A new paradigm for rating data models

Un nuovo paradigma per modelli su dati di preferenza

Domenico Piccolo

Abstract Rating data arise in several disciplines and the class of Generalized Linear Models (GLM) provides a consolidated methodology for their analysis: such structures (and a plethora of variants) model the cumulative probabilities of ordinal scores as functions of subjects' covariates. A different perspective can be adopted when considering that discrete choices as ordinal assessments are the result of a complex interaction between subjective perception and external circumstances. Thus, an explicit specification of the inherent uncertainty of the data generating process is needed. This paradigm has triggered a variety of researches and applications, conveying in the unifying framework of GEneralized Mixtures with uncertainty (GEM) which encompasses also classical cumulative models. Some critical discussions conclude the paper.

Abstract *Dati di preferenza sono presenti in differenti ambiti e la classe dei modelli lineari generalizzati fornisce una metodologia consolidata per la loro analisi. Una prospettiva differente deriva dal considerare le scelte discrete che producono dati ordinali come un processo derivante da (almeno) due componenti: una percezione soggettiva (feeling) ed una ineliminabile indecisione (uncertainty). Cosí, una specificazione esplicita dell'indecisione nel processo generatore dei dati di preferenza si é resa necessaria. Tale paradigma ha generato numerose varianti e sviluppi che includono anche l'approccio piú tradizionale. Il lavoro introduce i modelli* GEM *e si conclude con alcune considerazioni critiche.*

Key words: Scientific paradigm, Rating data, Empirical evidence, Statistical models, CUB models, GEM models

---

Domenico Piccolo
Department of Political Sciences, University of Naples Federico II, Naples, Italy. e-mail: domenico.piccolo@unina.it

1 Introduction

Statistical models are formal tools to describe, understand and predict real phenomena on the basis of some recognizable mechanism to be effectively estimated and tested on empirical data. This approach is ubiquitous in modern research and establishes the basis to advance disputable topics in any field: a model can be falsified and rejected to favour a more sustainable alternative, as it is for any progress in scientific research and human knowledge.

In this perspective, the specification step of a model derives from empirical evidence, rational deductions, analogy and similarities. Given a set of data and contextual information, statistical methods are nowadays sufficiently developed to provide suitable specifications: time series, longitudinal models, qualitative variables, count data, survival measures, continuous or discrete observations, experimental design, reliability and quality control, etc. are settings where consolidated theories and effective methods suggest consistent specifications.

These topics will be discussed in the field of ordinal data modelling, as ratings of objects or activities, opinions towards facts, expressions of preferences, agreement to sentences, judgements, evaluations of items, perceptions, subjective sensations, concerns, worry, fear, anxiety, etc. All these expressions are collected as ratings/scores by means of verbal or pseudo-verbal categories. Such data are available in different fields: Marketing researches, Socio-political surveys, Psychological enquiries, Sensory sciences and Medicine, among others. For these analyses, several approaches are available as log-linear and marginal models [48], contingency tables inference [47], and so on. In fact, the leitmotif of our discussion is that observed rating data are realizations of a (genuine) discrete or of a discretized process derived by an intrinsically continuous (latent) variable [6].

2 The classical paradigm

Emulating the approach introduced for the logistic regression, the paradigm of cumulative models considers the probability of responses less or equal to a rating r as a function of selected regressors. A convenient link is necessary to establish a one-to-one correspondence between a 0-1 measure (the probability) and a real quantity (the value of the regressors); as a consequence, probit, logit, complementary log-log links, etc. are proposed to specify the corresponding models.

The vast quantity of results derived from such procedures represents the dominant paradigm in theoretical and empirical literature about the statistical analysis of rating data. The process of data generation is introduced by invoking -for each subject- a latent variable Y_i^*, taking values along the real line and related to the values $\boldsymbol{t}_i$ of p explanatory subjects' covariates via the classical regression model: $Y_i^* = \boldsymbol{t}_i\boldsymbol{\beta} + \varepsilon_i, i = 1,2,\ldots,n$. For a given m, the relationship with the discrete ordinal variable R_i, defined over the discrete support $\{1,2,\ldots,m\}$, is provided by:

$$\alpha_{r-1} < Y_i^* \le \alpha_r \qquad \Longleftrightarrow \qquad R_i = r\,, \quad r = 1,2,\ldots,m\,,$$

where $-\infty = \alpha_0 < \alpha_1 < \ldots < \alpha_m = +\infty$ are the thresholds (cutpoints) of the continuous scale of the latent variables Y_i^*.

