A mixture model for preference data analysis. Comput Stat Data Anal

ARTICLE *in* COMPUTATIONAL STATISTICS & DATA ANALYSIS · JUNE 2005

Impact Factor: 1.4 · DOI: 10.1016/j.csda.2004.06.012

2 AUTHORS:

Angela D'Elia
European Commission

Domenico Piccolo
University of Naples Federico II

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Computational Statistics & Data Analysis 49 (2005) 917 – 934

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

A mixture model for preferences data analysis☆

Angela D'Elia*, Domenico Piccolo

Dipartimento di Scienze Statistiche, Università di Napoli Federico II, Via Leopoldo Rodinò 22, 80138 Napoli, Italy

Received 15 April 2003; received in revised form 14 June 2004; accepted 15 June 2004
Available online 8 July 2004

Abstract

A mixture model for preferences data, which adequately represents the composite nature of the elicitation mechanism in ranking processes, is proposed. Both probabilistic features of the mixture distribution and inferential and computational issues arising from the maximum likelihood parameters estimation are addressed. Moreover, empirical evidence from different data sets confirming the goodness of fit of the proposed model to many real preferences data is shown.

Keywords: Mixture model; Preferences data; Rankings

1. Introduction

Ranks data can be found in several situations: in particular, they are widely used in order to express the preferences/evaluations of a group of raters towards one or more items/services. In these settings many probabilistic models and statistical tools have been proposed and developed for describing the ranking process and/or analyzing ranks data (for a wide review, see Fligner and Verducci, 1993; Marden, 1995).

☆ This research has been supported by Dipartimento di Scienze Statistiche, Università di Napoli Federico II, CFEPSR in Portici, and by funds of L.R. 5/2002.

* Corresponding author. Tel.: +39-0812537460; fax: +39-0812537466.
E-mail addresses: angela.delia@unina.it (Angela D'Elia), domenico.piccolo@unina.it (Domenico Piccolo).

0167-9473/$ - see front matter
doi:10.1016/j.csda.2004.06.012

Although it is often assumed that the population of raters is homogeneous, there are also many examples where we should consider likely the presence of heterogeneity among the judges (Marden, 1995, Chapter 10).

For this aim, mixture models have been proposed, where the population is composed of homogeneous sub-populations, and the same type of standard ranking model is used for each group, but the parameters are allowed to differ (Bockenholt, 1993). In the same vein, Croon and Luijkx (1993) developed latent structure models for ranking data in order to extend the applicability of the Bradley–Terry–Luce model, while Stern (1993) proposed mixture models for the analysis of data from elections, in order to consider the different socio-political attitudes of the population. More recently, Murphy and Martin (2003) have extended the use of mixtures to distance-based models, imposing constraints on the "precision" parameters and getting, in this way, a wide modelling flexibility.

All the previous developments have been focussed on mixture models where the component distributions are the same for all the sub-populations, but their parameters values differ. In this way, they describe the presence of heterogeneity among the raters.

In this work, instead, we propose a model for ranks data, as the mixture of two different probabilistic structures. Indeed, this choice is motivated not only by taking into account the presence of different sub-groups of judges, but also by considering the ranking (elicitation) process itself as the sum of two distinct components, as it will be described in detail below.

The paper is organized as follows. In Section 2, we briefly recall some preliminary results, related to the component distributions. Thus, in Section 3, we develop a mixture model for ranks, focussing on its probabilistic features. Some inferential and computational issues are discussed in Sections 4.1 and 4.2, respectively. The evidence obtained from some real data sets are shown in Section 5, and further developments are considered in Section 6.

2. Background setting

Let r be the rank assigned by a rater to a given item among m. Following a *paired comparisons* criterion, D'Elia (2000) proposed to consider r as the realization of a Shifted binomial random variable $R \sim SB(\xi, m)$, with probability mass function

$$Pr(R = r) = \binom{m-1}{r-1} (1-\xi)^{r-1} \xi^{m-r}, \quad r = 1, 2, \ldots, m$$

and mean value and variance

$$E(R) = \xi + m(1-\xi), \quad Var(R) = (m-1)\xi(1-\xi).$$

Assuming that $R = 1$ means "most preferred", and $R = m$ means "least preferred", it is easily shown that the parameter $\xi \in [0, 1]$ increases with the liking feeling towards the item. Moreover, $Var(R)$ is maximum when $\xi = \frac{1}{2}$, when there is the greatest uncertainty in assigning a rank to the item. We would like to stress that this uncertainty component plays an important role in the elicitation process, especially when the raters rank items towards whom there is not a strong liking or disliking feeling.

An important feature of the SB random variable is that it allows the presence of an intermediate mode, and thus it results in a convincing tool for representing empirical preferences data. Indeed another probabilistic model, based on the inverse hypergeometric (IHG) random variable (Guenther, 1975), has been successfully proposed for analyzing ranks data (D'Elia, 1999; 2003a): however, the IHG distribution has the drawback of a monotonic shape of its probability mass function.

On the other hand, if we consider the case when there is a sort of indifference or *equi-preference feeling* towards a given item, then it seems appropriate to model ranks by means of a uniform discrete distribution: $U \sim \mathrm{Ud}(m)$, with probability mass function: $Pr(U = r) = 1/m$, $r = 1, 2, \ldots, m$.

This means to assume that the item has the same probability of receiving any rank $r \in [1, m]$. Of course, this happens if there is a total uncertainty with regard to the feeling towards an item: in fact, the Uniform random variable maximizes the entropy, among all the discrete distributions with finite support $\{1, 2, \ldots, m\}$, for a fixed m (Papoulis, 1984, pp. 514–515).

3. A mixture model for ranks data

Following the lines of the previous section, we propose that the ranking of an item may be represented as the mixture of two components: the liking/disliking feeling and the uncertainty of the choice process.

Indeed, especially for those items that do not excite strong liking or disliking feelings, it seems plausible to assume that the elicitation mechanism exhibits a greater uncertainty, in addition to that typical of all the individual choices and behaviors. This can be easily shown if, for instance, we consider a repeated ranking of several items by the same group of raters: while it is expected that the extreme ranks will remain unchanged, it is very plausible that the middle part of the ranking will change somewhere.

Furthermore, a mixture model for preferences data seems an adequate tool also for representing the heterogeneity due to the presence of two different sub-groups of raters: a *thoughtful* and an *instinctive* one, for whom there is a different weight of the uncertainty component.

As a consequence, we propose to take into account the composite nature of the ranking (or elicitation) process by means of a mixture model, whose components are the SB and the Ud random variables, respectively.

We assume that the rank assigned to a given item can be considered as of a realization of a mixture of a uniform and a shifted binomial (MUB) distribution

$$Pr(R = r) = \pi p_{\mathrm{B}}(r) + (1 - \pi) p_{\mathrm{U}}(r), \quad r = 1, 2, \ldots, m,$$

where $p_{\mathrm{B}}(r)$ and $p_{\mathrm{U}}(r)$ represent the probability mass functions of the SB and Ud random variables, respectively.

