

New developments in ranks data modelling with covariates¹

Nuovi sviluppi nella modellistica per l'analisi dei ranghi con covariate

Angela D'Elia

Dipartimento di Scienze Statistiche
Università degli Studi di Napoli Federico II
Via L. Rodinò 22, 80138 Napoli
angela.delia@unina.it

Riassunto: In questo lavoro si presentano alcuni sviluppi metodologici relativi ai modelli per variabili rango, con inclusione di covariate. In effetti, l'analisi statistica di graduatorie di preferenze e/o di valutazione è utile in numerosi contesti, nei quali è importante tener conto anche delle variabili che influiscono sulla formulazione delle graduatorie stesse. In tale ottica, l'articolo illustra un modello statistico mediante il quale è possibile stimare l'impatto delle variabili esplicative sia sul livello atteso di gradimento (e/o di valutazione) che sull'incertezza che accompagna la formulazione del giudizio in termini di assegnazione di rango.

Key words: Ranks data; Statistical models; Preferences and Evaluation analyses.

1. Introduction

There are several different contexts where people are asked to express their judgements, by means of ranks data. Among these, we can highlight two main areas, that are the surveys, respectively: *i*) on the preferences of a group of raters towards a set of mutually exclusive items; *ii*) on the evaluations with respect to a set of complementary but not-alternative items.

The first case concerns situations where people express their liking or disliking feeling towards m objects, yielding a ranking of these objects: this process causes a permutation of the first m integers, since usually no ties are allowed. This is typical of marketing analyses, but also of political and social choices science surveys and, more generally, of psychological and behavioural studies (staff selections, policy alternatives assessment, etc.). We'll refer to these fields as *Preferences data analysis*.

On the other hand, there exist situations where people express, by means of ranks, their evaluation about the perceived level of some latent variable: for instance, the satisfaction level with respect to a given facility (customers and users' satisfaction), the pain intensity suffered for a given pathology (medical studies), the agreement rate towards a given opinion (political polls), the measure of pupils' gained knowledge and skills (by means of examination grades, in educational researches), etc. In all these cases, the raters do not yield a ranking of m items (since they are, usually,

¹ This research was financially supported by Dipartimento di Scienze Statistiche, Università degli Studi di Napoli Federico II and by funds of L.R. 5/2002. In addition, the Author utilized research and scientific facilities provided by Centro per la Formazione in Economia e Politica dello Sviluppo Rurale, Portici.

complementary), but for each of them they choose a level of evaluation among m possible ones. We'll refer to these fields as *Evaluations data analysis*.

Both *preferences* and *evaluations* fields share several aspects and, as a consequence, the methodological issues raised by their statistical analysis may be faced by following a unique approach for both of them, introducing minor modifications for the specific considered cases, as we are going to show in this paper.

In particular, with respect to the literature about these issues, we present a new analysis setting, based on a model that preserves the discrete nature of the ranks, and explicitly links them to the raters' covariates.

The paper is organised as follows. In Section 2, we explicit some preliminary settings for the ranks data analysis; then, in Section 3 we discuss the psychological mechanism that is assumed to underlie the elicitation of ranks. On these bases, in Sections 4 and 5, we illustrate a statistical model for ranks data with covariates, stressing the issues related to the parameters estimation and their meanings. Some empirical evidence are discussed in Section 6; finally, further considerations about these topics are developed in Section 7.

2. Main settings about ranks data analysis

Henceforth, in the case of *preferences data analysis*, we refer to this kind of setting: a group of n raters are asked to express a ranking among m alternatives items ($O_1, \dots, O_j, \dots, O_m$), assigning –with no ties– a rank $r \in [1, m]$ to each of them (Fligner and Verducci, 1993; Marden, 1995). Of course, the observed ranks $\mathbf{r} = (r_1, r_2, \dots, r_n)'$ for a given item O_j are related in some way to the ranks assigned to the other $(m - 1)$ objects; however, we follow a univariate approach, considering only *one* item at a time. This stems from two considerations: 1) there are several real situations where the research interest concerns the liking or disliking feeling towards *one* specific item (D'Elia, 2003); 2) both in the statistical and psychometric literature about preferences data, it is generally accepted the hypothesis of *independence of irrelevant alternatives* (Arrow, 1951; Luce, 1959): that is, the odds that item h is preferred to item j do not depend on the ranks given to the other items.

On the other hand, *evaluations data* derive from raters who choose for every item/issue a level of satisfaction/agreement/feeling among m possible levels. In this case, the rank r assigned to a given item does not express the position in a preference ranking, but the level of some feeling. Thus, in the evaluation context we can properly consider the observed ranks $\mathbf{r}$ for a single issue in a univariate setting: indeed, strictly speaking, we do not have a ranking procedure (yielding a permutation), but just a choice among m different evaluation levels.