Then, if $\varepsilon_i \sim F_\varepsilon(.)$, the probability mass function of R_i, for $r = 1,2,\ldots,m$, is:

$$Pr(R_i = r) = Pr(\alpha_{r-1} < Y_i^* \le \alpha_r) = F_\varepsilon(\alpha_r - \boldsymbol{t}_i\boldsymbol{\beta}) - F_\varepsilon(\alpha_{r-1} - \boldsymbol{t}_i\boldsymbol{\beta}), \qquad (1)$$

where $Pr(R_i \le r|\boldsymbol{\theta},\boldsymbol{t}_i) = F_\varepsilon(\alpha_r - \boldsymbol{t}_i\boldsymbol{\beta})$ for $i = 1,2,\ldots,n$ and $r = 1,2,\ldots,m$.

This specification requires the knowledge of $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')$ parameters, that is the $(m-1)$ intercept values $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_{m-1})'$ in addition to the explicit covariate parameters $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)'$.

When selecting logistic random variables ε_i, the probability of a single rating turns out to be:

$$Pr(R_i = r|\boldsymbol{\theta},\boldsymbol{t}_i) = \frac{1}{1+\exp(-[\alpha_r - \boldsymbol{t}_i\boldsymbol{\beta}])} - \frac{1}{1+\exp(-[\alpha_{r-1} - \boldsymbol{t}_i\boldsymbol{\beta}])}\,, \qquad (2)$$

for $i = 1,2,\ldots,n$ and $r = 1,2,\ldots,m$. As a consequence of proportionality properties, these models are known as *proportional odds models* (POM) [1, 66].

Cumulative models have been embedded into the GLM perspective [51] and generalized in several directions: multilevel, varying choice of thresholds, multivariate setting, *conditional logits* including both subjects' and objects' covariates [52], variable effect $\boldsymbol{t}_i\boldsymbol{\beta}_j$ as in *stereotype* models [5], or thresholds which depend on covariates as in *partial proportional odds models* [57], and models with dispersion effects as the *location-scale models* [50] or *location-shift models* [67].

Quite often, the interpretation of these models takes advantage of odds and log-odds measures which are quantities easily manageable by Medicine and Biomedical researchers; similarly, plotting devices explain the direction and the effect of significant covariates on the ordinal responses and new graphical solutions have been recently advanced [68]. In order to anchor the estimated results to easier and more interpretable indexes [2, 3] some difficulties have been emphasized.

Indeed, some drawbacks of the classical specification should be noticed:

- Data generating process refers to a latent variable whose unobservable distribution defines the discrete distribution for the observable ratings.
- It is difficult to accept that -in any circumstance- subjects perform choices by considering ratings not greater than a fixed one, whereas it is more common to consider choices as determined by the "stimulus" associated to a single category and its surrounding values. In fact, the relationship (2) is difficult to manage for deriving immediately the effect of a set of covariates on the categorical response.
- If ratings generated by several items have to be clustered, ranked or classified, unconditionally from covariates, the classical setting leads to a saturated model, which implies an arithmetic equivalence between observed and assumed distributions.

A different paradigm has been proposed for rating data [58, 28] which is substantially based on the explicit modelling of the data generating process of the ordinal observations [44].

3 A generating process for rating data

Finite mixtures have been advanced by several Authors for analysing ordinal data (see [62] for a review) and most of them are motivated for improving fitting. They assume a convex combinations of probability distributions belonging to the same class of models or discretize continuous random variables to get appropriate probability mass functions for the observed ratings: in this vein, the reference to the Beta distribution with several variants is frequent: [55, 32, 65, 31, 70], among others. A different proposal arises from stochastic binary search algorithm [8].

