



Discrete Beta and Shifted Beta-Binomial models for rating and ranking data

Mariangela Sciandra¹ · Salvatore Fasola¹ · Alessandro Albano^{1,2} · Chiara Di Maria¹ · Antonella Plaia^{1,2}

Received: 13 October 2023 / Accepted: 13 December 2023 / Published online: 27 January 2024
© The Author(s) 2024

Abstract

Ranking and rating methods for preference data result in a different underlying organization of data that can lead to manifold probabilistic approaches to data modelling. As an alternative to existing approaches, two new flexible probability distributions are discussed as a modelling framework: the *Discrete Beta* and the *Shifted Beta-Binomial*. Through the presentation of three real-world examples, we demonstrate the practical utility of these distributions. These illustrative cases show how these novel distributions can effectively address real-world challenges, with a particular focus on data derived from surveys concerning environmental issues. Our analysis highlights the new distributions' capability to capture the inherent structures within preference data, offering valuable insights into the field.

Keywords Discrete beta · Ranking · Rating · Shifted Beta-Binomial

Handling Editors: Giada Adelfio and Francesco Lagona

✉ Mariangela Sciandra
mariangela.sciandra@unipa.it

Salvatore Fasola
salvatorefasola@gmail.com

Alessandro Albano
alessandro.albano@unipa.it

Chiara Di Maria
chiara.dimaria@unipa.it

Antonella Plaia
antonella.plaia@unipa.it

¹ Department of Economics, Business and Statistics, University of Palermo, Viale delle Scienze, Building 13, 90128 Palermo, Italy

² Sustainable Mobility Center (Centro Nazionale per la Mobilità Sostenibile—CNMS), Milan, Italy

1 Introduction

In the literature on measurement of individual preferences and values, both rating and ranking score systems are frequently used. In most applications, observations are typically ranked; however, ratings have desirable executive and statistical properties. An ongoing debate exists between the proponents of the ranking method (Falahee and MacRae 1997) and those of the rating one (Villanueva et al. 2005).

Ranking data arise when n individuals are asked to order a set of K objects, or *items* (e.g., wines, movies, means of transport, etc.), according to some criterion, such as preference, importance and satisfaction. Subjects are required to observe all items and to establish their position relative to others, so that items can be ordered from the most to the least preferred. In the rating method, subjects are asked to assess each object (item) independently of the others, by scoring it on a response scale with K ordered categories, such as a Likert scale. The use of rating scales in measuring attitudes or values is very popular because they are easy to understand and less demanding than rankings, both in collecting and analyzing data.

Both rankings and ratings are widely used in surveys aiming at investigating people's opinions and attitudes. This is particularly important for policymakers can use rankings and ratings in the case of questions on sensitive topics, such as civil rights or environmental issues. The latter is the focus of the present paper.

A crucial aspect of the debate about ratings and rankings has been highlighted by Ovadia (2004) and concerns the cognitive organization of the values. A different set of assumptions about the nature and structure of data underlies each of the two approaches. When preference data are collected according to the ranking method, an intrinsic assumption of strict hierarchy between values is assumed because of the interdependence between ranks. This is no longer true for the rating score system in which each value is independent of each other and, as each object receives an absolute value, has possible occurrence of ties (an issue not addressed in this paper). These underlying assumptions about the organization of values have some methodological implications: rating score data do not ask for complex statistical methods, then they are advocated for their simplicity in execution and analysis (Agresti 2010); on the contrary, ranking methods result in data sets that cannot be analyzed with standard statistical methods because of the interdependence of the ranks (Alvo and Philip 2014). For this reason, ranking data modelling has received a lot of attention and many models have been proposed over the years (Dwass 1957; Critchlow et al. 1991; Lee and Philip 2010; Yu et al. 2019).

Some approaches have focused on modelling the approval rate of a single item, by relying on the *Shifted Binomial* (D'Elia 2000) and the *Inverse Hypergeometric* (D'Elia 2003) distributions, but these distributions are not sufficiently flexible to fit the empirical rank distributions. For this reason, mixtures of Discrete Uniform and Shifted Binomial random variables (MUB models) have been proposed to deal with both the selection mechanism and uncertainty in the ranking process

(D’Elia and Piccolo 2005). However, a discrete parametric distribution able to assume also “J” and “U” (non-monotonic convex) shapes is not yet present among the existing ones, as discussed by Punzo and Zini (2012). Based on the assumption that a single unidimensional latent variable governs the responses, two new probability distributions are discussed in this paper. The proposals apply to rankings and ratings, being quite flexible in shape while preserving simplicity, and also in terms of parameter interpretation.

The paper is organized as follows: in Sect. 2, we introduce notation and discuss recent findings in the rating and ranking literature. In Sect. 3, we present the theoretical definition of the proposed *Discrete Beta* and *Shifted Beta-Binomial* distributions with some inferential results. In Sect. 4, through the use of three real data sets coming from surveys addressing environmental issues, we obtain insights into the relative merits of the proposed distributions in dealing with rankings and ratings, respectively. We conclude with a discussion and directions for future research.

2 Theoretical background

Both rating and ranking data are often analyzed through a probabilistic modelling approach. This requires data scientists to make formal assumptions about the essential nature of the unobservable process that generated the data. However, mechanisms generating rating and ranking data may not be exactly distinguished. Indeed, assuming an underlying continuous scale of measurement, both rating and ranking data can be derived: an ordinal rating scale arises by partitioning the real axis into adjacent intervals and assigning them consecutive numerical scores, while ranks arise by determining the relative positions of two or more objects placed along the axis. For example, supposing that continuous numerical scores are assigned to seven objects and partitioning the real axis into five intervals, rating and ranking data are generated as illustrated in Fig. 1.

The following subsection discusses the schemes of item presentation and the most popular model approaches used for ranking and rating systems.

2.1 Rating and ranking systems

A rating score system assumes that subjects independently assess each object using a common measurement scale with K ordered categories. However, this system has a major drawback as judges tend to assign scores toward the extremes of the scale without discriminating between items (Tourangeau et al. 2000; de Rezende and de Medeiros 2022). Additionally, it can lead to reduced motivation to discriminate between

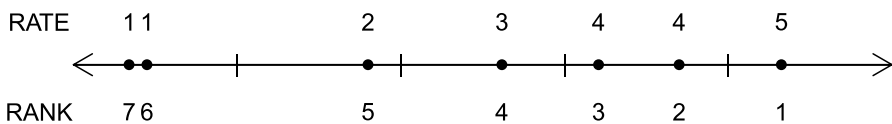


Fig. 1 Example of rating and ranking generating mechanism

items, potentially resulting in all items receiving the same score (Wind 2020; Ovadia 2004). Furthermore, personal bias can affect rating ordinal scales, leading to individual clustering of responses influenced by temperament, cultural background, and personal interests of judges (Kemmelmeier 2016; Harzing et al. 2009).