As far as it concerns the meanings of the assigned ranks, we assume that $r=1$ means “most preferred” or “maximum level of something”, for the *preferences* or *evaluations data*, respectively; while $r=m$ means “least preferred” or “minimum level”. Thus, for instance, in a customers' satisfaction context, an observed rank equal to 1 would mean that there is maximum satisfaction with respect the given item. However, this assumption (which naturally refers to a preference ranking) does not contradict the practice of the evaluations surveys, where generally the higher the rank the greater the satisfaction, since in this case we can reverse the observed ranks by means of the relation: $r^*=m-r+1$. Specifically, the model we are going to develop in Section 4 has got such a reversibility property.

A part from the differences described above, it should be quite clear that the statistical analyses (support definition, probability distribution, inferential issues, etc.) for both the problems (*preferences and evaluations analyses*) can be driven and faced in a common framework, if we refer the observed results to the ranks elicitation mechanism.

3. The mechanism of ranks elicitation

The procedure by means of which a rater assigns a rank to a given item entails, of course, an elicitation strategy, either referring to a preference order or being induced from an evaluation measurement.

The elicitation can be performed in different ways, as it has so far been pointed out in the psychometric literature (starting from the seminal papers by Thurstone, 1927, and Daniels, 1950, to Critchlow *et al.*, 1991). In this field, particular attention has been devoted to the so called *criterion of paired comparisons* (Bradley and Terry, 1952), assuming that the rank given to an item comes from the comparisons of each item with all the others. This means that in formulating a preference ranking of m items, each rank results from $m(m-1)/2$ comparisons among all the possible items pairs.

In fact, we can more realistically assume that this strategy is not performed with respect to all the items, but in order to assign the ranks to a subset of them, which are neither strongly liked nor strongly disliked². Indeed, especially for large m , it is quite natural to figure out that, mostly in the middle of the ranking, the elicitation mechanism is based on a paired comparisons criterion.

In the case of evaluations, the observed ranks can be thought of as results of a paired comparison strategy, too: in fact, in this context, the rater does not compare items, but levels of satisfaction (for instance) and then chooses the one nearest to his/her perception and/or judgement.

Of course, both *preferences* and *evaluations* formulations are, in some extent, affected by the uncertainty, as it happens in most individual behaviours and choices. This is especially true in this context, since we deal –by discrete tools (the ranks)– with feelings (preferences expression, satisfaction or agreement level, pain assessment, etc.) that are intrinsically continuous, but do not admit a direct measurement (as we do, for instance, with temperature or with weight): this fact determines a sort of fuzziness in both the ranking and evaluation procedures.

Then, in order to take into account the latter consideration, we can assume that the ranks elicitation is a composite procedure, made up of two different elements. In particular, they can be described as follows:

- 1) comparison, by means of which the rater assesses his/her preference order or his/her evaluation level;
- 2) uncertainty, that makes fuzzy and not completely settled the results of the comparison.

With respect to these components, it is important to stress that: *i*) the comparison takes place among the items in a *preference analysis* context, whilst the comparison concerns the levels in an *evaluation analysis* setting; *ii*) despite of the presence of the uncertainty, we generally observe discrete realisations (the ranks): this means that it becomes

² In the case of a strong feeling towards a given item, a more natural strategy is to follow the sequential comparisons criterion (D'Elia, 2003).

necessary to develop a statistical framework in order to make explicit the weight of the uncertainty component in a discrete model and, then, to allow its estimation.

In particular, from the above considerations, it emerges that a proper structure for describing and representing the ranks elicitation mechanism (in both *preferences* and *evaluations* contexts) should rely on a mixture distribution of the two components (comparison and uncertainty), as we are going to illustrate in the next Section.

4. A model for ranks data

Since ranks data are generally used for expressing continuous and, often, fuzzy feelings, most of the statistical literature about this topics has developed models for rankings based on underlying continuous random variables: e.g. the Gamma distribution (Henery, 1983; Stern, 1990), and the Normal and multivariate Normal distributions (Dansie, 1986; Yu, 2000). In fact, this kind of assumption also occurs when discrete choice models have been specified (like logistic regression): in the latter case, indeed, a continuous random term is usually introduced to model choice probabilities, by means – often – of the Gumbel distribution (McFadden, 1973; Chapman and Staelin, 1982), or of a mixture of Gumbel and Normal components (Skrondal and Rabe-Hesketh, 2003).

