&i);
                w&i=ww&i/&k;
        %end;
        array neww(&k) b1-b&k;
        array minim(&k) min1-min&k;
        array maxim(&k) max1-max&k;
        array pesi(&k) w1-w&k;
        do i=1 to &k;
                neww(i)=-pesi(i)*log(1-provv(i));
                minim(i)=-pesi(i)*log(1-(0.5/((&m+0.5*(&u/100))-(1+0.5*(1-
(&l/100)))+1)));
                maxim(i)=-pesi(i)*log(1-(((&m+0.5*(&u/100))-(1+0.5*(1-
(&l/100)))+0.5)/((&m+0.5*(&u/100))-(1+0.5*(1-(&l/100)))+1)));
        end;
        y=%do i=1 %to %eval(&k-1); b&i+ %end; b&k;
        ymin=%do i=1 %to %eval(&k-1); min&i+ %end; min&k;
        ymax=%do i=1 %to %eval(&k-1); max&i+ %end; max&k;
        y_def=(y-ymin)/(ymax-ymin);
run;
%mend;
```

References

Aiello F., Attanasio M. (2004), How to transform a batch of simple indicators to make up a unique one?, *Atti della XLII riunione scientifica della Società Italiana di Statistica*, Cleup, Padova, 327-338.

Arboretti Giancristofaro R., Pesarin F., Salmaso L. (2005), Non parametric approaches for multivariate testing with mixed variables and for ranking on ordered categorical variables with an application to the evaluation of PhD programs, in S. Sawilowsky (ed.): *Real Data Analysis*, American Educational Research Association, Age Publishing, 331-357.

Attanasio M., Capursi V. (1997), Graduatorie sulla qualità della vita: prime analisi di sensibilità delle tecniche adottate, *Rivista Italiana di Economia Demografia e Statistica*, LI, 331-342.

Bird S.M., Cox D., Farewell V.T., Goldstain H., Holt T., Smith P.C. (2005), Performance indicators: good, bad and ugly, *Journal of the Royal Statistical Society*, 168, 1-27.

Blalock H.M. (1982), *Conceptualization and Measurement in the Social Sciences*, Sage, Beverly Hills.

Bland J.M., Altman D.G. (1995), Multiple significance tests: the Bonferroni method, *Biometrical Journal*, 3, 310-370.

Cadeo R. (a cura di), Dossier sulla qualità della vita 2003, *Il Sole 24 ore*, 22 dicembre 2003.

D'Esposito M.R., Ragozini G. (2004), Multivariate Ordering in Performance Analysis, *Atti della XLII riunione scientifica della Società Italiana di Statistica*, Cleup, Padova, 51-54.

Fayers P.M., Hand D.J. (2002), Causal variables, indicator variables and measurement scales: an example from quality of life, *Journal of the Royal Statistical Society A*, 165, 233-261.

Her Majesty's Treasury, Cabinet Office, National Audit Office, Audit Commission and Office for National Statistics (2001), *Choosing the Right Framework: a Framework for Performance Informantion,* Stationery Office, London, available on-line at http://www.hm-treasury.gov.uk/documents/ public_spending_and_services/ pss_index.cfm.

Lago A., Pesarin F. (2000), Nonparametric combination of dependent rankings with application to the quality assessment of industrial products, *Metron*, LVIII, 39-52.

Merlini A. (2001), La qualità della vita: le funzioni di ordinamento, *Rivista Italiana di Demografia e Statistica*, LV, 2, 119-159.

Ministero dell'Università e della Ricerca Scientifica e Tecnologica. Osservatorio per la valutazione del sistema universitario (1998), *Indicazioni* *per la preparazione delle relazioni dei Nuclei di valutazione interna e insieme minimo di indicatori,* on-line at http://www.cnvsu.it/publidoc/comitato/.

Pagnotta S.M. (2003), Una generalizzazione dei ranghi per standardizzare i dati, *Quaderni di Statistica*, 5, 49-64.

Terzi S., Moroni L. (2004), Graduatorie della qualità della vita e loro sensibilità al pre-trattamento delle variabili che la definiscono: alcune critiche al dossier de "Il Sole 24 ore", *Quaderni di Statistica*, 6, 105-128.

Vitali O., Merlini A. (1999), La qualità della vita: metodi e verifiche, *Rivista Italiana di Demografia e Statistica*, LIII, 2, 5-92.

Quaderni di Statistica
Vol. 9, 2007

A general approach for modelling individual choices

Domenico Piccolo
Dipartimento di Scienze Statistiche, Università di Napoli Federico II
E-mail: domenico.piccolo@unina.it

Summary: In this article, we generalize the approach for deriving the Inverse Hyper-Geometric (IHG) random variable proposed by Ridout (1999). After reviewing some preliminary results, the paper focuses on a general relationship between a statement on the sequential odds of the choices and the probability distribution of preferences. Finally, some statistical consequences of these findings are discussed.

Keywords: Odds and probability, IHG random variable, Preferences distributions.

1. Introduction

The statistical approach to ordinal data is mainly based on Generalized Linear Models (GLM) proposed by McCullagh (1980) and discussed by Agresti (2002). In that context, for model specification, a relationship among log-odds of cumulative probability and a linear function of covariates is assumed.

From a different viewpoint, some models have been introduced in literature order to explain the behaviour of respondents when faced to multiple choices. In this vein, among others, D'Elia and Piccolo (2005) and Piccolo and D'Elia (2007) proposed a new class of models, MUB and CUB models, respectively, which proved to be useful in several fields of applications.

When ranks data are characterized by a unique mode, the Inverse Hy-