Thus, we define $R \sim MUB(m, \pi, \xi)$ if

$$Pr(R = r) = \pi \binom{m-1}{r-1} (1 - \xi)^{r-1} \xi^{m-r} + (1 - \pi) \frac{1}{m}, \quad r = 1, 2, \ldots, m.$$

From a computational point of view, the following recursive relations are more efficient:

$$Pr(R=1)=\pi\xi^{m-1}+(1-\pi)\frac{1}{m},$$

$$Pr(R=r+1)=Pr(R=r)\left(\frac{1-\xi}{\xi}\right)\left(\frac{m-r-1+\pi}{r}\right)+\left(\frac{1-\pi}{m\,\xi}\right),$$
$$r=1,2,\ldots,m-1.$$

The two components of our distribution have weights that depend upon π and $1-\pi$, respectively, with $\pi \in [0,1]$. In particular, we get the following cases:

- $\pi \to 0$: then, R tends to behave as a uniform distribution, and the rank assigned to a given item depends only upon the numbers m of the items. This is the case of total uncertainty, or *equi-preference feeling*.
- $\pi \to 1$: then, R tends to behave as a shifted binomial distribution, and its features depend only upon the parameter ξ. This case is analogous to that of a preferences order arising from a paired comparisons criterion.
- $\pi \in (0,1)$: then, $(1-\pi)$ measures how the uncertainty affects the elicitation mechanism and, as a consequence, the ranking.

Of course, if we assume the co-existence of two different sub-group of raters, the coefficients π and $1-\pi$ would represent their proportion in the population, respectively.

If we let μ_{B}, μ_{U} the mean values of the two components of the mixtures, the mean value of the MUB distribution is

$$E(R)=\pi\mu_{\mathrm{B}}+(1-\pi)\mu_{\mathrm{U}}=\pi(m-1)\left(\frac{1}{2}-\xi\right)+\frac{m+1}{2}$$

that reduces to $E(R)=(m+1)/2$ when $\xi=\frac{1}{2}$ (symmetric distribution).

As far as it concerns the variance of the mixture, with an obvious notation, we get

$$\begin{aligned}Var(R)&=\pi\sigma_{\mathrm{B}}^2+(1-\pi)\sigma_{\mathrm{U}}^2+\pi(1-\pi)(\mu_{\mathrm{B}}-\mu_{\mathrm{U}})^2\\&=(m-1)\left\{\pi\xi(1-\xi)+(1-\pi)\frac{m+1}{12}+\pi(1-\pi)\frac{(m-1)(2\xi-1)^2}{4}\right\}\end{aligned}$$

that reduces to $Var(R)=(m-1)[\pi/4+(1-\pi)(m+1)/12]$, for $\xi=\frac{1}{2}$.

If we consider the behavior of $E(R)$ and $Var(R)$ over the parametric space, we notice that

(i) $E(R)$ decreases when both π and ξ tend to 1: this implies a strong liking towards the item;

(ii) $Var(R)$ decreases monotonically with increasing $\mid \xi-\frac{1}{2} \mid$ for a fixed π, and decreases with π for a fixed ξ, (see the following Fig. 1).

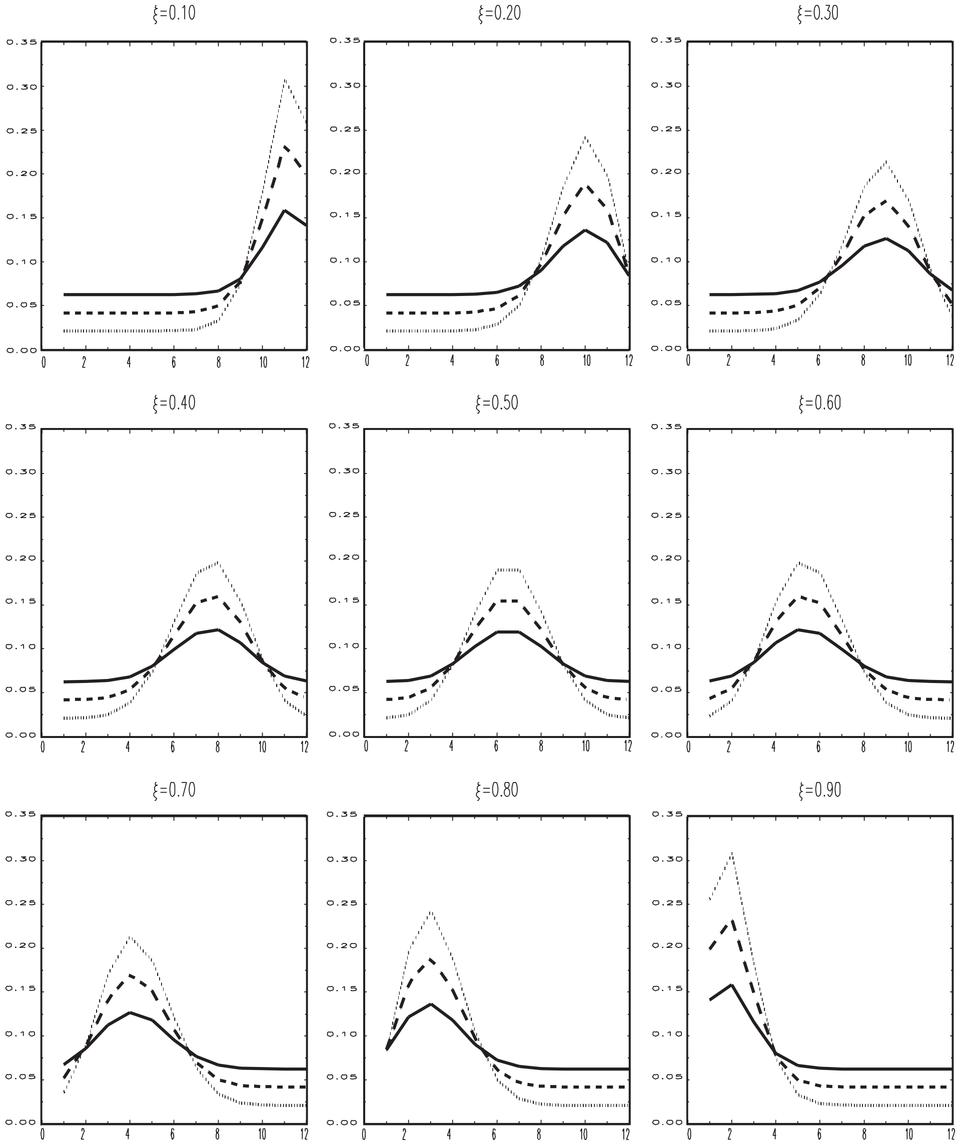

Fig. 1. Probability distribution functions of the MUB random variable, for $m = 12$, $\pi = \frac{1}{4}$ (solid line), $\pi = \frac{1}{2}$ (dashed line), $\pi = \frac{3}{4}$ (dotted line).

The main features of the MUB distribution are

- the MUB distribution admits an intermediate mode;
- the presence of the *uncertainty share*, measured by $(1 - \pi)/m$, makes more heavy the tails of the distribution;

- when $\xi = \frac{1}{2}$, the MUB distribution has a symmetric shape, being a convex linear combination of two symmetrical distributions; the asymmetry sign depends upon $(\xi - \frac{1}{2})$;
- the MUB model is reversible: that is, if $R \sim MUB(m, \pi, \xi)$ then $(m - R + 1) \sim MUB(m, \pi, 1 - \xi)$.