Differently from this practice, we think that a useful modelling strategy for ranks data can be based on a discrete probabilistic model, that takes into account both the discrete nature of the observed realizations, and the composite nature of the ranks elicitation mechanism, following the lines developed in the previous Section.

Thus, let r be the rank assigned to a given item/issue in a *preference/evaluation* context: in both the cases, we assume that the rank is the expression of a choice performed either among m positions in a (*preference*) ranking, or among m (*evaluation*) levels.

Following D’Elia and Piccolo (2003), we consider r as a realization of a mixture of a discrete Uniform and a shifted Binomial distributions (henceforth, called MUB, that is: $R \sim \text{MUB}(m, \pi, \xi)$, if:

$$\Pr(R = r) = \pi P_B(r) + (1 - \pi) P_U(r), \quad r = 1, 2, \dots, m,$$

where P_B and P_U represent the probability mass functions of the shifted Binomial and of the Uniform random variable, respectively:

$$P_B(r) = \binom{m-1}{r-1} (1 - \xi)^{r-1} \xi^{m-r}; \quad P_U(r) = \frac{1}{m}; \quad r = 1, 2, \dots, m;$$

and the parameters $\xi \in [0, 1]$; $\pi \in [0, 1]$.

The two components of the mixture have been introduced in order to adequately represent the corresponding components of the ranks elicitation mechanism. Indeed, the shifted Binomial random variable describes the observed ranks as resulting from a paired comparisons selection strategy (D’Elia, 2000); on the other hand, the discrete Uniform distribution plays the role of the uncertainty component, maximizing the entropy on the finite support $\{1, \dots, m\}$ (Papoulis, 1984, pp. 514-515). Of course, the relevance of this component is inversely related to the value of the π parameter, varying from maximum uncertainty in assigning ranks ($\pi=0$) to a perfect paired comparisons scheme ($\pi=1$).

Thus, by means of the MUB model, it is possible to deal with ranks, preserving both their extrinsic discrete nature and their origin from a continuous and fuzzy process.

An important feature of the MUB model is that –changing the parameters’ values– its shape is adequate for fitting several different empirical distributions (asymmetrical, skewed, flat, etc.); moreover, it allows to deal with cases where both the body and the tails of the distribution are heavy.

In order to better understand the meaning of the parameters π and ξ , it is useful to consider the expectation and the variance of R , that are, respectively:

$$E(R) = \pi(m-1)\left(\frac{1}{2} - \xi\right) + \frac{(m+1)}{2};$$

$$Var(R) = (m-1)\left\{\pi\xi(1-\xi) + (1-\pi)\frac{m+1}{12} + \pi(1-\pi)\frac{(m-1)(2\xi-1)^2}{4}\right\}.$$

As far as it concerns the expectation, it emerges that $E(R)$ decreases when the ξ parameter tend to 1, keeping fixed π , while the variance decreases with increasing values of π , keeping fixed ξ . Since we assumed that lower values of the ranks mean greater preference or greater evaluation, we are able to get the following interpretations: ξ is a measure of *preference* or of a positive *evaluation*, while $(1-\pi)$ is a measure of uncertainty in both the settings.

Of course, from a joint point of view, it results that when both the parameters values increase there is a strong feeling (great preference or satisfaction, for instance) towards the considered item/issue, whose rank is assigned with low uncertainty. On the other hand, it is more difficult to jointly interpret the roles of the parameters when both tend to zero. The following scheme depicts some extreme situations.

MUB parameters	<i>Preference data</i>	<i>Evaluation data</i>
$\xi \rightarrow 1; \pi \rightarrow 1$	maximum liking and minimum uncertainty in formulating rankings	maximum satisfaction/agreement/pain and minimum uncertainty in choosing levels
$\xi \rightarrow 0; \pi \rightarrow 1$	minimum liking and minimum uncertainty in formulating rankings	minimum satisfaction/agreement/pain and minimum uncertainty in choosing levels
$\pi \rightarrow 0$ $\forall \xi \in [0,1]$	equi-preference feeling	indifference feeling

It is worth to notice that, since the MUB model is reversible (that is: if $R \sim \text{MUB}(m, \pi, \xi)$, then $(m-R+1) \sim \text{MUB}(m, \pi, 1-\xi)$), this property makes easy to adapt the above scheme to those situations where we may assume, especially in an *evaluation* context, that low ranks are associated with minimum feeling (satisfaction, agreement, etc.).