In this scenario, psychological rationale related to the choice of ordinal scores leads to the introduction of CUB models [58]. A mixture distribution is specified to model the *propensity* to adhere to a meditated choice (formally described by a shifted Binomial random variable) and to a totally random one (described by a discrete Uniform distribution). Original motivations for the selection of these random variables were mostly based on a simplicity criterion [28]. However, the Binomial random model may be interpreted as a counting process of a sequential selection among the m categories and accounts for the genuine feeling of the response (appendix of [43]). Then, the Uniform distribution has been introduced as the most extreme among all discrete alternatives and accounts for the inherent uncertainty of the choice [39, 34, 63]. Independently, a psychological support to a direct modelling of a discrete component in the ordinal choices has been recently proposed by Allik [4].

Then, for a known $m > 3$, given a matrix $\boldsymbol{T}$ of values for subjects' covariates, a CUB model for the i-th subject implies both stochastic and deterministic relationships defined, respectively, as:

$$\begin{cases} Pr\,(R_i = r \mid \boldsymbol{T}) = \pi_i\, b_r(\xi_i) + (1-\pi_i)\, p_r^U\,, & r = 1,2,\dots,m; \\ \\ \pi_i = \pi_i(\boldsymbol{\beta}) = \dfrac{1}{1+e^{-\boldsymbol{z}_i\boldsymbol{\beta}}}\,; \qquad \xi_i = \xi_i(\boldsymbol{\gamma}) = \dfrac{1}{1+e^{-\boldsymbol{w}_i\boldsymbol{\gamma}}}\,; & i = 1,2,\dots,n\,. \end{cases} \tag{3}$$

We set $b_r\,(\xi_i) = \binom{m-1}{r-1}\xi_i^{m-r}\,(1-\xi_i)^{r-1}$ and $p_r^U = 1/m$, $r = 1,2,\dots,m$, for the shifted Binomial and Uniform probability mass functions introduced to model feeling and uncertainty, respectively. Here, $(\boldsymbol{z}_i', \boldsymbol{w}_i')'$ are the values of the subjects' covariates extracted from the subjects' covariates data matrix $\boldsymbol{T}$. The acronym CUB stands for *C*ombination of a (discrete) *U*niform and (shifted) *B*inomial distribution.

Given the parameterization (3), it should be noticed that:

- the set of values $\boldsymbol{z}_i$ and $\boldsymbol{w}_i$ may be coincident, overlapped or completely different;

- whereas $1-\xi_i$ characterize the Binomial distribution of the mixture and are immediately related to the strength of the feeling component (they involve the modal value of the responses), the uncertainty parameters $1-\pi_i$ are just the weights of the Uniform distribution assumed for the indecision in the responses (thus, they are not involved in the specification of p_r^U);
- although psychological arguments are sufficient for motivating uncertainty as an important component of an ordinal choice, it is possible to show that uncertainty may be also generated when genuine Binomial choices manifest small variations in the feeling parameters.
- CUB models are able to detect a possible relationship between feeling and uncertainty when a common covariate is significant for those components.

Any one-to-one mapping $\mathbb{R}^p \leftrightarrow [0,1]$ between parameters and covariates is legitimate but the logistic function is to be preferred for simplicity and robustness properties [42]. Thus, the relationships:

$$logit\,(1-\pi_i) \,=\, -\boldsymbol{z}_i\boldsymbol{\beta}\,; \quad logit\,(1-\xi_i) \,=\, -\boldsymbol{w}_i\boldsymbol{\gamma}; \qquad i=1,2,\ldots,n\,. \tag{4}$$

immediately relates uncertainty weights and feeling measures to subjects' covariates.

Although model (3) has been introduced with covariates, one may specify a CUB distribution without such a constraint:

- if $\pi = aver(\pi_i)$ and $\xi = aver(\xi_i)$ are some averages of the individual parameters, the parameters $(\pi,\, \xi)$ can be used to compare the responses to different items;
- for a given i-th subject, the features of the implied CUB model conditional to $(\boldsymbol{z}_i,\, \boldsymbol{w}_i)$ may be investigated by letting $\pi_i = \pi$ and $\xi_i = \xi$.