In order to show these features, Fig. 1 plots the probability mass functions of the MUB distribution for $m = 12$, and varying values of the parameters π, ξ.

As a consequence, the MUB model results in a very flexible tool for describing and analyzing preferences data, as it will be confirmed by the fitting to several real data sets (Section 5).

As a final issue, we need to discuss the meaning of the parameters π and ξ. In first instance, the π parameter is inversely related to the uncertainty of the probabilistic model, since $(1-\pi)/m$ is the proportion of the uniform component spread out over all the elements of the support. Instead, a more difficult task is to give a precise meaning to the parameter ξ. Indeed, from the probability mass function of the MUB distribution, we get

$$\xi = \left\{ \frac{1}{\pi} \left[Pr(R = 1) - \frac{1 - \pi}{m} \right] \right\}^{1/(m-1)}.$$

Thus, although ξ is related in some way to the liking feeling towards the item (since it increases with $Pr(R=1)$), it is not immediate to consider it as a direct preference measure. In a sense, we could say that the joint increase of both π and ξ means a greater preference feeling (since, this fact lowers the mean value of R); on the other hand, it is not so easy to establish a precise meaning when both the parameters tend to zero.

4. Inferential and computational issues

Let $(R_1, R_2, \ldots, R_n)$ be a random sample of i.i.d. MUB random variables, and $(r_1, r_2, \ldots, r_n)'$ be the observed ranks assigned by n raters to a given item among m, being m fixed and known. For inferential purposes it is important to notice that we can use equivalently the observed frequencies vector $(n_1, n_2, \ldots, n_m)'$, where n_r is the frequency of $R = r$, $r = 1, 2, \ldots, m$. This leads to a great gain in the computational effort of the estimation, since in real data sets we have $m \ll n$.

4.1. The maximum likelihood estimation

Let $\theta = (\pi, \xi)'$ be the unknown parameters vector; then, the log-likelihood function for the MUB model is

$$\log L(\theta) = \sum_{r=1}^{m} n_r \log(p_r(\theta)),$$

where we let $p_r(\theta) = Pr(R=r \mid \theta)$ for easiness of notation. Of course, this expression seems difficult to deal with and, for this reason, we will get the maximum likelihood estimates of θ by means of the E-M algorithm, as it will be discussed in the Section 4.2.

As far as it concerns the standard errors of the maximum likelihood estimates for grouped data (Rao, 1973, pp. 367–368), generalizing the arguments developed by D'Elia (2003a), we get

$$E\left(-\frac{\partial^2 \log L(\theta)}{\partial \pi^2}\right) = n \sum_{r=1}^{m} \frac{\left\{\frac{\partial p_r(\theta)}{\partial \pi}\right\}^2}{p_r(\theta)} = n d_{\pi\pi},$$

$$E\left(-\frac{\partial^2 \log L(\theta)}{\partial \xi^2}\right) = n \sum_{r=1}^{m} \frac{\left\{\frac{\partial p_r(\theta)}{\partial \xi}\right\}^2}{p_r(\theta)} = n d_{\xi\xi},$$

$$E\left(-\frac{\partial^2 \log L(\theta)}{\partial \pi \partial \xi}\right) = n \sum_{r=1}^{m} \frac{\left\{\left(\frac{\partial p_r(\theta)}{\partial \pi}\right)\left(\frac{\partial p_r(\theta)}{\partial \xi}\right)\right\}}{p_r(\theta)} = n d_{\pi\xi}.$$

Thus, the asymptotic variance and covariance matrix of the maximum likelihood estimators results

$$\mathbf{V} = \frac{1}{n}\begin{bmatrix} d_{\pi\pi} & d_{\pi\xi} \\ d_{\pi\xi} & d_{\xi\xi} \end{bmatrix}^{-1}.$$

From a computational point of view, the **V** matrix can be easily computed since, from the peculiar nature of the MUB distribution, we obtain

$$\frac{\partial p_r(\theta)}{\partial \pi} = q_r - \frac{1}{m}, \quad \frac{\partial p_r(\theta)}{\partial \xi} = \pi q_r \frac{m - \xi(m-1) - r}{\xi(1-\xi)}, \quad r = 1, 2, \ldots, m,$$

where $q_r = Pr(R = r \mid \pi = 1)$ is recursively computed by

$$q_1 = \xi^{m-1}, \quad q_r = q_{r-1}\frac{1-\xi}{\xi}\left(\frac{m}{r-1} - 1\right), \quad r = 2, 3, \ldots, m.$$

Of course, we substitute in the previous expressions the E-M estimates of the parameters π and ξ.

From this result, asymptotic tests about parameters values and related confidence intervals are easily obtained. For instance, the confidence ellipse at $100(1-\alpha)\%$ level is given by

$$\left\{(\pi, \xi) : d_{\pi\pi}(\hat{\pi} - \pi)^2 + 2d_{\pi\xi}(\hat{\pi} - \pi)(\hat{\xi} - \xi) + d_{\xi\xi}(\hat{\xi} - \xi)^2 \leqslant -\frac{2}{n}\log(\alpha)\right\}.$$

4.2. The E-M estimates

The maximum likelihood estimates of the parameters of the mixture model MUB can be obtained by means of the E-M algorithm, originally introduced by Dempster et al. (1977) for dealing with missing values, and then widely developed for the fitting of mixture models (for recent reviews, see McLachlan and Krishnan (1997), McLachlan and Peel (2000) and the special issue of *Computational Statistics and Data Analysis*, 2003, 41 (3–4)).

In order to get the E-M estimates of the parameters of the MUB model, we can consider the ranks data as incomplete, because the appropriate mixture component for each rater is unknown. Then, let us introduce the latent variables Z_{gi} $(g=1, 2;\ i=1, 2, \ldots, n)$, such that $Z_{gi}=1$ if the ith rater’s preferences come from the gth distribution, and $Z_{gi}=0$ otherwise.

Thus, letting $\mathbf{r}=(r_1, r_2, \ldots, r_n)'$, and $\mathbf{z}=(z_{11}, z_{12}, \ldots, z_{1n}; z_{21}, z_{22}, \ldots, z_{2n})'$, the complete data log-likelihood function is

$$\log L_c(\theta; \mathbf{r}, \mathbf{z}) = \sum_{g=1}^{2}\sum_{i=1}^{n} z_{gi}\{\log(\pi_g) + \log(p_g(r_i; \xi_g))\},$$

where $p_g(r_i; \xi_g)$ represents the probability mass function for the ith observation from the gth component of the mixture; in particular, we have: $p_g(r_i; \xi_g) = p_{\mathrm{B}}(r_i; \xi)$ for $g=1$; $p_g(r_i; \xi_g) = p_{\mathrm{U}}(r_i)$ for $g=2$.