Given this probabilistic framework, there are two possible ways to perform a *preferences/evaluations* data analysis, from an inferential point of view. Indeed, in first instance, we can rely on the observed ranks $\mathbf{r} = (r_1, r_2, \dots, r_n)'$ assigned to a given

item/issue by a group of n raters, and from this data we estimate both ξ and π (as in D'Elia and Piccolo, 2003). Then, if we get information on the raters' features too, we can develop models linking the expressed ranks to the individual covariates. In this paper we focus only on this second framework, in order to illustrate how to measure the impact of the covariates on both the components of the ranks elicitation process.

5. The role of covariates

In this Section, we are going to illustrate an extension of the MUB model, allowing in its specification the inclusion of raters' covariates: indeed, we explicitly link the ranks to a set of variables. In such a way, we can model the ranks elicitation process by relating its components (comparison and uncertainty) to a set of individual features of the raters. Since the two components are described by the parameters ξ and π , respectively, that means to introduce a link between the parameters and covariates, developing a framework by means of which we can explain the expressed ranks (both in a *preference* and in an *evaluation* setting) on the basis of specific characteristics of the raters (e.g. age, gender, educational level, income, etc.).

In fact, the inclusion of the covariates in the MUB model can be accomplished in two different ways.

In first instance, we can model a parameter at a time, as in D'Elia (2003), specifying two kinds of structures:

- a model (MUB_π) where the π parameter depends upon the raters' covariates, but ξ is kept fixed. This is the case when we are interested in modelling only the uncertainty component;
- a model (MUB_ξ) where the ξ parameter depends upon the raters' covariates, but π is kept fixed. This is the case when either we assume that the uncertainty component is not affected by covariates, or we are not interested in modelling it.

Then, letting $\mathbf{X}$ to be the $(n \times p+1)$ design matrix of raters' covariates, with generic row $\mathbf{x}_i = (1, x_{i1}, x_{i2}, \dots, x_{ip})'$ for $i = (1, 2, \dots, n)$, in order to have a mapping from $\square$ to the interval $[0, 1]$ we can apply the logit link function and get the following relations:

$$(\pi | \mathbf{X} = \mathbf{x}_i) = \frac{1}{1 + \exp(-\mathbf{x}_i \boldsymbol{\beta})}; \quad (\xi | \mathbf{X} = \mathbf{x}_i) = \frac{1}{1 + \exp(-\mathbf{x}_i \boldsymbol{\gamma})};$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)'$ and $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \dots, \gamma_p)'$ are the covariates coefficient vectors, respectively. Of course, the sets of covariates that are significant in the MUB_ξ and in the MUB_π models could be completely overlapping, partially overlapping, or completely disjoint.

In order to better assess the role of the covariates in the ranks elicitation procedure, it is useful to consider the following expressions, where we make explicit the impact of just one covariate –for sake of simplicity– on π for the MUB_π model, and on ξ for the MUB_ξ model, separately. In detail, letting $\mathbf{x}_i = (1, x_{i1})'$ and $\mathbf{u} = (0, 1)'$, we get:

$$\Delta_\pi = (\pi | \mathbf{X} = \mathbf{x}_i + \mathbf{u}) - (\pi | \mathbf{X} = \mathbf{x}_i) = \left[\frac{e^{-\beta_0 - \beta_1 x_{i1}} (1 - e^{-\beta_1})}{(1 + e^{-\beta_0 - \beta_1 x_{i1}})(1 + e^{-\beta_0 - \beta_1 x_{i1} - \beta_1})} \right];$$

$$\Delta_{\xi} = (\xi | \mathbf{X} = \mathbf{x}_i + \mathbf{u}) - (\xi | \mathbf{X} = \mathbf{x}_i) = \left[\frac{e^{-\gamma_0 - \gamma_1 x_{i1}} (1 - e^{-\gamma_1})}{(1 + e^{-\gamma_0 - \gamma_1 x_{i1}})(1 + e^{-\gamma_0 - \gamma_1 x_{i1} - \gamma_1})} \right].$$

Thus, in the MUB _{ξ} model it results:

$$E(R | \mathbf{X} = \mathbf{x}_i + \mathbf{u}) - E(R | \mathbf{X} = \mathbf{x}_i) = -\pi(m-1) \cdot \Delta_{\xi}.$$

Then, modelling separately the parameters π and ξ , we obtain the following scheme, by means of which we can assess the effect of a generic x_{ij} unitary increment (*ceteris paribus*) on the ranks elicitation procedure.