In this way, a CUB probability distribution is defined as:

$$Pr\,(R=r\mid \boldsymbol{\theta}) = \pi\, b_r\,(\xi) + (1-\pi)\; p_r^U\,, \qquad r=1,2,\ldots,m. \tag{5}$$

where $\boldsymbol{\theta} = (\pi,\xi)' \in \Omega\,(\boldsymbol{\theta}) = \{(\pi,\xi):\; 0<\pi\leq 1,\, 0\leq \xi\leq 1\}$ and the parameter space $\Omega\,(\boldsymbol{\theta})$ is the (open left) unit square.

A CUB model is *identifiable* [36] for any $m>3$ whereas $m=3$ implies a saturated model. Not all the well-defined discrete distributions are *admissible* for a CUB model as, for instance, bimodal probability functions. However, if multimodality is a consequence of latent classes then a CUB model with an explanatory variable related to those classes is adequately fitted according to (3).

A qualifying feature of CUB models is their visualization: given the one-to-one correspondence between the probability mass function (5) and a single point in the parameter space $\Omega\,(\boldsymbol{\theta})$, it is immediate to compare different items in terms of feeling and uncertainty; furthermore, by including subjects' covariates as in (3), the effect of each covariate on either components is visualized and response profiles can be identified. Finally, when CUB models estimated for different clusters –specified by time, space, circumstances and/or covariates– the changes in the responses (in terms of feeling and uncertainty) are immediately shown in the same parameter space. In this respect, also more refined devices –as, for instance, Scatter Plot of Estimates (*SPE*)– have been advanced [45].

Inferential issues have been solved by asymptotically efficient maximum likelihood methods performed by EM procedures [59] which are implemented in an R package [45]. Residual analysis may be conducted [29]. Alternative inferential procedures as Bayesian approaches [25] and permutation tests [9] have been also established.

Successful applications of CUB models have been obtained in different fields and for dissimilar objectives [21, 26, 10, 15, 11, 24, 13, 17]. In addition, formal and empirical comparisons with cumulative models have been extensively discussed on both real and simulated data sets [62]. Especially, the performance of the novel class of models has been checked when heterogeneity is a heavy component: in these circumstances, the better performance of CUB model as measured by *BIC* criterion, for instance, is the consequence of parsimony. Moreover, CUB models correctly assign heterogeneity to a possible uncertainty as related to explanatory covariates, whereas in POM heterogeneity is scattered over the categories by means of estimated cutpoints.

In this class of models, feeling causes no interpretation problem since it is immediately related to the attraction towards the item (modal values of the distribution and feeling parameters are strictly related); on the contrary, some doubts concern the nature of the uncertainty component [33], especially when no subjects' covariates are introduced. Surely, the measure conveyed by $1-\pi$ includes at least three meanings:

- *subjective indecision*: the measure $1-\pi_i$ is related to the personal indecision of the *i*-th respondent;
- *heterogeneity*: for CUB model without covariates, $1-\pi$ summarizes heterogeneity of the responses with respect to a given item [12, 16];
- *predictability*: in case of predicting ordinal outcomes, π is a direct measure of predictability of the model with respect to two extremes: a discrete Uniform ($\pi \to 0$) and shifted Binomial ($\pi = 1$) distributions.

Indeed, the special mixture proposed for ordinal data allows for a single rating to convey information on both *feeling* and *uncertainty* parameters:

i) the observed r is directly related to ξ since the probability of high/low values of r may increase or decrease with $1-\xi$;

ii) the observed r has also a bond with π since, for each score r, it increases or decreases the relative frequencies of: $\left|Pr(R=r)-\frac{1}{m}\right| \propto \pi$. Then, modifying the distance from the Uniform situation, each response modifies also the information concerning the uncertainty.

4 The family of CUB models

Previous models have been generalized in several directions to adequately fit rating data and exploit parsimony in different contexts. Among the many variants, we mention the inclusion of *shelter effect* [37], also with covariates (GeCUB) [43], the cases when almost one dominant preference is expressed (CUSH) [16] or some proposal with varying uncertainty (as VCUB [34] and CAUB [63] for *response style effects*) and also by including random effects in the components (HCUB [38] and RCUB [64]) and connections with fuzzy analysis of questionnaires [30].