Given this setting, the E-M algorithm for the MUB model involves the iteration (until convergence) of the following steps:

- *E-Step* (kth iteration): Compute, preliminarily

  $$E(Z_{gi} \mid \mathbf{r}; \theta^{(k)}) = \frac{\pi_g^{(k)} p_g(r_i; \xi_g^{(k)})}{\pi_1^{(k)} p_1(r_i; \xi^{(k)}) + \pi_2^{(k)} p_2(r_i)} = \tau_g(r_i; \theta^{(k)})$$

  and then

  $$E(\log L_c(\theta^{(k)}; \mathbf{r}, \mathbf{z})) = \sum_{g=1}^{2}\sum_{i=1}^{n} \tau_g(r_i; \theta^{(k)})\{\log(\pi_g^{(k)}) + \log(p_g(r_i; \xi_g^{(k)}))\}$$

  for $g=1,2$; $i = 1, 2, \ldots, n$. Here, $\tau_g(r_i; \theta^{(k)})$ represents the estimate of the (posterior) probability that the R_i random variable belongs to the gth component of the mixture $(g = 1, 2)$, given the current estimates of the parameters $\theta^{(k)}$.
- *M-Step* (kth iteration): Compute new maximum likelihood estimates of the parameters of the mixture, by maximizing the function

  $$\begin{aligned} Q(\theta^{(k)}) =& \sum_{i=1}^{n} \{\tau_1(r_i; \theta^{(k)}) \log(\pi^{(k)}) + \tau_2(r_i; \theta^{(k)}) \log(1 - \pi^{(k)})\} \\ &+ \sum_{i=1}^{n} \{\tau_1(r_i; \theta^{(k)}) \log[p_1(r_i; \xi^{(k)})] + \tau_2(r_i; \theta^{(k)}) \log[p_2(r_i)]\}. \end{aligned}$$

  In particular, we get an explicit expression of the estimate of π to be used in the $(k+1)$th iteration of the algorithm

  $$\pi^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n} \tau_1(r_i; \theta^{(k)}) = \frac{1}{n}\sum_{i=1}^{n} w_i,$$

  where we let $w_i = \tau_1(r_i; \theta^{(k)})$, $i = 1, 2, \ldots, n$, for simplicity.

For the peculiar nature of our mixture distribution, the estimate of the parameter ξ at the kth iteration comes from $\sum_{i=1}^{n} w_i (\partial \log[p_1(r_i;\xi)]/\partial\xi) = 0$, that is

$$\sum_{i=1}^{n} w_i \left[\frac{1}{1-\xi} + \frac{m}{\xi}\right] = \sum_{i=1}^{n} w_i \left[\frac{r_i}{\xi(1-\xi)}\right].$$

Then

$$\xi^{(k+1)} = \frac{m\sum_{i=1}^{n} w_i - \sum_{i=1}^{n} r_i w_i}{(m-1)\sum_{i=1}^{n} w_i} = \frac{m - \sum_{i=1}^{n} r_i w_i / \sum_{i=1}^{n} w_i}{m-1}.$$

This result allows a useful interpretation of the ξ parameter estimate. Indeed, by letting $\bar{R}_n(p) = \sum_{i=1}^{n} r_i w_i / \sum_{i=1}^{n} w_i$, from the last expression, we get

$$\xi^{(k+1)} = \frac{m - \bar{R}_n(p)}{m-1}.$$

Now, $\bar{R}_n(p)$ represents the average of the observed ranks, weighted with the posterior probability that r_i is a realization of the SB distribution, given the current data. Of course, when $w_i = w, \forall i = 1, 2, \dots, n$, then $\bar{R}_n(p) = \bar{R}_n$ (the average rank), and $\xi^{(k+1)}$ is the same as the maximum likelihood estimator $\hat{\xi}$ in the SB model (D'Elia, 2000).

As far as it concerns the choice of the starting values vector $\theta^{(0)} = (\pi^{(0)}, \xi^{(0)})'$ for initializing the E-M algorithm, in absence of any a priori information, the previous results suggest to use: $\pi^{(0)} = \frac{1}{2}$; $\xi^{(0)} = (m - \bar{R}_n)/(m-1)$. Finally, in order to assess the achievement of the convergence of the E-M estimates, we used (for all the data sets in Section 5) the criterion of stopping the iterations when: $\log L(\theta^{(k+1)}) - \log L(\theta^{(k)}) < 10^{-10}$. The finite samples performance of maximum likelihood estimators, obtained by the reported E-M algorithm, has been checked (D'Elia, 2003b) by means of an extensive Monte Carlo study. These results supported the consistency of our approach, also for small samples.

5. Empirical evidence

The proposed MUB model has been fitted to three different data sets, in order to check the adequateness of our proposal and for highlighting some related critical issues. Indeed, as it will appear from the evidence shown in this section, there is not a best model for all ranks data, representing preferences and/or evaluations, since in many case the behavior of the sub-populations is a main problem.

In the following subsections we are going to discuss the results for the APA election data (5.1), the main evidence from a study about the evaluation given by the students to the University they attend (5.2), and finally the results obtained in a survey on the preferences towards different living places (5.3). The data sets used in Sections 5.2 and 5.3 are freely accessible on the web-site: http://www.dipstat.unina.it/spazioricerca.htm. The E-M algorithm for obtaining the ML estimates was implemented by means of the programming language $GAUSS^{®}$ 5.0, on a Pentium 4, 256 Mb. For all the estimated models the algorithm convergence time was less than 0.01 seconds.

Table 1
Estimation results of MUB and IHG models, fitted to APA election data

Candidate	MUB				IHG	Fitting measures			
	$\hat{\pi}$	$\hat{\xi}$	$\frac{1-\hat{\pi}}{m}$	$Pr(R=1)$	$Pr(R=1)$	χ^2_{MUB}	χ^2_{IHG}	$AICC_{\text{MUB}}$	$AICC_{\text{IHG}}$
A	0.305 (0.020)	0.643 (0.012)	0.139	0.191	0.232	5.8	129.2	18236.3	18358.8
B	0.339 (0.020)	0.393 (0.011)	0.132	0.140	0.189	3.6	245.4	18212.3	18455.0
C	0.101 (0.011)	0.998 (0.016)	0.180	0.280	0.199	145.1	362.7	18261.8	18471.9
D	0.076 (0.014)	0.104 (0.029)	0.185	0.185	0.183	20.0	26.0	18430.7	18434.7
E	0.057 (0.021)	0.638 (0.062)	0.189	0.198	0.200	8.8	5.8	18476.8	18471.9

The estimated parameters standard errors are in parentheses.

5.1. APA election data

The American Psychological Association (APA) is a professional organization, that elects a president every year by asking each member to rank (according to a preference order) five candidates (A, B, C, D, E). During the 1980 election, about 15 000 members voted, but only 5738 ballots resulted complete.

This data set has been analyzed by many Authors (see, for instance, Diaconis, 1989; McCullagh, 1993; Stern, 1993; and more recently, Yu, 2000; Murphy and Martin, 2003). All of them found that these data seem to be the preferences expression of a mixture of at least two groups of voters: the academic and the clinical psychologists. For this reason, our mixture (MUB) model seems to be a good candidate for describing the APA data set.

In Table 1 , we show the main results obtained by fitting the MUB model to the APA election data. For a comparison purpose, also the results from the IHG model are illustrated.