	Uncertainty		Preference	Evaluation
$\beta_j > 0$	decrease	$\gamma_j > 0$	improvement	increase
$\beta_j = 0$	no effect	$\gamma_j = 0$	no effect	no effect
$\beta_j < 0$	increase	$\gamma_j < 0$	worsening	decrease

In fact, whilst the uncertainty component depends only upon the π parameter (and then on Δ_{π} and the $\boldsymbol{\beta}$ coefficients vector), in the MUB model $E(R)$ depends on both the parameters ξ and π . This means that in assessing the impact of the covariates on the expected *preference* or *evaluation*, the above scheme is valid, for any value of γ_j , only when $\beta_j = 0$.

As a consequence, it seems important to develop a further extension of the MUB framework, allowing a specification where both the model's parameters are simultaneously related to the raters' covariates. Then, it is useful to introduce a different notation, in order to clearly distinguish the sets of covariates affecting π and ξ . In particular, we let $\mathbf{Z}$ to denote the subset of $(p_{\pi} + 1)$ variables related to π , and $\mathbf{W}$ the subset of $(p_{\xi} + 1)$ ones related to ξ , moreover, since $\mathbf{Z}$ and $\mathbf{W}$ may be overlapping or not, we let $\boldsymbol{\rho}_i$ to be the generic linear predictor for the i -th rater ($i = 1, 2, \dots, n$).

Then, we can specify a MUB model with inclusion of covariates for both the parameters $\boldsymbol{\theta} = (\xi, \pi)'$ as:

$$(\boldsymbol{\theta} | \boldsymbol{\rho}_i) = \left(\frac{1}{1 + \exp(-\mathbf{z}_i \boldsymbol{\beta})}, \frac{1}{1 + \exp(-\mathbf{w}_i \boldsymbol{\gamma})} \right).$$

Given this framework, the log-likelihood function results:

$$\log L(\boldsymbol{\theta}) = \sum_{i=1}^n \log \left\{ \frac{1}{1 + \exp(-\mathbf{z}_i \boldsymbol{\beta})} \left(\binom{m-1}{r_i-1} \frac{\exp(-\mathbf{w}_i \boldsymbol{\gamma})(r_i-1)}{(1 + \exp(-\mathbf{w}_i \boldsymbol{\gamma}))^{m-1}} \right) - \frac{1}{m} \right\}.$$

From a computational point of view, in order to get the Maximum Likelihood estimates of the coefficient vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we rely on the Expectation-Maximization (E-M)

algorithm, since it is considered one of the best tool for dealing with likelihood functions raising from mixture models (McLachlan and Peel, 2000).

This stems from the fact that in such models we can consider the observed data as incomplete, since the appropriate mixture component for each subject is unknown: for instance, in our case, we do not know the single rater's amount of uncertainty in expressing *preference* or *evaluation* ranks.

As a consequence, the complete data log-likelihood function factorises in two separate parts, on which two alternating steps of the E-M algorithm are performed³, until to the convergence, as follows (for any iteration $k = 0, 1, 2, \dots$):

- *E-step*: compute the conditional expectation of the complete data log-likelihood function, by means of the current parameters estimates:

$$\begin{aligned} Q(\boldsymbol{\beta}^{(k)}, \boldsymbol{\gamma}^{(k)}) &= Q_1(\boldsymbol{\beta}^{(k)}) + Q_2(\boldsymbol{\gamma}^{(k)}) = \\ &= - \sum_{i=1}^n \left\{ \log \left(1 + \exp(-\mathbf{z}_i \boldsymbol{\beta}^{(k)}) \right) + (1 - \tau_i^{(k)}) \mathbf{z}_i \boldsymbol{\beta}^{(k)} \right\} + \\ &\quad - \sum_{i=1}^n \tau_i^{(k)} \left\{ (r_i - 1) \mathbf{w}_i \boldsymbol{\gamma}^{(k)} + (m - 1) \log \left(1 + \exp(-\mathbf{w}_i \boldsymbol{\gamma}^{(k)}) \right) \right\}, \end{aligned}$$

where $\tau_i^{(k)}$ represent the estimated posterior probability that r_i belongs to the shifted Binomial component of the MUB model;

- *M-step*: maximize the $Q(\boldsymbol{\beta}^{(k)}, \boldsymbol{\gamma}^{(k)})$ function, in order to get updated estimates of the coefficients: $\boldsymbol{\beta}^{(k+1)}, \boldsymbol{\gamma}^{(k+1)}$.

Since the log-likelihood splits in two functions $Q_1(\boldsymbol{\beta}^{(k)})$ and $Q_2(\boldsymbol{\gamma}^{(k)})$, this allows easy derivations of the standard errors, as shown in D'Elia (2003) and Piccolo (2003).

6. Some empirical evidences

In order to better highlight the flexibility of the model with covariates we proposed in the previous Section, we show two different examples of analysis on really observed ranks data. The chosen results are a sample (only for illustrative purposes) from two wider case-studies, related to *Preference data analysis* (6.1) and *Evaluation data analysis* (6.2), respectively.