Then, CUB models have been usefully applied in presence of "don't know" answers (DK-CUB [54, 41]), for detecting latent classes (LC-CUB [35]), for managing missing values in ordered data [24], for considering a MIMIC jointly with a CUB model [19] and for implementing model-based trees [18]. Noticeable is the discussion of similarities with log-linear models [56].

Further significant issues are the multi-items treatments, when also objects' characteristics are considered [61], and some genuine and promising multivariate proposals [7, 22, 23, 20].

An important advancement considers *overdispersion* as a further component to be examined. Thus, the feeling parameter is interpreted as a random variable specified by a Beta distribution: so, a (shifted) Beta-Binomial distribution is considered for the feeling component and CUBE models are introduced [39, 60, 46]. This class is truly general since it includes several previous specifications (as IHG [27], for instance); in addition, it adds higher flexibility to CUB models with an extra parameter. Since CUB are nested into CUBE models, relative testing is immediate.

The importance to take explicitly uncertainty into account in designing statistical models for rating data has received an increasing consideration (as in CUP models: [69]); indeed, the inclusion of classical approach in the emerging paradigms is a desirable opportunity.

In fact, due to the complex functioning of decision making process, two components are definitely dominant: i) a primary attraction/repulsion towards the item which is related to the personal history of the subject and his/her beliefs with respect to the item; ii) an inherent indecision derived by the circumstances surrounding the choice. This argument is grounded on empirical evidence since any choice is a human decision, where taste, emotion, sentiment are substantial issues which manifest themselves as a random component denoted as *feeling*. Moreover, choice changes over time, with respect to the environment, the modality (verbal, visual, numerical, etc.) and the tool (face-to-face response, telephone, questionnaire, online, etc.) by which data are collected: these circumstances cause an inherent indecision denoted as *uncertainty*.

Then, each respondent provides observations which are realizations of a choice R_i generated by mixing the feeling X_i and the uncertainty V_i. Formally, let m be the number of categories and $\boldsymbol{t}_i^{(X)} \in \boldsymbol{T}$, $\boldsymbol{t}_i^{(V)} \in \boldsymbol{T}$, where $\boldsymbol{T}$ includes the values of the selected covariates to explain feeling and uncertainty, respectively.

For any well-defined discrete distributions for feeling X_i and uncertainty V_i, respectively, a *GEneralized Mixture model with uncertainty* (GEM) is defined as follows:

$$Pr\,(R_i = r \mid \boldsymbol{\theta}) = \pi_i\, Pr\left(X_i = r \mid \boldsymbol{t}_i^{(X)}, \boldsymbol{\Psi}\right) + (1-\pi_i)\, Pr\,(V_i = r)\,, \tag{6}$$

for $i = 1,\dots,n$ and $r = 1,\dots,m$. Here, $\pi_i = \pi(\boldsymbol{t}_i^{(V)}, \boldsymbol{\beta}) \in (0,1]$ are introduced to weight the two components and $\boldsymbol{\Psi}$ includes all the parameters necessary for a full specification of X_i. The probability distribution of the *uncertainty* component V_i is assumed known over the support $\{1,\dots,m\}$ on the basis of *a priori* assumptions.

GEM models (6) are a very flexible class of models to interpret and fit rating data since they may specify: main flavour of the responses and their uncertainty/heterogeneity, overdispersion in respondents' feeling, presence of a possible inflated category and different probability distributions for given covariates: a distribution function for a latent variable to be discretized (classical approach); a probability mass function (novel approach). Both of them may include uncertainty in their specification for better fitting and interpretation.

Thus, GEM is the new paradigm since it is an *inclusive class of models for ordinal data* which encompasses both classical and mixture models in a unique representation where distributions are not compelled to belong to the exponential family.

5 Occam's razor: believe it or not

Occam's razor refers to one of the first principle in scientific research: "*Non sunt multiplicanda entia sine necessitate*" (Entities are not to be multiplied without necessity).

Cumulative models without covariates are saturated and this circumstance causes logical problems which have not been considered too much in the literature. The point is that estimable thresholds act as frequency parameters to perfectly fit the observed distribution; the introduction of explanatory variables modifies this perfect fit by adding variability to data. The paradox is that the starting point is a deterministic model and this structure becomes a stochastic one by adding further information.