The most popular rating system is the Likert scale, where subjects assign scores based on their experiences, feelings, or emotions within a defined interval between minimum and maximum anchor points. Often, mean or median values are used to rank objects, though this practice has questionable assumptions. Sometimes, Likert-type responses are treated as categorical, as in Paired Comparison (PC) pattern models (Dittrich et al. 2007; Sullivan and Artino 2013). The basic idea of transforming Likert-scale data into paired comparisons is simple: for any two Likert items j and k , if the response to the first item is higher on the numeric scale than the second ($j > k$), then the first item is preferred and the transformed response is $y_{jk} = 1$. Conversely, if item k is preferred ($k > j$) due to a lower score, $y_{jk} = -1$. If responses are equal, it results in an undecided preference ($y_{jk} = 0$). Modelling the joint distribution of these transformed random variables involves assumptions about the probabilistic mechanism, correspondence between Likert responses and derived paired comparison patterns (Dittrich et al. 2002), and potential subject and item covariates. However, unmeasured characteristics of subjects may also influence responses, leading to the use of random effects models to account for such heterogeneity (Francis et al. 2010; Schauburger and Tutz 2022). Another way to evaluate and compare multiple items or objects is through ranking systems where the emphasis is on the order of preference or importance among the items. In the case of ranking data, subjects are asked to order a set of K objects based on a criterion such as preference, importance, or satisfaction. Subjects observe all items and establish their positions relative to each other, resulting in a hierarchy of items without ties. Four classes of probability models for ranking data are order statistic models, paired comparison models, distance-based models, and multistage models.

Order statistic models extend Thurstone's order statistic model and assume that judges' preferences change based on random utilities associated with each item. The Plackett–Luce ranking model is a common choice, representing items with non-negative real-valued parameters on a ratio scale (Buchholz et al. 2022; Gorantla et al. 2023).

The class of paired comparison models converts ranked data into paired comparison data and then works on them, often using Thurstone and Bradley–Terry models as foundations. Various extensions of these models exist (Maydeu-Olivares and Böckenholt 2005; Dittrich et al. 1998, 2002, 2007).

The distance-based models measure the discrepancy between two rankings using distance measures. The probability of observing a ranking vector is inversely proportional to its distance to the mode of the ranking distribution (*modal ranking*), i.e., the permutation that has the highest probability to be generated. The Mallows ϕ -model is a well-known example parameterized by a modal ranking and a dispersion parameter (Vitelli et al. 2018; Feng and Tang 2022). However, distance-based models do not easily accommodate covariates as they generally depend on numerical feature values and are not well-suited for dealing with categorical or discrete covariates. To use these models with categorical data,

it is often necessary to apply preprocessing techniques, such as one-hot encoding (Cerda et al. 2018), to transform categorical variables into a format that can be effectively used in distance calculations. When the task involves incorporating covariates into clustering or classification activities, it is often more appropriate to explore alternative methods such as regression-based clustering, mixture models, or decision tree-based algorithms like Random Forests or Gradient Boosting Trees (D'Ambrosio and Heiser 2016; Dery and Shmueli 2020; Plaia et al. 2022; Albano et al. 2023). These alternative approaches can naturally handle covariates and offer greater flexibility in capturing intricate relationships between covariates and the target variable.

In the end, multistage models decompose the ranking process into independent stages. For example, the Plackett–Luce model belongs to this class, and it decomposes ranking into stages (Fligner and Verducci 1988; Critchlow and Fligner 1991, 1993).

Like rating data, ranking data can also include information about objects and judges, such as prices, brands, demographics, etc. Various developments exist for incorporating covariates into ranking models, including rank-ordered logit models and models with subject-specific covariates (Schauberger and Tutz 2017; Fok et al. 2012; Shen et al. 2021; Lee and Philip 2010; Yu 2000).

3 Methods

In this section, we reconsider the two distributions introduced by Fasola and Sciandra (2013, 2015), by focusing on their theoretical foundations (see also Ursino and Gasparini 2018; Iannario 2014). Discrete Beta and Shifted Binomial distributions are two different probability distributions that can be useful for modelling different types of data, including rating data and ranking data, under certain conditions. However, it is important to note that these distributions are not universally applicable to all scenarios, and their usefulness depends on the specific characteristics of the data you are dealing with. The Discrete Beta distribution (DBT) is often used to model discrete data that represent proportions or *ratings*, typically with a limited number of categories or levels. It is generally used when data fall within a fixed range or when one wants to model the distribution of ratings or proportions. The Shifted Binomial distribution (SBB) can be useful for modelling *ranking* data, especially when one aims to understand the probability distribution of ranks or positions within a set of items or alternatives. It is typically used when one has a fixed number of items and the goal is to estimate the probability of an item being ranked at a certain position. In summary, the choice between the Discrete Beta distribution and the Shifted Binomial distribution depends on the nature of data and on the specific research question or analysis one is conducting. Discrete Beta is useful for modelling ratings or proportions within a fixed range, while Shifted Binomial is useful for modelling ranking data where one is interested in the distribution of ranks or positions. It is important to carefully consider the characteristics of the data under investigation

and the goals of the analysis when selecting the appropriate distribution for one's modelling needs.

3.1 The Discrete Beta model for rating data

Let us define a *rate* as a response variable R assuming values on finite integer scale $R = \{1, \dots, j, \dots, K\}$. To introduce our first distribution proposal we make the typical assumption that the rate represents an indicator of the extent to which a certain *attribute* is present on a given object, expressed on an ordinal scale (Salzberger 2010; Linacre 2002). For example, raters can be asked to express degree of agreement (disagreement) with some statement or their level of satisfaction (dissatisfaction) with a given service. Clearly, different rates indicate that different raters assign different levels of the attribute to the items, but nothing can be said in terms of the magnitude of such differences. However, the inability to evaluate differences among rates can be traced back to the measurement scale: the continuous attribute is latent, while its discretized counterpart is observed.

Let R^* be the latent attribute and assume it is a random variable with *p.d.f.* $f_{R^*}(r^*, \theta)$ where θ is a vector of unknown parameters. The rate R , namely the discretized counterpart of R^* , is therefore ruled by

$$P(R = j) = P(\gamma_{j-1} < R^* \leq \gamma_j), \quad j = 1, 2, \dots, K,$$

where the γ values are $K + 1$ threshold values. Clearly, the *p.d.f.* for the rate can be defined as

$$P(R = j) = F_{R^*}(\gamma_j, \theta) - F_{R^*}(\gamma_{j-1}, \theta) \quad (1)$$

where $F_{R^*}(\cdot)$ is the distribution function of R^* . Making inference on such models translates into estimation of both the γ 's and the parameter vector θ .

The traditional approach used to estimate such models consists of choosing an underlying probability distribution ruling R , fixing the parameter vector θ , assuming $\gamma_0 = -\infty$ and $\gamma_K = +\infty$, and estimating the unknown thresholds γ_j , $j = 1, 2, \dots, K - 1$. A typical choice is the standard logistic distribution underlying proportional odds models (Agresti 2011). However such models turn out to be little parsimonious, especially when the rating scale is wide (K is large).