6.1 Preference data analysis

A survey has been conducted on 183 students attending the Faculty of Political Sciences (University of Naples Federico II), concerning their preferences towards future jobs. In particular, each rater was asked to express a ranking (with no ties) of $m = 14$ professions, supposed to be typical for this kind of Faculty.

Here we focus on the results obtained for the item "Political activity", whose average rank is 8.175, and the standard deviation is 4.509. For the ranks assigned to this item, two covariates were found significant: the Age for the π parameter, and the Gender (Male=0, Female=1) with respect to ξ (Table 1).

³ For more details about this point, see D'Elia (2003) and Piccolo (2003).

Table 1: Coefficient estimates for the item “Political activity”.

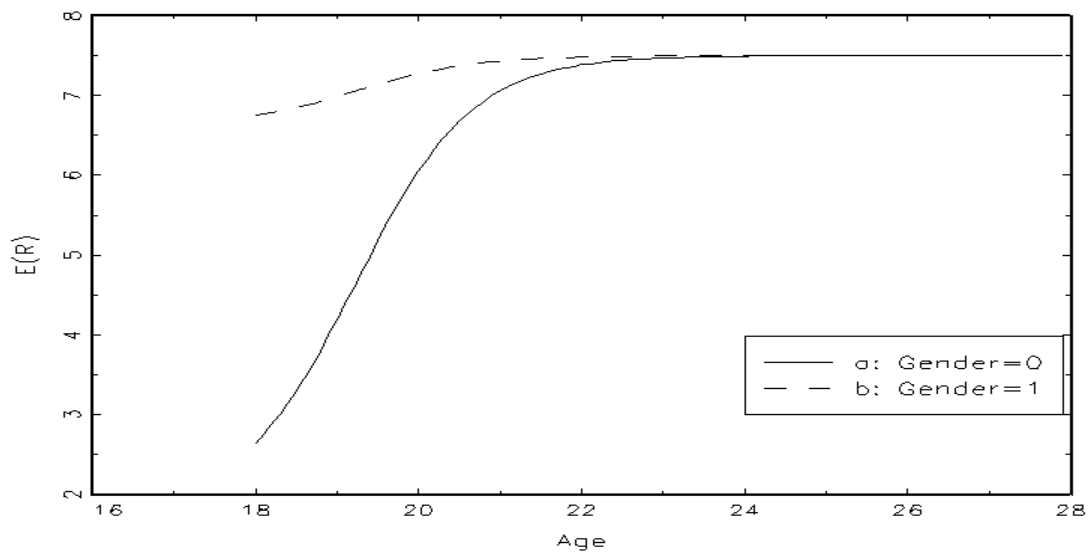
Covariates	Estimates	Standard errors
Constant ($\hat{\beta}_0$)	26.953	9.536
Age ($\hat{\beta}_1$)	-1.402	0.465
Constant ($\hat{\gamma}_0$)	2.753	0.372
Gender ($\hat{\gamma}_1$)	-2.479	0.425

In particular, it emerges that the Age has a negative impact on π , and thus older the age greater the uncertainty in assessing a rank for the considered item; on the other hand, being Female (Gender=1) negatively affects ξ and then makes higher the expected rank (*preference* worsening) being the Age fixed, as it can be seen from Table 2, where we show 4 different profiles of students (raters).

Table 2: Expected ranks for the item “Political activity”.

Age	Gender	$\hat{\pi}$	$\hat{\xi}$	E(R)
19	0	0.580	0.940	4.181
19	1	0.580	0.568	6.987
22	0	0.020	0.940	7.384
22	1	0.020	0.568	7.482

However, it is interesting to notice that the effect of the covariate Gender on the expected rank is stronger for younger raters, while it becomes lighter for older ones. Indeed, an increasing Age makes much more greater the uncertainty ($1 - \hat{\pi} = 0.980$) of the ranking process, and this fuzziness makes less neat the differences due to Gender (as it is confirmed by Figure 1).

**Figure 1:** Expected ranks for the item “Political activity” as functions of Age and Gender.

6.2 Evaluation data analysis

In recent years, great attention has been paid to the problem of *quality evaluation* measurement and, then, to the statistical analyses related to this kind of topic. In this field, all the issues concerning the assessment of quality in public services, such as educational institutions, health-care structures, etc. (Gori and Vittadini, 1999) have particular relevance, since it is not proper to evaluate their performances in profit terms. With respect to such contexts, indeed it is considered more important to take into account and to analyse the satisfaction of the users of the services, as it emerges by the *evaluations* questionnaires they compile. Of course, it seems also important to detect any possible relation between the expressed judgement and the rater's features, in order to better plan future changes for improving the service itself.