Comparing the fitting measures (χ^2 and the corrected Akaike information criterion: $AICC = -2\log L(\theta) + 2k + 2k(1+k)/(n-k+1)$, where k is the number of θ's parameters: Hurvich and Tsay, 1989; Burnham and Anderson, 2002, pp. 60–67), it emerges that (with the exception of candidate E) the MUB model fits the data better than the IHG one.

As far as it concerns the components of the MUB distribution, we can notice that the uncertainty share $(1-\pi)/m$ is greater in the ranking of candidates C, D and E. This might mean that for these three candidates there was a stronger contrast between the two sub-groups of raters (indeed, it is known that clinic psychologists preferred C, while academic ones preferred D and E), leading to a greater uncertainty in the ranks they received, as confirmed by their U-shaped observed frequencies distributions (see Fig. 2, where observed and MUB expected frequencies have been plotted). In particular, it is worth noticing that the candidate E, whose observed ranks distribution exhibits the more irregular shape, is also the one with the greatest uncertainty share (0.189).

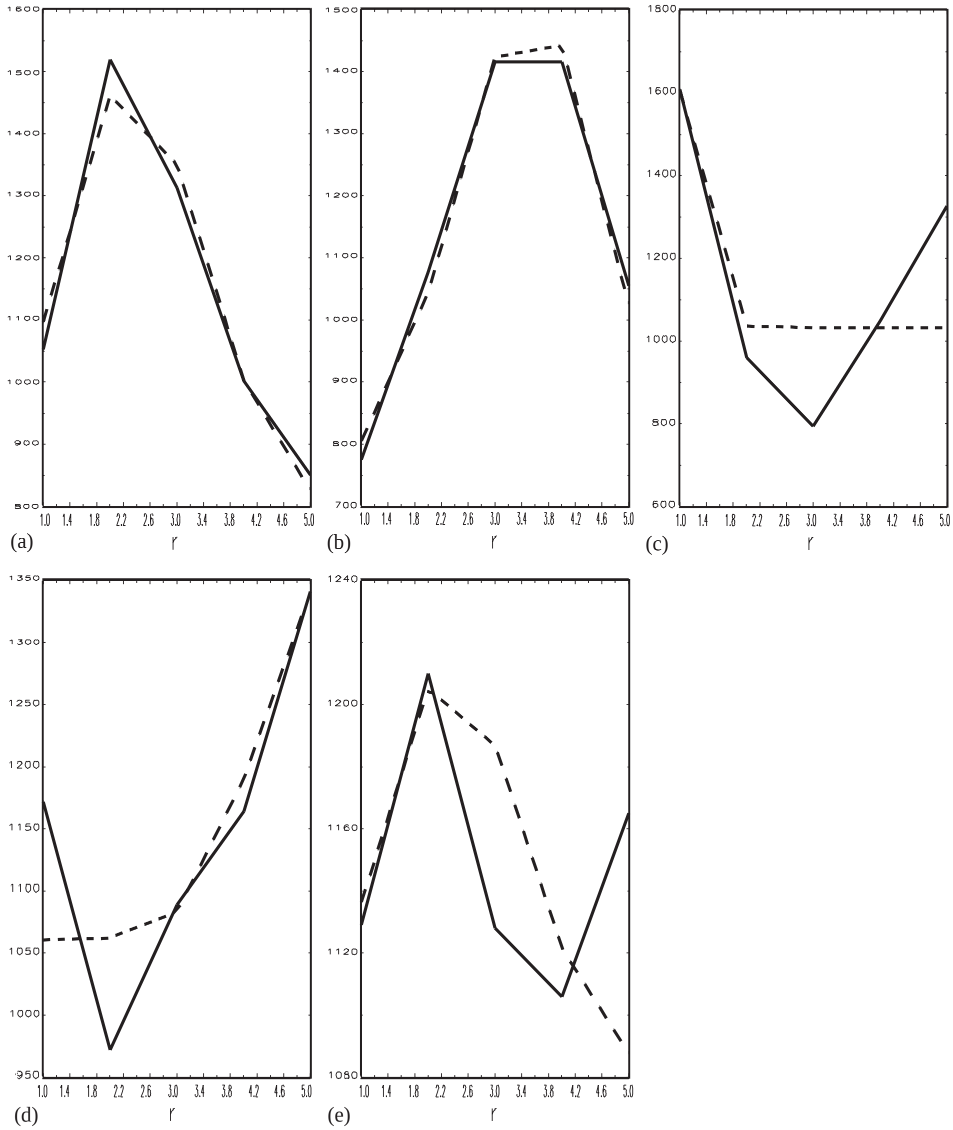

Fig. 2. APA data: observed (solid line) and expected frequencies for MUB (dashed line) model.

Finally, if we consider the estimated $Pr(R=1)$ for all the candidates, we get the following preference order: {C E A D B}, that is consistent with those obtained by Stern (1993) using a mixture of two Bradley–Terry–Luce models, and by Yu (2000) using a multivariate normal order-statistics model (MVNOS), as shown in Table 2 . This result is also consistent with the actual winner of the APA election, who was candidate C on the basis of the Hare system (Hare, 1865).

Table 2
Observed relative frequencies and estimated $Pr(R=1)$, under MUB, MVNOS and Stern's models

Candidate	n_1/n	$Pr(R=1)$		
		MUB	MVNOS	Stern
A	0.184	0.191	0.193	0.199
B	0.135	0.140	0.130	0.153
C	0.280	0.280	0.276	0.276
D	0.204	0.185	0.198	0.186
E	0.197	0.198	0.200	0.186

Table 3
Estimation results of MUB and IHG models, fitted to evaluation data

Topic	MUB				IHG		
	$\hat{\pi}$	$\hat{\xi}$	$\frac{1-\hat{\pi}}{m}$	$Pr(R=1)$	$Pr(R=1)$	n_1/n	$\bar{R}_n$
Classrooms	0.291	0.795	0.101	0.175	0.187	0.177	3.472
	(0.006)	(0.003)					
Exams dates	0.354	0.543	0.092	0.101	0.146	0.061	4.083
	(0.010)	(0.005)					
Exams information	0.350	0.713	0.093	0.139	0.186	0.142	3.535
	(0.007)	(0.003)					
Workshops	0.534	0.786	0.066	0.193	0.239	0.208	3.054
	(0.006)	(0.002)					
Lessons schedule	0.494	0.746	0.072	0.158	0.212	0.151	3.286
	(0.006)	(0.002)					
Lessons calendar	0.744	0.868	0.037	0.354	0.366	0.370	2.315
	(0.004)	(0.001)					
Global satisfaction	0.755	0.718	0.035	0.139	0.256	0.120	3.024
	(0.005)	(0.002)					

The estimated parameters standard errors are in parentheses.

5.2. University evaluation data

During the year 2002, an evaluation survey was conducted on the students attending the University of Naples Federico II (the largest University in the South of Italy). Among these, about 27 000 students completed the assigned questionnaire.