The main idea developed in this paper consists of reversing this approach with the goal of reducing the number of parameters preserving, at the same time, the flexibility of the resulting distribution. Indeed, we propose to fix the threshold values γ_j and estimate the parameter vector θ . In particular, we believe that the Beta distribution $X \sim B(\alpha, \beta)$, $\alpha > 0$, $\beta > 0$ is very appropriate for this scope, in terms of parsimony, flexibility and ease of interpretation. Indeed, its bounded range of variation allows one to fix the $K - 1$ unknown thresholds at evenly spaced values spanning between $\gamma_0 = 0$ and $\gamma_K = 1$ namely $\gamma_j = j/K$, $j = 1, 2, \dots, K - 1$. This will induce specular, flexible shapes as the continuous Beta distribution and, more importantly, since

$$P(R = j) = F_B\left(\frac{j}{K}; \alpha, \beta\right) - F_B\left(\frac{j-1}{K}; \alpha, \beta\right), \tag{2}$$

it requires the estimation of only two parameters as in the most natural competitor MUB model (D’Elia and Piccolo 2005; Piccolo and D’Elia 2008), assuming a mixture of a uniform and a Shifted Binomial distribution. Supposing n individuals have rated an object, we denote by r_i the rate assigned by the i -th individual to the object, being R_i the random variable of interest. Estimates of α and β can be obtained via numerical maximization of the likelihood function:

$$L(\alpha, \beta; \mathbf{r}) = \prod_{i=1}^n \prod_{j=1}^K \left[F_B\left(\frac{j}{K}; \alpha, \beta\right) - F_B\left(\frac{j-1}{K}; \alpha, \beta\right) \right]^{I(r_i=j)}. \tag{3}$$

Alternatively, estimates can be obtained via the method of moments; in particular, denoting the expected value and the variance of R by $E[R|\alpha, \beta]$ and $V[R|\alpha, \beta]$, respectively, estimates can be obtained solving for α and β the system of equations:

$$\begin{cases} E[R|\alpha, \beta] = \bar{r} \\ V[R|\alpha, \beta] = v \end{cases}$$

where \bar{r} and v are the empirical mean and variance of the observed vector of rates, respectively.

The expected value of the Discrete Beta distribution is derived as follows:

$$\begin{aligned} E[R|\alpha, \beta] &= \sum_{j=1}^K j \left[F_B\left(\frac{j}{K}; \alpha, \beta\right) - F_B\left(\frac{j-1}{K}; \alpha, \beta\right) \right] \\ &= \sum_{j=1}^K j F_B\left(\frac{j}{K}; \alpha, \beta\right) - \sum_{j=1}^K j F_B\left(\frac{j-1}{K}; \alpha, \beta\right) \\ &= \sum_{j=1}^K j F_B\left(\frac{j}{K}; \alpha, \beta\right) - \sum_{j=0}^{K-1} (j+1) F_B\left(\frac{j}{K}; \alpha, \beta\right) \\ &= K + \sum_{j=1}^{K-1} j F_B\left(\frac{j}{K}; \alpha, \beta\right) - \sum_{j=1}^{K-1} (j+1) F_B\left(\frac{j}{K}; \alpha, \beta\right) \\ &= K - \sum_{j=1}^{K-1} F_B\left(\frac{j}{K}; \alpha, \beta\right). \end{aligned} \tag{4}$$

Similarly, for the second moment,

$$\begin{aligned}
 E[R^2|\alpha, \beta] &= K^2 + \sum_{j=1}^{K-1} j^2 F_B\left(\frac{j}{K}; \alpha, \beta\right) - \sum_{j=1}^{K-1} (j+1)^2 F_B\left(\frac{j}{K}; \alpha, \beta\right) \\
 &= K^2 - \sum_{j=1}^{K-1} (1+2j) F_B\left(\frac{j}{K}; \alpha, \beta\right),
 \end{aligned} \tag{5}$$

so that the variance is

$$V[R|\alpha, \beta] = K^2 - \sum_{j=1}^{K-1} (1+2j) F_B\left(\frac{j}{K}; \alpha, \beta\right) - E_R^2(\alpha, \beta). \tag{6}$$

This expression turns out to be difficult to handle mathematically, so that estimation via the method of moments requires the application of numerical algorithms like Newton–Raphson. Because of the mathematical complexity of the expected value and the variance of the discrete distribution, it turns out to be more convenient to base inference on the summary measures $E[B|\alpha, \beta]$ and $V[B|\alpha, \beta]$ of the underlying Beta; in particular, a commonly used reparametrization when dealing with the Beta distribution is

$$\eta = \text{logit}\left(\frac{\alpha}{\alpha + \beta}\right), \quad \gamma = -\text{logit}\left(\frac{1}{1 + \alpha + \beta}\right). \tag{7}$$

Such reparametrization allows for unconstrained estimation since both η and γ belong to $(-\infty, +\infty)$. Parameter η is directly linked to the expected value of the underlying Beta distribution, and can therefore be considered a *liking indicator*, since we expect higher rates as $E[B|\alpha, \beta] = \frac{\alpha}{\alpha + \beta}$ increases. Similarly, since $\gamma = \log(\alpha + \beta)$ and

$$\alpha + \beta = \frac{E[B|\alpha, \beta](1 - E[B|\alpha, \beta])}{V[B|\alpha, \beta]} - 1,$$

the parameter γ is inversely linked to the variance of the underlying Beta random variable, and can be considered as an *agreement indicator* (the degree of agreement between individuals in rating the item). Finally, a covariate vector \mathbf{x}_i can also be introduced in the model:

$$\eta_i = \mathbf{x}_i^T \boldsymbol{\lambda}, \quad \gamma_i = \mathbf{x}_i^T \boldsymbol{\omega}, \tag{8}$$

leading to the same structure and interpretation as in usual logit or log-linear models. For example, with a unique explanatory variable x in the linear predictor, we have

$$\lambda = \log \left[\frac{\alpha(x_i + 1)/\beta(x_i + 1)}{\alpha(x_i)/\beta(x_i)} \right]$$

which resembles a log-odds ratio and

$$\omega = \log \left[\frac{\alpha(x_i + 1) + \beta(x_i + 1)}{\alpha(x_i) + \beta(x_i)} \right]$$

which resembles a log-rate ratio. Of course, alternative reparameterizations are possible.

These two quantities, expressed as the logarithm of the odds ratio and log-rate ratio, respectively, are often used to represent data with a wide range of values or to make comparisons more meaningful. In particular, attention should focus on the sign of the logarithm: if the logarithm is positive, it indicates a surplus or an increase compared to the reference odds; if the logarithm is negative, it signifies a deficit compared to the reference odds. To achieve a more meaningful interpretation, always compare the odds to a reference value. In some cases, one might be interested in the absolute value of the logarithm of the odds ratio, especially if the aim is to measure the intensity of the change regardless of its direction (increase or decrease).

3.2 The Shifted Beta-Binomial model for ranking data

Let’s now define a *rank* as a response variable R still assuming values on the finite integer scale $R = \{1, \dots, j, \dots, K\}$. Differently from rates, ranking data arise when n individuals are asked to order a set of K objects from the most to the least preferred. When focusing on a single object, the response variable R is still given by a random integer between 1 and K , assuming that ties cannot occur.

To introduce our second distribution proposal, we assume a paired comparison perspective (Bradley 1916): the unconscious mechanism underlying ranking of item k consists of the execution of $K - 1$ comparisons between item k and the others (Marden 1996). Such a frame involves a vector of random variables

$$W_j = \begin{cases} 1 & \text{if item } j \text{ is preferred over item } k \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\text{Prob} \{W_j = 1\} = \psi \quad \forall j \neq k.$$

By construction, probability ψ can be considered as a *disliking indicator* for item k . Note that, if W_j and $W_{j'}$ are mutually independent (Dittrich et al. 1998) for each $j' \neq j$, it follows

$$R = \sum_{j \neq k} W_j + 1$$

and ranks assigned to the item follow the *Shifted Binomial* distribution (Li and Chen 2023; Oh 2014; D’Elia 2000). However, as for the rate distribution, there is a simple way to obtain a more flexible distribution using a continuous Beta. Indeed, from the basics of Bayesian statistics, if we assume the Shifted-Binomial to be just the

conditional distribution of R given $\Psi = \psi$, and $\Psi \sim B(\alpha, \beta)$, the marginal distribution of R will be given by

$$\text{Prob}(R = j) = \binom{K-1}{j-1} \frac{B(\alpha + j - 1, \beta + K - j)}{B(\alpha, \beta)}, \quad (9)$$

namely R will follow a *Shifted Beta-Binomial* distribution ($B(\cdot)$ is the *Beta* function). Note that, for $\alpha = 1$, this reduces to the *Inverse Hypergeometric* distribution discussed by D'Elia (2003) and Ouimet (2023).