One of the main field where this kind of analysis has become relevant is the *evaluation* of the University system (AA.VV., 2002). At this regard, we discuss a limited subset of the results obtained in the *evaluation* survey conducted by the Università di Napoli Federico II during 2001.

For shortness, we limit ourselves to consider just one class, for which 137 students were asked to express their global satisfaction by means of a rank $r^* \in [1, 7]$. On these data, preliminarily we applied the transformation: $r = m - r^* + 1$.

The global satisfaction received an average rank equal to 4.703, while the standard deviation is 1.440. With respect to this issue, we found that the ξ parameter depends upon the covariate: Work (student is a worker: no=0; yes=1); instead, the uncertainty seems to not depend upon students' features (Table 3).

Table 3: *Coefficient estimates for the issue "Global satisfaction".*

Covariates	Estimates	Standard errors
Constant ($\hat{\gamma}_0$)	-0.005	0.276
Work ($\hat{\gamma}_1$)	-0.750	0.301

In particular, it emerges that the condition of Work negatively affects ξ , and thus yields a worsening in the expected *evaluation*. Indeed, considering the expected ranks conditioned to the two profiles (no-worker and worker), we can notice that $E(R) \cong 4$ for "full time" students, while $E(R) \cong 5$ for those who have a work activity, too (Table 4 and Figure 2). This evidence, if confirmed for other classes/Faculties, seems very interesting, since it points out the existence of some degree of dissatisfaction among worker-students, that raises the necessity of planning specific proposals towards this subset of University's users.

Table 4: *Expected ranks for the issue "Global satisfaction".*

Work	$\hat{\pi}$	$\hat{\xi}$	E(R)
0	0.818	0.499	4.006
1	0.818	0.320	4.884

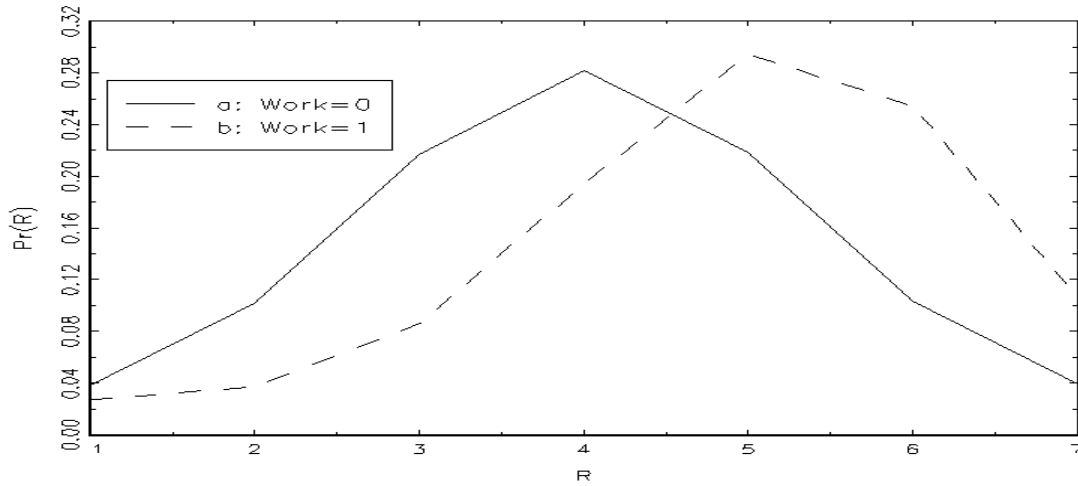


Figure 2: *Estimated rank probability distributions for workers and no-workers.*

7. Further considerations

In this paper, we have developed a common framework in order to perform statistical analyses for both *preferences* and *evaluations* data. The unifying aspect stems from taking into account the composite procedure underlying the elicitation of ranks: on this ground, we derived a mixture model of two discrete components, that allows the analysis of observed data, also in a setting where covariates are included. In this way, we get a tool that allows the estimation of the impact of the raters' covariates on the judgements they express, by means of ranks data.

Of course, this tool seems adequate for studying ranks data also in fields other than the *preferences* and *evaluations* ones; a first evidence, in this vein, has been obtained in the context of linguistic analyses where, considering the rankings of perceived synonymy with respect to a target word, Cappelli and D'Elia (2004) found significant relations between the educational and cultural features of the raters and the words they selected.