In particular, each student was asked to express his/her satisfaction (on a range from 1 = "very satisfied" to 7 = "not at all satisfied") with respect to different aspects related to the teaching organization and to the provided educational facilities. In detail, the issues concerned: the adequateness of classrooms, the professors' compliance of the schedule and calendar of the lessons, the usefulness of workshops, the information received about the exams and the respective dates, and finally a measure of the overall satisfaction.

From the estimation of the MUB model for this data set, we get the results shown in Table 3.

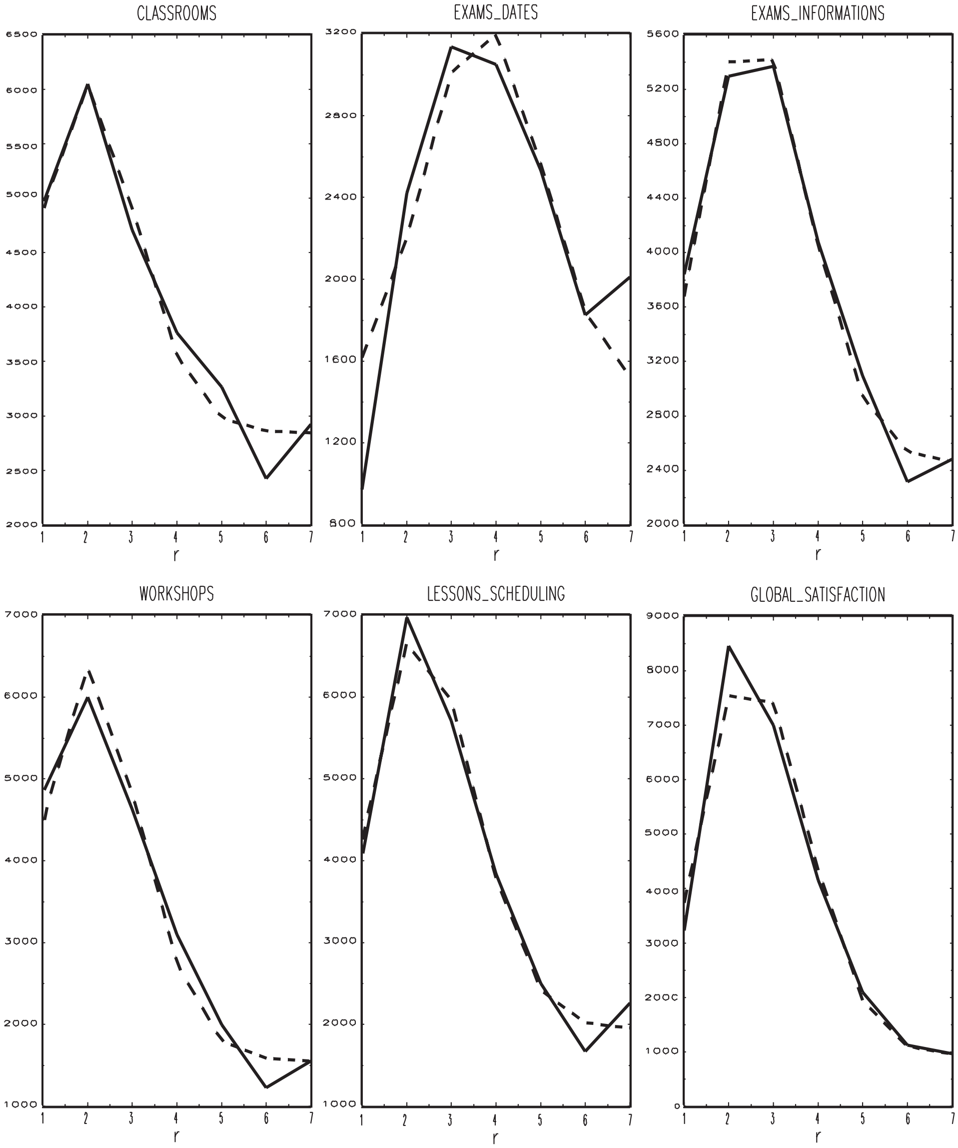

Fig. 3. Evaluation data: observed (solid line) and expected frequencies for MUB (dashed line) model.

Thus, it appears that the evaluation towards the adequateness of classrooms has the greatest uncertainty share: this might be due to the presence inside the same University of very heterogeneous buildings (old and historical ones, for humanistic Faculties, and modern and efficient ones for scientific Faculties). On the other hand, the smallest uncertainty share happens for the professors' compliance of lessons' calendar and for the global

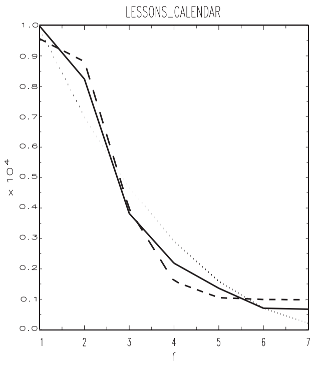

Fig. 4. Evaluation data: observed (solid line) and expected frequencies for MUB (dashed line) and IHG (dotted line) models.

satisfaction, showing that for these aspects there is an homogeneous opinion among the students.

It is important to notice the professors' compliance of lessons' calendar is also the item that receives the lowest average rank $\bar{R}_n$ ($=2.315$, that is the highest evaluation), and gets the highest estimated $Pr(R=1)=0.354$.

As far as it concerns the goodness of fit of the MUB model, since $n \simeq 27\,000$, the χ^2 measure would result inflated by n and, thus, is unsuitable. For this reason, we prefer to assess the fitting by means of the visual inspection of Fig. 3.

Then, it emerges that the MUB model fits very well the evaluations data, and it seems capable to catch also different shapes of the observed frequencies distributions.

Finally, with respect to the professors' compliance of lessons' calendar, also the IHG model fits well the observed data (Fig. 4): this might depend on the strong satisfaction feeling of the students towards this item, leading to a monotonic observed frequencies distribution and, as a consequence, to the adequateness of the IHG model, too.

5.3. Preferences towards living places

An analysis was conducted in order to study the preferences of young people (living in Naples, Italy) towards several Italian cities. In detail, each rater was asked to rank $m=12$ cities from the most preferred as living place, until to the least preferred on the basis of the same criterion. The results obtained from the fitting of MUB models—and for comparison from the IHG one—are shown in Table 4 and Figs. 5 (level curves of the log-likelihood function) and 6. The cities order is based on the increasing values of their average rank $\bar{R}_n$.

Thus, we can notice that the MUB model exhibits an adequate fit for almost all the observed preferences. In particular, some issues are worth to be considered.