Supposing n individuals have ranked K objects, we denote by r_i the rank assigned by the i -th individual to a given object of interest, R_i being the relevant random variable. Once again, estimates of α and β can be obtained via numerical maximization of the likelihood function:

$$L(\alpha, \beta; \mathbf{r}) = \prod_{i=1}^n \prod_{j=1}^K \left[\binom{K-1}{j-1} \frac{B(\alpha + j - 1, \beta + K - j)}{B(\alpha, \beta)} \right]^{I(r_i=j)}.$$

Alternatively, the system of equations to solve when applying the method of moments is now linear in α and β , leading to explicit solutions. Indeed, the expected value and the variance of R are, respectively,

$$E[R|\alpha, \beta] = (K-1) \frac{\alpha}{\alpha + \beta} + 1$$

and

$$V[R|\alpha, \beta] = \frac{(E[R|\alpha, \beta] - 1)(K - E[R|\alpha, \beta])}{K - 1} \frac{\alpha + \beta + K - 1}{\alpha + \beta + 1},$$

leading to (see Appendix)

$$\hat{\beta} = \frac{v - (K - \bar{r})(\bar{r} - 1)}{(\bar{r} - 1) - v(K - 1)/(K - \bar{r})} \quad (10)$$

and

$$\hat{\alpha} = \frac{\bar{r} - 1}{K - \bar{r}} \hat{\beta}. \quad (11)$$

Once again, inference can be based on the summary measures of an underlying Beta, using the same reparameterization as for the Discrete Beta distribution, and also covariates can be included in the model. However, differently from the rate distribution, η should now be considered as a *disliking indicator*, since we expect higher (worse) ranks as $E[B|\alpha, \beta]$ increases.

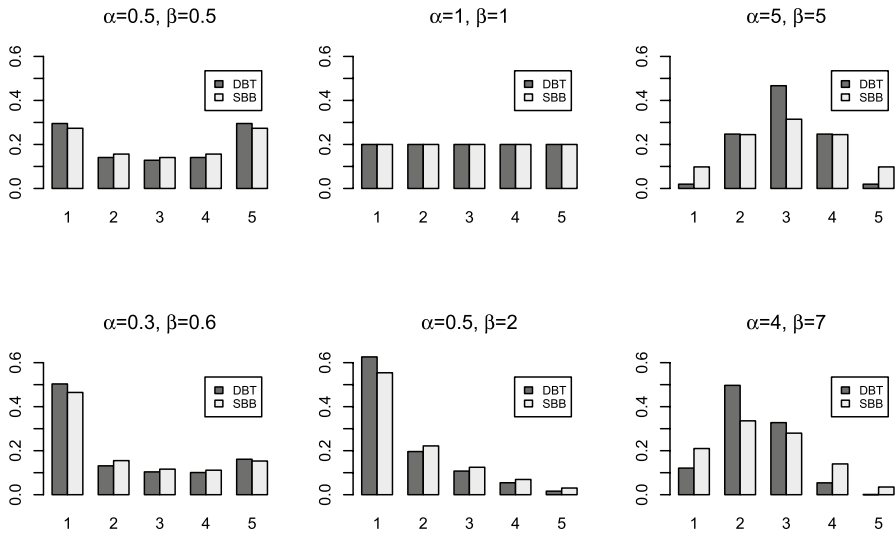


Fig. 2 Shapes of the Discrete Beta (DBT) and the Shifted Beta-Binomial (SBB) distributions for different values of α and β

3.3 Shape comparisons

In order to compare the two proposed distributions in terms of shapes, we simulated data from six different scenarios obtained by varying α and β . Figure 2 represents six double bar plots corresponding to the two proposed distributions in the different scenarios considered.

When $\alpha = \beta = 1$ the two distributions coincide and follow a discrete uniform distribution $U(K)$, where the finite number of values are equally likely to happen. When $\alpha = \beta$, the two distributions are symmetric. When $\alpha \neq \beta$, several shapes can be obtained by properly changing the parameters. Specifically, when $\alpha < \beta$, both distributions display positive skewness, intensifying as the gap between the two parameters widens. Conversely, when $\alpha > \beta$, negative skewness characterizes both distributions.

The distributions take on a “U” or “J” shape when α and β are lower than 1, with the probability mass becoming more concentrated in the tails as these parameters decrease. In contrast, a concave shape is observed when α and β exceed 1, concentrating the probability mass around the central category as both parameters increase.

Regardless of the specific shape attained under varying parameter values, the SBB distribution maintains, compared to the DBT, a higher probability mass in rare or extreme events. The resulting flexibility is an important strength of the proposed distributions because it allows one to deal with differently shaped distributions.

The space of admissible expected values, $E(R)$, and variances, $V(R)$, for the Discrete Beta, the Shifted Beta-Binomial and the competitor MUB model are graphically displayed in Fig. 3. Parametric spaces are obtained by solving the method of moment equations for particular pairs of true values of $E(R)$ and $V(R)$. If the pair

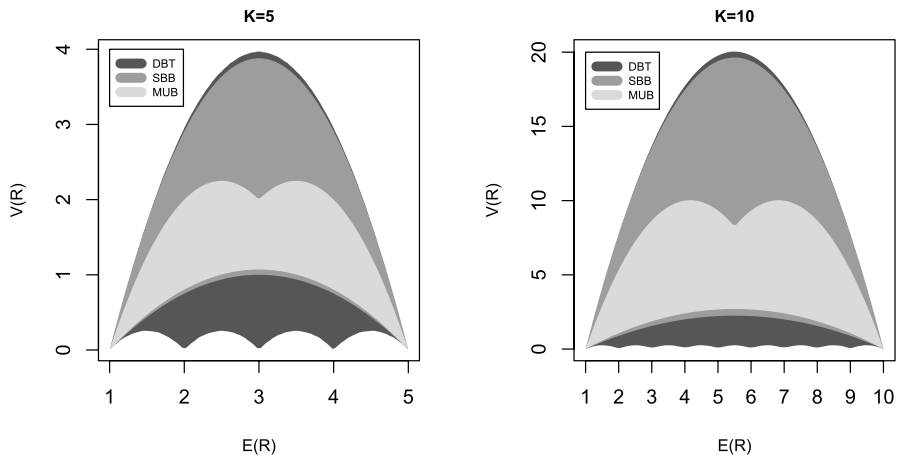


Fig. 3 Space of admissible expected values, $E(R)$, and variances, $V(R)$, for the Discrete Beta (DBT), the Shifted Beta-Binomial (SBB), and the Mixture of Uniform and Binomial (MUB)

of $E(R)$ and $V(R)$ is outside the coloured regions, the equations deriving from the method of moments return not admissible α and β values (negative or infinite); on the contrary, the inner areas lead to admissible pairs of parameters. The comparison is carried out for two different numbers of total objects ($K = 5$ and $K = 10$).

The parametric space of the MUB distribution (light grey area) is a subset of the parametric space of the SBB distribution (darker grey), and, in turn, the parametric space of the SBB distribution is a subset of the parametric space of the DBT distribution (black area). More interestingly, the parametric spaces confirm the main advantages of the two proposed distributions: they overcome the problem of dealing with non-monotonic convex shapes. For example, when $E(R)$ takes the mid-span value and $V(R)$ takes a value lower than the variance of the Uniform distribution $\left(\frac{K^2-1}{12}\right)$, then a corresponding MUB distribution does not exist, while this is not true for the two proposed distributions.