As far as it concerns methodological developments, there are several issues worth to be considered. In first instance, it would be interesting to explore the possibility of using different link functions in the model specification, in order to make even more explicit the role and the meaning of the covariates impact on the expected ranks. On the other hand, also the problem of an adequate model specification for multimodal empirical distributions should be considered.

From an inferential point of view, alternative parameters estimation methods (e.g. based on minimum distance criteria) could be taken into account. At the same time, computational issues (as the better choice of starting values for implementing the E-M algorithm) should be faced, too.

In addition to all these issues, a further modelling development seems interesting. Indeed, in several real situations, especially related with *preferences* contexts, we can assume that the elicitation mechanism of ranks occurs by means of a hierarchical procedure: the rater first detects k sets of items, that differ from each other on the basis of liking/disliking feeling; then, later, the rater assigns ranks to the items inside each set, by means of a fuzzy paired comparisons strategy, as the one modelled in this paper. In such a case, a hierarchical or multi-step model could be developed for a better fitting of ranks data.

References

- AA. VV. (2002) Valutazione della didattica e dei servizi nel sistema Università, *Atti della Giornata di Studio*, CUSL, Salerno.
- Arrow K. J. (1951) *Social Choice and Individual Values*, Wiley, New York.
- Bradley R. A., Terry M. B. (1952) The rank analysis of incomplete block designs. I. The method of paired comparisons, *Biometrika*, 39, 324–345.
- Cappelli C., D’Elia A. (2004) La percezione della sinonimia: un’analisi statistica mediante modelli per ranghi, *Le poids des mots – Actes de JADT04*, (Purnelle G., Fairon C. and Dister A. Eds.), Presses Universitaires de Louvain, Belgium, 129-140.
- Chapman R. G., Staelin R. (1982) Exploiting rank ordered choice set data within the stochastic utility model, *Journal of Marketing Research*, 14, 288-301.
- Critchlow D. E., Fligner M. A. and Verducci J. S. (1991) Probability models on ranking, *Journal of Mathematical Psychology*, 35, 294-318.
- Daniels H. E. (1950) Rank correlation and population models (with discussion), *Journal of the Royal Statistical Society B*, 12, 171-191.
- Dansie B. R. (1986) Normal order statistics as permutation probability models, *Applied statistics*, 35, 269-275.
- D’Elia A. (2000) Il meccanismo dei confronti appaiati nella modellistica per graduatorie: sviluppi statistici ed aspetti critici, *Quaderni di Statistica*, 2, 173-203.
- D’Elia A. (2003) Modelling ranks using the Inverse Hypergeometric distribution, *Statistical Modelling: an International Journal*, 3, 65-78.
- D’Elia A. (2003) A mixture model with covariates for ranks data: some inferential developments, *Quaderni di Statistica*, 5, 1-25.
- D’Elia A., Piccolo D. (2003) A mixture model for preference data analysis, *Preliminary Report*.
- Fligner M. A., Verducci J. S. (1993) *Probability models and statistical analyses for ranking data*, Lecture Notes in Statistics, Springer-Verlag, New York.
- Gori E., Vittadini G. (a cura di) (1999) *Qualità e valutazione nei servizi di pubblica utilità*, Etas Libri, Milano.
- Henery R. J. (1981) Permutation probabilities as models for horse races, *Journal of the Royal Statistical Society B*, 20, 822-834.
- Luce D. R. (1959) *Individual Choice Behavior*, Wiley, New York.
- Marden J. I. (1995) *Analyzing and Modeling Rank Data*, Chapman & Hall, London.
- McFadden D. (1973) Conditional logit analysis of qualitative choice behavior, in *Frontiers in Econometrics* (Zarembka P. ed.), Academic Press, New York, 105-142.
- McLachlan G., Peel D. (2000) *Finite Mixture Models*, J. Wiley & Sons, New York.
- Papoulis A. (1984) *Probability, Random Variables, and Stochastic Processes* (2nd edition), Mc Graw-Hill Int. Book Co., Singapore.
- Piccolo D. (2003) Computational issues in the E-M algorithm for ranks model estimation with covariates, *Quaderni di Statistica*, 5, 27-48.
- Skrondal A., Rabe-Hesketh S. (2003) Multilevel logistic regression for polytomous data and rankings, *Psychometrika*, 68, 267-287.
- Stern H. (1990) Models for distribution on permutations, *Journal of the American Statistical Association*, 85, 558-564.
- Thurstone L. (1927) A law of comparative judgement, *Psychol. Review*, 34, 273-286.
- Yu P. L. H. (2000) Bayesian analysis of order-statistic models for ranking data, *Psychometrika*, 65, 281-299.