- The MUB model has the best fitting for the city of Venice, while the IHG one shows the best fit for Naples. Indeed, the heterogeneity of the raters with respect to the sex (and, thus, with regard to their emotional feeling) makes a mixture model more adequate for representing the preferences towards Venice. On the other hand, the fact that the survey has been conducted on people living in Naples, has determined an obvious homogeneity

Table 4
Estimation results of MUB and IHG models, fitted to cities data

City	MUB				IHG	Fitting measures			
	$\hat{\pi}$	$\hat{\xi}$	$\frac{1-\hat{\pi}}{m}$	$Pr(R=1)$	$Pr(R=1)$	χ^2_{MUB}	χ^2_{IHG}	$AICC_{\text{MUB}}$	$AICC_{\text{IHG}}$
Florence	0.834 (0.035)	0.879 (0.008)	0.014	0.215	0.305	29.5	28.8	803.8	793.8
Rome	0.750 (0.041)	0.889 (0.009)	0.021	0.226	0.271	18.6	35.8	843.4	840.9
Naples	0.566 (0.051)	0.862 (0.013)	0.0.36	0.146	0.191	35.2	19.6	965.1	948.7
Bologna	0.539 (0.053)	0.847 (0.014)	0.038	0.125	0.176	22.6	21.6	974.9	968.7
Venice	0.535 (0.064)	0.579 (0.020)	0.039	0.040	0.105	9.1	52.2	1009.0	1054.9
Genoa	0.642 (0.060)	0.489 (0.017)	0.030	0.030	0.092	25.0	91.5	989.1	1063.6
Milan	0.155 (0.065)	0.303 (0.060)	0.070	0.070	0.064	34.7	26.3	1061.7	1052.3
Verona	0.168 (0.040)	0.014 (0.013)	0.069	0.060	0.050	36.8	28.9	1031.8	1017.1
Palermo	0.631 (0.057)	0.287 (0.015)	0.031	0.031	0.058	36.6	80.7	987.3	1030.2
Turin	0.340 (0.055)	0.142 (0.020)	0.055	0.055	0.051	62.0	52.4	1033.0	1023.1
Catania	0.635 (0.053)	0.205 (0.013)	0.030	0.030	0.050	36.1	75.3	962.7	1008.9
Bari	0.672 (0.048)	0.165 (0.012)	0.027	0.027	0.042	35.5	74.3	933.6	972.4

The estimated parameters standard errors are in parentheses.

of the feeling expressed toward this city, making the IHG model a good representation for its observed preferences data.

- For Rome, the χ^2 and the *AICC* measures give discording results. However, the inspection of Fig. 6 shows that both the models have a good fit to the observed ranks.
- The greatest uncertainty share is present for: Milan (0.070), Verona (0.069) and Turin (0.055). Indeed, the observed frequencies distributions of the ranks assigned to these cities show multimodal shapes, that might be due to the co-existence of different feelings towards them (e.g. the liking as far as it concerns the job opportunities, and the disliking for northern places).

6. Further developments

In this paper, we have proposed, as a tool for analyzing ranks data, a mixture model (MUB) whose component distributions represent the preference towards an item and the uncertainty feeling, respectively.

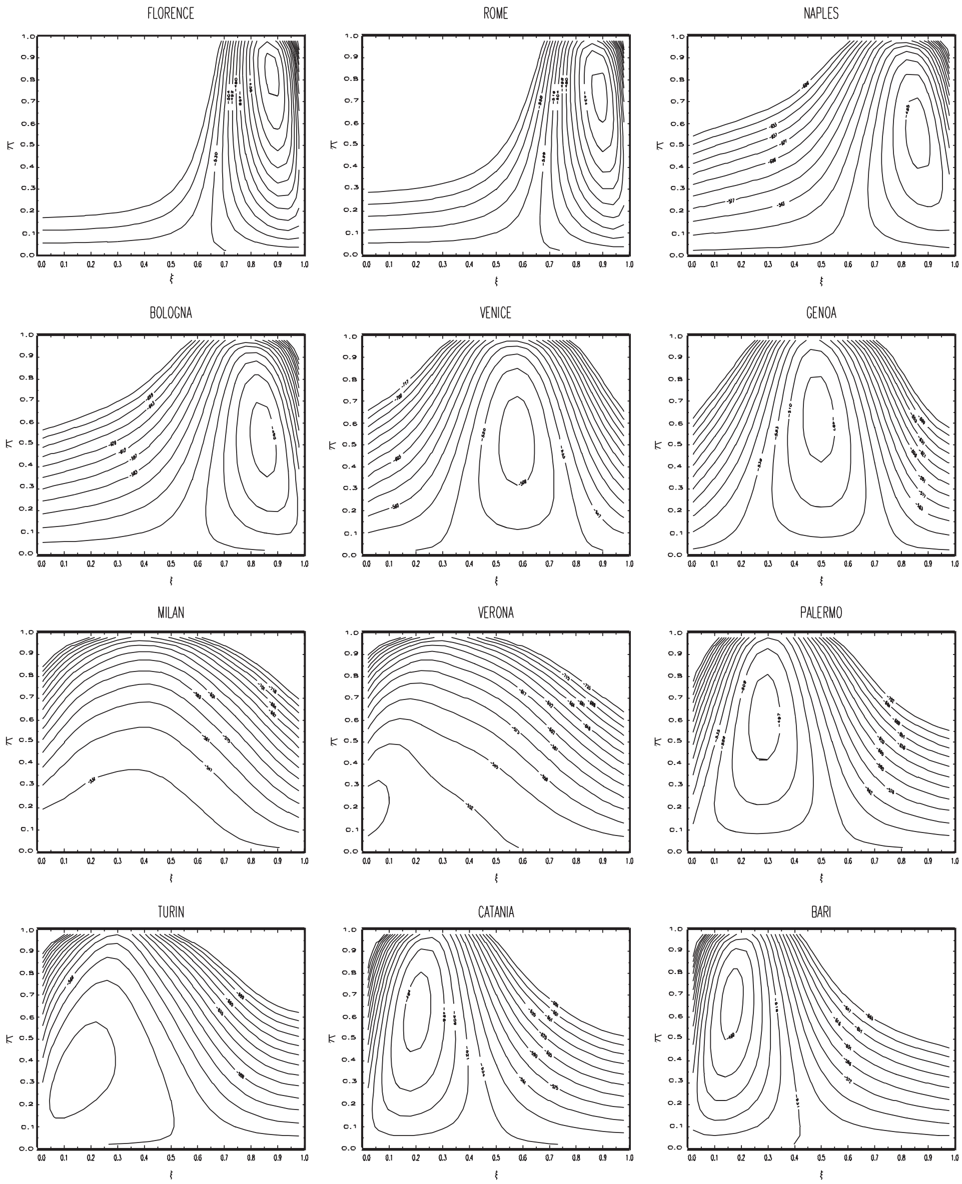

Fig. 5. Cities data: level curves of the MUB log-likelihood function.

Several extensions of this framework are worth to be considered. In first instance, a useful development would be to include covariates in the MUB model, following a Generalized Linear Models approach, as already shown for similar models in D'Elia (1999, 2000). In this way, it might be possible to study the effect of raters' features on the expressed preferences, and thus to improve the interpretation and the fitting to the observed data (e.g. in the case of multimodal shapes).