4 Case studies

In this section, the performance of the two proposed distributions is shown through three real examples concerning the attitude of different people towards climate change and pollution.

Human activities are damaging the environment in different ways, from climate change to endangered animal species. Although data indicate an increase of the average global temperature in the last decades and despite the growing number of extreme events like floods, violent storms or heat waves, not everybody believes that climate change is taking place. Analysing data about people's perception of climate change is then important to understand how information

on this theme is conveyed and how much of a priority action is perceived by governments to tackle the phenomenon.

The data used in the first example are ratings and come from a survey carried out by Eurobarometer in 2009, which provides information about Europeans’ opinions on climate change. In the other applications, we focused on rankings, showing a model without covariates applied on a data set containing answers to an online survey about climate change and marine ecosystems conducted at the end of 2021, and a model including covariates for ranking data about air pollution collected in Utah between 2018 and 2020.

4.1 Eurobarometer rating data

The Eurobarometer rating data contain the results of a survey on the attitudes of Europeans’ towards climate change which was carried out in late August and September 2009. This survey investigated the opinion of Europeans on several climate change-related topics, and subjects were asked to answer different questions about their perception of climate change using an ordinal scale. Among them, we focused on the following question:

How serious a problem do you think climate change is at this moment? Please use a scale from 1 to 10, with “1” meaning it is “not at all a serious problem” and “10” meaning it is “an extremely serious problem”.

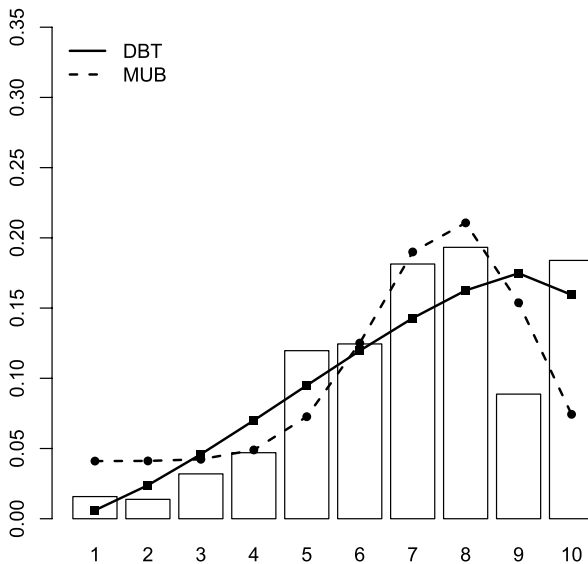


Fig. 4 Observed and fitted distributions of the rates assigned to the seriousness of climate change by participants in the 2009 Eurobarometer survey

The Discrete-Beta model is here applied to the rates assigned by $n = 25,862$ individuals. The observed rate frequencies are: 1 = 408, 2 = 357, 3 = 826, 4 = 1216, 5 = 3095, 6 = 3219, 7 = 4691, 8 = 4998, 9 = 2295, 10 = 4757. Figure 4 shows the empirical distribution of the rates. Due to the non-monotonic convex shape, we believed appropriate to hypothesize a random distribution based on an underlying Beta. We also estimated a MUB model, and the fitted values provided by the two models are superimposed in Fig. 4.

The DBT yields $\hat{\alpha} = 2.33$ and $\hat{\beta} = 1.18$ through maximum likelihood estimation (MLE). The estimated ratio $\frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta}}$ is 0.66, which compared with the intermediate scenario corresponding to a Uniform distribution (0.5), indicates a general interest towards the climate change issue. Additionally, the estimated variance, which is equal to 4.96, reflects the degree of disagreement among the respondents in their ratings, where the maximum variance is obtained in the case of complete polarization of the votes. In comparison, if we were to assume uniformly distributed frequencies, yielding $\hat{\alpha} = \hat{\beta} = 1$, the variance would be $\frac{K^2 - 1}{12} = 8.25$. The observed variance of 4.96 in this case suggests that, while there is a moderate level of agreement among the raters, opinions on the seriousness of climate change vary to a certain extent among the surveyed individuals. Furthermore, it is crucial to emphasize the observed shape of the ratings distribution, which exhibits a convex pattern, and this explains why the MUB model has therefore lower fitting performance in this case ($AIC_{DBT} = 108537.6, AIC_{MUB} = 111533.8$).

4.2 Online survey about climate change and coastal ecosystems

The online survey was carried out within the European Commission H2020-funded research project on “Marine Coastal Ecosystems Biodiversity and Services in a Changing World” (MaCoBioS, Fonseca et al. 2023). The questionnaire consists of 20 questions about perceptions of climate change and the health of coastal ecosystems. It was administered from November 2021 to February 2022 on the Qualtrics platform and response was voluntary, with a total of 709 valid responses. The volunteers were asked

Which three of the following key goals should be of greatest priority for governments to address?

- (i) Protecting natural resources and ecosystems
- (ii) Improving human health and healthcare systems
- (iii) Improving management of agriculture and forestry
- (iv) Improving education opportunities
- (v) Managing the environment to improve human health
- (vi) Addressing climate change
- (vii) Reducing economic insecurity and inequality
- (viii) Improving management of marine fisheries.

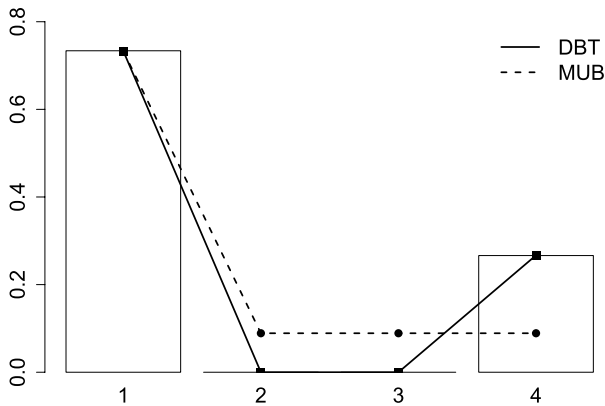


Fig. 5 Observed and fitted distributions of the ranks assumed by protecting natural resources in the ranking vectors observed on 709 respondents

To accommodate respondents' preferences, the survey allowed for the expression of incomplete rankings. In such cases, options that were not selected were treated as ties and placed at the bottom of the ranking. In this way, the unranked items are considered equally preferred or, equivalently, it is assumed that respondents have no preference among them (Li et al. 2019).

Our specific focus in this analysis centered on the preferences expressed towards “i) Protecting natural resources and ecosystems.” We examined the observed ranks and fitted values obtained from two models: the SBB model and the MUB model without the inclusion of covariates. The results of this comparison are displayed in Fig. 5.

Evidently, the SBB model demonstrated a superior fit compared to the MUB model, as confirmed by the respective AIC values: $AIC_{DBT} = 825.5642$ and $AIC_{MUB} = 1240.8397$. This outcome arises from the specific nature of this scenario. The data displayed a distinct U-shaped pattern, which signifies polarized preferences among the survey respondents. The pronounced U-shaped pattern in the data suggests that respondents had clear and divergent views on the importance of protecting natural resources and ecosystems, with a significant portion expressing strong support while others held opposite opinions. Such polarized preferences often pose a challenge for models like the MUB, which are designed to handle monotonic, gradual shifts in preferences, making it less suitable for handling scenarios with sharp and non-monotonic changes in rankings. In cases where preferences exhibit complexity, with a substantial concentration of rankings at both extremes as seen in this dataset, the SBB model provides a notably superior fit to the data.

Furthermore, both $\hat{\alpha}$ and $\hat{\beta}$ are close to 0, obtained through MLE, while the estimated ratio $\frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta}}$, is computed to be 0.27. This ratio is lower than the intermediate scenario (0.5), indicating an attitude in recognizing the importance of protecting natural resources, reflecting that a wider proportion of respondents held a preference for this goal in comparison to other options presented in the survey.

4.3 Utah air quality risk and behavioral action survey

The survey was administered to Utahns between November 2018 and January 2020. It includes more than 60 questions about demographics, daily habits, opinions about air pollution and attitudes towards government interventions to reduce it (Benney et al. 2020). The sample is made up of 1160 subjects and we analysed the answers to the question:

Please rank the following by which causes the most air pollution in Utah:

- (i) Wood burning
- (ii) Automobile exhaust
- (iii) Buildings (e.g. businesses and homes)
- (iv) Major industries (e.g. mining, airport, energy)
- (v) Home chemicals (e.g. aerosols, paints)
- (vi) Environment (e.g. wind blowing dust)
- (vii) Agriculture (e.g. farm equipment, animal byproducts, etc.)
- (viii) Government (e.g. offices, agencies, industries)

We focused on the rankings assigned to the fourth item, “Major industries (e.g. mining, airport, energy)” and investigated if they depend on some explanatory variables, specifically income and political ideology. Both are categorical variables, the former with eight categories (< 25,000\$, 25,000–35,000\$, 35,000–50,000\$, 50,000–75,000\$, 75,000–100,000\$, 100,000–150,000\$, ≥150,000\$, and Not Declared) and the latter with five categories (Democrat, Independent, Libertarian, Republican and Not Declared).

Figure 6 illustrates the observed ranks and fitted values derived from both the SBB model and the MUB model without covariates. Notably, the SBB model demonstrates a superior fit compared to the MUB model, as supported by the AIC values: $AIC_{DBT} = 3789.732$ and $AIC_{MUB} = 3922.127$.

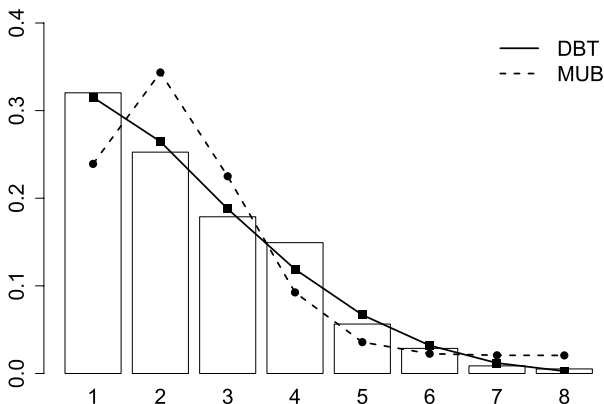


Fig. 6 Observed and fitted distributions of the ranks assumed by protecting natural resources in the ranking vectors observed on 1160 respondents

Table 1 Estimates, standard errors (S.E.) and p-values of the model coefficients for the disliking indicator using Utah survey data

	Estimate	S.E.	$Pr(z > z)$
Intercept	- 1.511	0.109	$< 2e^{-16}***$
[25,000, 35,000)	0.269	0.138	0.051
[35,000, 50,000)	0.168	0.137	0.219
[50,000, 75,000)	- 0.147	0.119	0.217
[75,000, 100,000)	0.241	0.136	0.077
[100,000, 150,000)	0.228	0.146	0.119
$\geq 150,000$	0.351	0.167	0.035*
Income not declared	0.177	0.190	0.351
Independent	0.019	0.113	0.865
Libertarian	- 0.360	0.319	0.259
Republican	0.194	0.090	0.030*
PO not declared	0.203	0.119	0.089

*p = 0.01–0.05; **p = 0.001–0.01; ***p < 0.001

Table 2 Estimates, standard errors (S.E.) and p-values of the model coefficients for the accuracy indicator using Utah survey data

	Estimate	S.E.	$Pr(z > z)$
Intercept	1.767	0.260	$< 2e^{-16}***$
[25,000, 35,000)	0.037	0.295	0.900
[35,000, 50,000)	0.479	0.347	0.167
[50,000, 75,000)	1.409	0.433	0.001**
[75,000, 100,000)	0.066	0.298	0.825
[100,000, 150,000)	0.084	0.312	0.788
$\geq 150,000$	0.817	0.584	0.162
Income not declared	0.515	0.514	0.317
Independent	0.377	0.291	0.195
Libertarian	0.617	0.629	0.327
Republican	0.051	0.264	0.847
PO not declared	0.693	0.283	0.014*

*p = 0.01–0.05; **p = 0.001–0.01; ***p < 0.001

We included the variables “Income” and “Political Orientation (PO)” into the linear predictor additively. Table 1 summarizes the output for η (the disliking indicator), while Table 2 summarizes the output for γ (the agreement indicator) for the selected model; standard error estimates are derived from a numerical (observed) information matrix. When considering the question of air pollution responsibility, the alternatives in first positions are typically seen as the most accountable. In this context, the disliking indicator can be interpreted similarly to the rating framework, as it reveals aversion patterns towards these alternatives. Essentially, the alternatives receiving greater disliking are positioned at the bottom of the rankings, indicating that they are perceived as less accountable. Subsequently, the accuracy indicator quantifies the extent of consensus or divergence in these perceptions among various demographic groups.

The intercept value of -1.511 pertains to individuals identifying as Democrats with an annual income below \$25,000, and it is statistically significant. Notably, the estimated ratio $\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}$, is 0.18, markedly lower than 0.5. This finding suggests a pronounced inclination for this demographic group to attribute a higher level of responsibility to major industries for air pollution compared to other options.

Regarding the influence of other political orientations, individuals who identify as Republicans exhibit less aversion to major industries compared to Democrats. This difference is reflected in the ratio $\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}$, which increases by 0.194 on a logit scale among Republicans. This implies that Republicans tend to place major industries closer to the end of the rankings compared to Democrats, indicating that they perceive it as less responsible for air pollution.

Similarly, a pattern emerges among individuals declaring an annual income exceeding \$150,000, where the ratio $\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}$ increases by 0.351 on a logit scale. Consequently, those with incomes above \$150,000 per year are less likely to perceive major industries as the primary culprits responsible for air pollution.

As regards the accuracy indicator, it measures the level of agreement among respondents in ranking major industries as a cause of air pollution. The accuracy coefficient for a person who identifies as a Democrat with an annual income of less than \$25,000, is 1.767. This indicates a high level of agreement within this demographic group regarding major industries' responsibility for air pollution in Utah. Moreover, the agreement between respondents significantly increases moving from the group earning less than \$25,000 to the group of those earning between \$50,000 and \$75,000. In this transition, the accuracy indicator experiences a notable increase on the logit scale of 1.409.

5 Conclusions

In this paper, we introduce two novel and flexible probability distributions designed for modeling rating and ranking data. These models offer a straightforward interpretation in terms of individuals' "liking" or "disliking" feelings toward specific items, as well as the "agreement" rate for the corresponding distribution of rates or ranks.