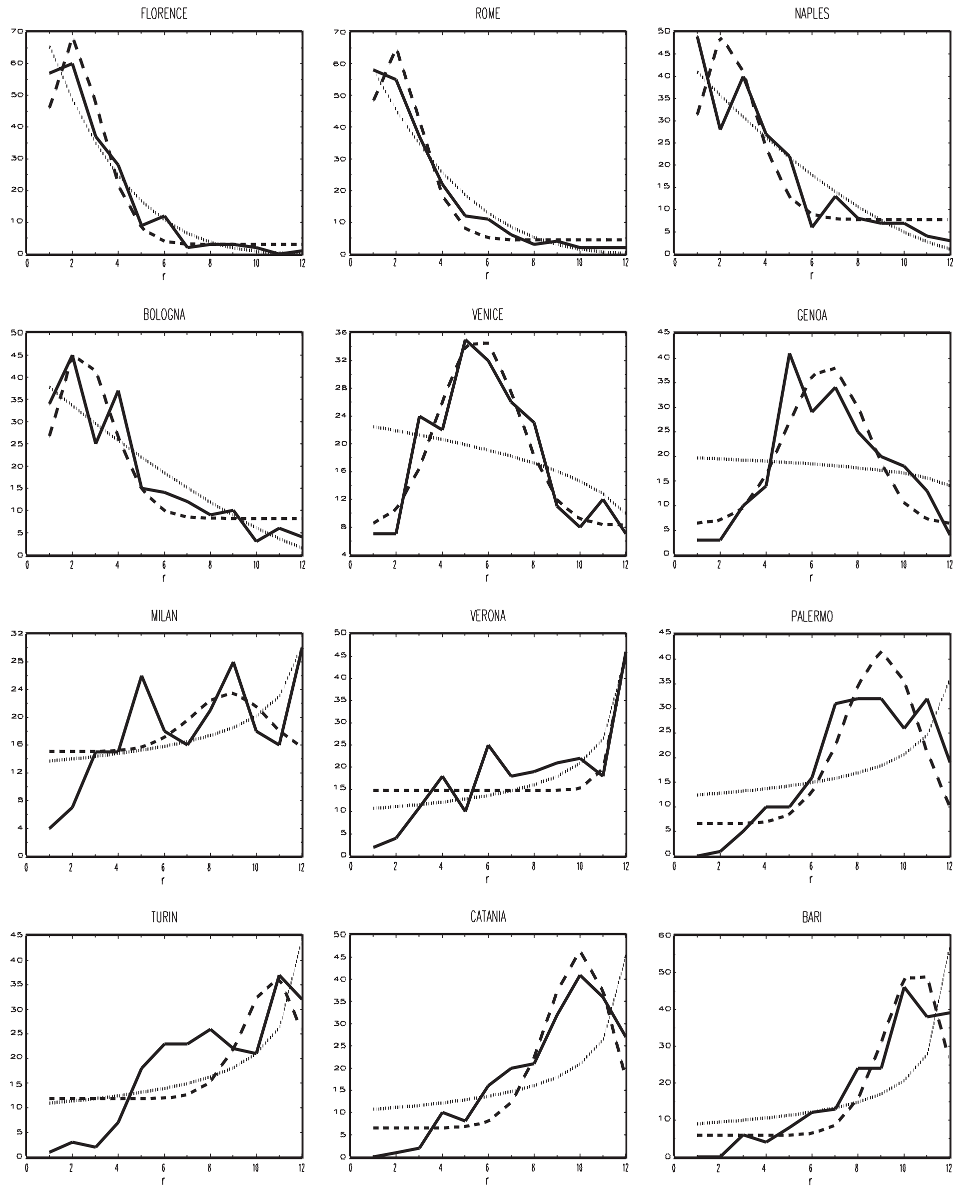

Fig. 6. Cities data: observed (solid line) and expected frequencies for MUB (dashed line) and IHG (dotted line) models.

In the same vein, also multilevel models should be considered, in order to take into account both the presence of heterogeneity and of hierarchical clusters among the raters: this last issue seems to be important especially in studies (as those on schools/Universities evaluation) where the raters are organized in clusters (e.g classes, Faculties, etc.), with homogeneous contents.

Finally, exploiting our results and the good performance of the proposed E-M estimators also in small samples, the planning of adequate and significant sampling experiments would be efficient in many fields as: marketing researches, opinion polls, evaluation enquiries, etc.

Acknowledgements

We thank the Associate Editor and Referees for their useful comments and suggestions.

References

Bockenholt, U., 1993. Applications of Thurstonian models to ranking data. in: Fligner, M.A., Verducci, J.S. (Eds.), Probability Models and Statistical Analyses of Ranking Data. Springer, New York, pp. 157–172.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodal Inference. 2nd Edition. Springer, New York.

Croon, M.A., Luijkx, R., 1993. Latent structure models for ranking data. in: Fligner, M.A., Verducci, J.S. (Eds.), Probability Models and Statistical Analyses of Ranking Data. Springer, New York, pp. 53–74.

D'Elia, A., 1999. A proposal for ranks statistical modelling. in: Friedl, H., Berghold, A.,Kauermann, G. (Eds.), Statistical Modelling. Graz, Austria, pp. 468–471.

D'Elia, A., 2000. A shifted binomial model for rankings. in: Nunez-Antón, V., Ferreira, E. (Eds.), Proceedings of the 15th International Workshop on Statistical Modelling, Servicio Editorial de la Universidad del Pais Vasco. Bilbao, pp. 412–416.

D'Elia, A., 2003a. Modelling ranks using the inverse hypergeometric distribution. Statist. Model. Internat. J. 3 (1), 65–78.

D'Elia, A., 2003b. Finite sample performance of the E-M algorithm for ranks data modelling. Statistica, forthcoming.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc. Ser. B 39 (1), 1–38.

Diaconis, P., 1989. A generalization of spectral analysis with application to ranked data. Ann. Statist. 17 (3), 949–979.

Fligner, M.A., Verducci, J.S., 1993. Probability Models and Statistical Analyses of Ranking Data. Springer, New York.

Guenther, W.C., 1975. The inverse hypergeometric—a useful model. Statist. Neerlandica 29, 129–144.

Hare, T., 1865. The Election of Representatives, Parlimentary and Municipal: A Treatise. 3rd Edition. Longman, Robert and Green, London.

Hurvich, C.M., Tsay, C.L., 1989. Regression and time series model selection in small samples. Biometrika 76, 297–307.

Marden, J.I., 1995. Analyzing and Modeling Rank Data. Chapman & Hall, London.

McCullagh, P., 1993. Permutations and regression models. in: Fligner, M.A., Verducci, J.S. (Eds.), Probability Models and Statistical Analyses of Ranking Data. Springer, New York, pp. 196–215.

McLachlan, G., Krishnan, T., 1997. The EM Algorithm and Extensions. Wiley, New York.

McLachlan, G., Peel, G.J., 2000. Finite Mixture Models. Wiley, New York.

Murphy, T.B., Martin, D., 2003. Mixtures of distance-based models for ranking data. Comput. Statist. Data Anal. 41, 645–655.

Papoulis, A., 1984. Probability, Random Variables, and Stochastic Processes. 2nd Edition. McGraw-Hill, Singapore.

Rao, C.R., 1973. Linear Statistical Inference and its Applications. 2nd Edition. Wiley, New York.

Stern, H., 1993. Probability models on rankings and the electoral process. in: Fligner, M.A., Verducci, J.S. (Eds.), Probability Models and Statistical Analyses of Ranking Data. Springer, New York, pp. 173–195.

Yu, P.L.H., 2000. Bayesian analysis of order-statistics models for ranking data. Psychometrika 65, 281–299.