One of the notable advantages of the proposed distributions lies in their simplicity and adaptability. Their simplicity relies on the fact that these distributions have only two parameters, while their applicability makes them particularly useful when the distributions assume "U" and "J" (not-monotonic convex) shapes, but we have shown that they also represent a good alternative to the MUB model (the most natural competitor) for concave, monotonic, and uniform distributions. Moreover, they readily accommodate the inclusion of covariates in the models, allowing for a more nuanced understanding of how various factors influence individuals' preferences and rankings. As our analysis has shown, both proposed distributions yield very similar inferential results and interpretations. Consequently, researchers can readily switch between these models to suit various types of data, whether they pertain to rankings or ratings.

Looking ahead, future research endeavors could focus on conducting extensive simulation studies to delve deeper into the fitting potential of the proposed models and their competitors. Moreover, a specific focus will be directed toward studying how to extend the proposed distributions to account for tied rankings. Such studies would further enhance their applicability in real-world scenarios.

Appendix

Let v denote $\text{var}(\mathbf{r})$.

$$\left\{ \begin{aligned} (K - 1) \frac{\alpha}{\alpha + \beta} + 1 &= \bar{r} \\ \frac{\left[(K - 1) \frac{\alpha}{\alpha + \beta} \right] \left[(K - 1) \frac{\beta}{\alpha + \beta} \right]}{K - 1} \frac{\alpha + \beta + K - 1}{\alpha + \beta + 1} &= v \end{aligned} \right.$$

Focusing on the first equation:

$$\begin{aligned} (K - 1)\alpha &= (\bar{r} - 1)(\alpha + \beta) \\ (K - \bar{r})\alpha &= (\bar{r} - 1)\beta \\ \alpha &= \frac{\bar{r} - 1}{K - \bar{r}}\beta. \end{aligned}$$

Focusing on the second equation:

$$\begin{aligned} (K - 1) \left[\frac{\frac{\bar{r}-1}{K-\bar{r}}\beta}{\frac{\bar{r}-1}{K-\bar{r}}\beta + \beta} \right] \left[\frac{\beta}{\frac{\bar{r}-1}{K-\bar{r}}\beta + \beta} \right] \frac{\frac{\bar{r}-1}{K-\bar{r}}\beta + \beta + K - 1}{\frac{\bar{r}-1}{K-\bar{r}}\beta + \beta + 1} &= v \\ (K - 1) \left[\frac{\frac{\bar{r}-1}{K-\bar{r}}\beta}{\frac{K-1}{K-\bar{r}}\beta} \right] \left[\frac{\beta}{\frac{K-1}{K-\bar{r}}\beta} \right] \frac{\frac{K-1}{K-\bar{r}}\beta + K - 1}{\frac{K-1}{K-\bar{r}}\beta + 1} &= v \\ \frac{(\bar{r} - 1)(K - \bar{r})}{K - 1} \frac{\frac{K-1}{K-\bar{r}}\beta + K - 1}{\frac{K-1}{K-\bar{r}}\beta + 1} &= v \\ \frac{(\bar{r} - 1)(K - \bar{r})}{K - 1} \frac{(K - 1)\beta + (K - 1)(K - \bar{r})}{(K - 1)\beta + (K - \bar{r})} &= v \\ \frac{(\bar{r} - 1)(K - \bar{r})(K - 1)\beta + (K - 1)(K - \bar{r})^2(\bar{r} - 1)}{(K - 1)^2\beta + (K - 1)(K - \bar{r})} &= v \\ \frac{v(K - 1)(K - \bar{r}) - (K - 1)(K - \bar{r})^2(\bar{r} - 1)}{(\bar{r} - 1)(K - \bar{r})(K - 1) - v(K - 1)^2} &= \beta \\ \frac{v - (K - \bar{r})(\bar{r} - 1)}{(\bar{r} - 1) - v(K - 1)/(K - \bar{r})} &= \beta. \end{aligned}$$

Acknowledgements The authors are grateful to Vito Muggeo and Gianfranco Lovison for their useful suggestions and comments. This research has been partially supported by the European Union - Next-GenerationEU - National Sustainable Mobility Center CN00000023, Italian Ministry of University and Research Decree n. 1033- 17/06/2022, Spoke 2, CUP B73C2200076000.

Author Contributions Conceptualization: MS and SF; Methodology: MS and SF; Data curation: AA and CDM; Formal analysis: AA and CDM; Visualization: AA and CDM; Writing—original draft: MS, SF, AA, CDM; Writing—review & editing: MS, SF, AA, CDM, AP; Supervision: Antonella Plaia.

Funding Open access funding provided by Università degli Studi di Palermo within the CRUI-CARE Agreement. This research has been partially supported by the European Union - NextGenerationEU - National Sustainable Mobility Center CN00000023, Italian Ministry of University and Research Decree n. 1033- 17/06/2022, Spoke 2, CUP B73C2200076000.

Data availability Eurobarometer survey data can be found at https://data.europa.eu/data/datasets/s703_72_1_ebs322?locale=en. Data of the online survey about climate change and coastal ecosystems are available in the Mendeley repository <https://data.mendeley.com/datasets/t82xdzpdh8/2>. Data from the Utah Air Quality Risk and Behavioral Action Survey are freely available in the ICPSR repository <https://www.openicpsr.org/openicpsr/project/117904/version/V1/view?path=/openicpsr/117904/fcr:versions/V1&type=project>.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agresti A (2010) Analysis of ordinal categorical data, vol 656. Wiley, Hoboken
- Agresti A (2011) Categorical data analysis. Springer, Berlin
- Albano A, Sciandra M, Plaia A (2023) A weighted distance-based approach with boosted decision trees for label ranking. *Expert Syst Appl* 213:119000. <https://doi.org/10.1016/j.eswa.2022.119000>
- Alvo M, Philip L (2014) Statistical methods for ranking data, vol 1341. Springer, New York
- Benney T, Chaney R, Singer P, Sloan C (2020) Utah air quality risk and behavioral action survey. Inter-university Consortium for Political and Social Research [distributor], Ann Arbor. <https://doi.org/10.3886/E117904V1>
- Bradley RA (1976) A biometrics invited paper. *Science, statistics, and paired comparisons*. *Biometrics* 32(2):213–239
- Buchholz A, Lichtenberg JM, Benedetto GD, Stein Y, Bellini V, Ruffini M (2022) Low-variance estimation in the Plackett-Luce model via quasi-Monte Carlo sampling. <https://arxiv.org/abs/2205.06024>

- Cerda P, Varoquaux G, Kégl B (2018) Similarity encoding for learning with dirty categorical variables. *Mach Learn* 107(8–10):1477–1494
- Critchlow DE, Fligner MA (1991) Paired comparison, triple comparison, and ranking experiments as generalized linear models, and their implementation on GLIM. *Psychometrika* 56(3):517–533
- Critchlow DE, Fligner MA (1993) Ranking models with item covariates. In: Fligner M, Verducci J (eds) *Probability models and statistical analyses for ranking data*. Springer, New York, pp 1–19
- Critchlow DE, Fligner MA, Verducci JS (1991) Probability models on rankings. *J Math Psychol* 35(3):294–318
- D'Ambrosio A, Heiser WJ (2016) A recursive partitioning method for the prediction of preference rankings based upon Kemeny distances. *Psychometrika* 81(3):774–794
- D'Elia A (2000) A shifted binomial model for rankings. In: Núñez-Antón V, Ferreira E (eds) *Statistical modelling, XV international workshop on statistical modelling*. New trends in statistical modelling. pp 412–416
- D'Elia A (2003) Modelling ranks using the inverse hypergeometric distribution. *Stat Model* 3(1):65–78
- D'Elia A, Piccolo D (2005) A mixture model for preferences data analysis. *Comput Stat Data Anal* 49(3):917–934
- de Rezende NA, de Medeiros DD (2022) How rating scales influence responses' reliability, extreme points, middle point and respondent's preferences. *J Bus Res* 138:266–274. <https://doi.org/10.1016/j.jbusres.2021.09.031>
- Dery L, Shmueli E (2020) BoostLR: a boosting-based learning ensemble for label ranking tasks. *IEEE Access* 8:176023–176032. <https://doi.org/10.1109/ACCESS.2020.3026758>
- Dittrich R, Hatzinger R, Katzenbeisser W (1998) Modelling the effect of subject-specific covariates in paired comparison studies with an application to university rankings. *J R Stat Soc Ser C* 47(4):511–525
- Dittrich R, Hatzinger R, Katzenbeisser W (2002) Modelling dependencies in paired comparison data: a log-linear approach. *Comput Stat Data Anal* 40(1):39–57
- Dittrich R, Francis B, Hatzinger R, Katzenbeisser W (2007) A paired comparison approach for the analysis of sets of Likert scale responses. *Stat Model* 7:3–28
- Dwass M (1957) On the distribution of ranks and of certain rank order statistics. *Ann Math Stat* 28(2):424–431
- Falahee M, MacRae A (1997) Perceptual variation among drinking waters: the reliability of sorting and ranking data for multidimensional scaling. *Food Qual Prefer* 8(5):389–394
- Fasola S, Sciandra M (2013) New flexible probability distributions for ranking data. In: Minerva T, Morlini I, Palumbo F (eds) *SIS CLADAG 2013, 9th scientific meeting of the classification and data analysis group of the Italian Statistical Society*. pp 191–194
- Fasola S, Sciandra M (2015) New flexible probability distributions for ranking data. In: Morlini I, Minerva T, Vichi M (eds) *Advances in statistical models for data analysis*. Springer, Cham, pp 117–124
- Feng Y, Tang Y (2022) On a Mallows-type model for (ranked) choices. *Adv Neural Inf Process Syst* 35:3052–3065
- Fligner MA, Verducci JS (1988) Multistage ranking models. *J Am Stat Assoc* 83(403):892–901
- Fok D, Paap R, Van Dijk B (2012) A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. *J Appl Econom* 27(5):831–846
- Fonseca C, Wood LE, Andriamahefazafy M, Casal G, Chaigneau T, Cornet CC, O'Leary BC (2023) Survey data of public awareness on climate change and the value of marine and coastal ecosystems. *Data Brief* 47:108924. <https://doi.org/10.1016/j.dib.2023.108924>
- Francis B, Dittrich R, Hatzinger R (2010) Modeling heterogeneity in ranked responses by nonparametric maximum likelihood: how do Europeans get their scientific knowledge? *Ann Appl Stat* 4(4):2181–2202
- Gorantla S, Bhansali E, Deshpande A, Louis A (2023) Optimizing group-fair Plackett-Luce ranking models for relevance and ex-post fairness. <https://arxiv.org/abs/2308.13242>
- Harzing A-W, Baldueza J, Barner-Rasmussen W, Barzantny C, Canabal A, Davila A et al (2009) Rating versus ranking: what is the best way to reduce response and language bias in cross-national research? *Int Bus Rev* 18(4):417–432
- Iannario M (2014) Modelling uncertainty and overdispersion in ordinal data. *Commun Stat - Theor Method* 43(4):771–786. <https://doi.org/10.1080/03610926.2013.813044>
- Kemmelmeier M (2016) Cultural differences in survey responding: issues and insights in the study of response biases. *Int J Psychol* 51(6):439–444

- Lee PH, Philip L (2010) Distance-based tree models for ranking data. *Comput Stat Data Anal* 54(6):1672–1682
- Li S, Chen J (2023) Mixture of shifted binomial distributions for rating data. *Ann Inst Stat Math* 75:833–853. <https://doi.org/10.1007/s10463-023-00865-7>
- Li X, Wang X, Xiao G (2019) A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. *Brief Bioinform* 20(1):178–189
- Linacre JM (2002) Optimizing rating scale category effectiveness. *J Appl Meas* 3(1):85–106
- Marden JI (1996) *Analyzing and modeling rank data*. CRC Press, Boca Raton
- Maydeu-Olivares A, Böckenholt U (2005) Structural equation modeling of paired-comparison and ranking data. *Psychol Methods* 10(3):285
- Oh C (2014) A maximum likelihood estimation method for a mixture of shifted binomial distributions. *J Korean Data Inf Sci Soc* 25(1):255–261
- Ouimet F (2023) Deficiency bounds for the multivariate inverse hypergeometric distribution. <https://arxiv.org/abs/2308.05002>
- Ovadia S (2004) Ratings and rankings: reconsidering the structure of values and their measurement. *Int J Soc Res Methodol* 7(5):403–414
- Piccolo D, D’Elia A (2008) A new approach for modelling consumers’ preferences. *Food Qual Prefer* 19(3):247–259
- Plaia A, Buscemi S, Fürnkranz J, Mencía EL (2022) Comparing boosting and bagging for decision trees of rankings. *J Classif* 39:78–99
- Punzo A, Zini A (2012) Discrete approximations of continuous and mixed measures on a compact interval. *Stat Pap* 53(3):563–575
- Salzberger T (2010) Does the Rasch model convert an ordinal scale into an interval scale? *Rasch Meas Trans* 24(2):1273–1275
- Schauberger G, Tutz G (2017) Subject-specific modelling of paired comparison data: a lasso-type penalty approach. *Stat Model* 17(3):223–243. <https://doi.org/10.1177/1471082X17693086>
- Schauberger G, Tutz G (2022) Multivariate ordinal random effects models including subject and group specific response style effects. *Stat Model* 22(5):409–429
- Shen H, Hong L, Zhang X (2021) Ranking and selection with covariates for personalized decision making. *INFORMS J Comput* 33(4):1500–1519
- Sullivan G, Artino A (2013) Analyzing and interpreting data from Likert-type scales. *J Grad Med Educ* 5(4):541–542. <https://doi.org/10.4300/JGME-5-4-18>
- Tourangeau R, Rips LJ, Rasinski K (2000) *The psychology of survey response*. Cambridge University Press, Cambridge
- Ursino M, Gasparini M (2018) A new parsimonious model for ordinal longitudinal data with application to subjective evaluations of a gastrointestinal disease. *Stat Methods Med Res* 27(5):1376–1393. <https://doi.org/10.1177/0962280216661370>
- Villanueva ND, Petenate AJ, Da Silva MA (2005) Performance of the hybrid hedonic scale as compared to the traditional hedonic, self-adjusting and ranking scales. *Food Qual Prefer* 16(8):691–703
- Vitelli V, Sørensen Ø, Crispino M, Frigessi A, Arjas E (2018) Probabilistic preference learning with the Mallows rank model. *J Mach Learn Res* 18(158):1–49
- Wind SA (2020) Do raters use rating scale categories consistently across analytic rubric domains in writing assessment? *Assess Writ* 43:100416. <https://doi.org/10.1016/j.asw.2019.100416>
- Yu PLH (2000) Bayesian analysis of order-statistics models for ranking data. *Psychometrika* 65(3):281–299
- Yu PL, Gu J, Xu H (2019) Analysis of ranking data. *WIREs Comput Stat* 11(6):e1483