Now, let $\mathfrak{E}(\mathscr{M} \mid A)$ be the measure of the explicative power of a model $\mathscr{M}$ when the information set is A: it may be log-likelihood, pseudo-R^2, fitting measure, divergence, BIC, deviance, etc. Then, let us denote by $\mathfrak{E}(\mathscr{M} \mid A \cup B)$ the same quantity when the information set is enlarged by a not-empty set B disjoint with A. Of course, $A \subset (A \cup B)$ and $\mathfrak{E}(\mathscr{M} \mid A) < \mathfrak{E}(\mathscr{M} \mid A \cup B)$ for any informative data set B. According to Occam's razor principle, a statistician should prefer a model based on $(A \cup B)$ if and only if the enlarged information improves the explicative power of A, otherwise B is a "*multiplicanda entia sine necessitate*".

In the class $\mathscr{M}$ of cumulative models, let A be the information set consisting of observed ratings and B the set of values of explanatory variables. Then, $\mathfrak{E}(\mathscr{M} \mid A \cup B) < \mathfrak{E}(\mathscr{M} \mid A)$ since any explicative power measure will reach its maximum in case of perfect coincidence of predictions and observations. Thus, under this perspective, the application of these models might appear controversial.

These considerations favour the novel paradigm since in the parametric mixture models modelling it is possible to fit and interpret data by splitting the explicative power of the probability structure from the added contribution given by explanatory covariates. This approach starts with a stochastic model and improves its performance by adding genuine information.

6 Conclusions

The core of the paper is to underline the paradigmatic nature of the modelling approach to rating data. According to Kuhn [49], a paradigm includes "the practices that define a scientific discipline at a certain point in time". Indeed, changes in paradigms have been recently proposed in order to better comprehend the subjective mechanism of a discrete selection out of a list of ordinal categories. The crisis of the established approach is more evident for clustering of items, visualization and interpretation purposes, especially in the big data era. Simulated and real data analysis support the usefulness of the chosen mixture modelling approach as a prominent alternative, thus fostering an integrated rationale.

Although some of them are useful, "all models are substantially wrong" and this aphorism is valid also for a novel paradigm. The substantive problem is to establish the starting point for further advances in order to achieve better models which should ever be improved. As statisticians, we must be aware of the role and importance of uncertainty in human decisions and CUB models may be considered as building blocks of more complex statistical specifications; above all, they act as a benchmark for more refined analyses.

An open-minded research process implies that a new paradigm is to be contrasted by the current one, being able to pursue all previous commitments and even solve new challenges. A convincing proposal easily captures new followers; it may be applied in different circumstances and former paradigms may be more and more critically considered. Unfortunately, since models are questionable by definition and given that human mind has inertial attitude towards novelties (it is a heavy effort to assume a new paradigm in consolidated procedures), the breaking point is often deferred in time.

Probably, time is not ripe yet for a paradigm shift. Nevertheless, a comprehensive family of models with appealing interpretation and parsimony features, a number of published papers supporting the new approach in different fields, an increasing diffusion of models which include uncertainty with a prominent role, the availability of a free software which effectively performs inferential procedures and graphical analysis are convergent signals that the prospective paradigm is slowly emerging.

Acknowledgements: Author thanks Stefania Capecchi, Francesca Di Iorio and Rosaria Simone for stimulating discussions and critical comments which improved the final version of the paper. This work has been partially supported by CUBREMOT project, University of Naples Federico II.

References

1. Agresti, A. (2010). *Analysis of Ordinal Categorical Data*, 2^{nd} edition. Hoboken: Wiley.
2. Agresti, A. and Kateri, M. (2017). Ordinal Probability Effect Measures for Group Comparisons in Multinomial Cumulative Link Models. *Biometrics*, **73**, 214–219.
3. Agresti, A. and Tarantola, C. (2018). Simple ways to interpret effects in modeling ordinal categorical data. *Statistica Neerlandica*, 1–14. DOI: 10.1111/stan.12130.
4. Allik, J. (2014). A mixed-binomial model for Likert-type personality measures. *Frontiers in Psychology*, **5**, 371. Published online 2014 May 9. doi: 10.3389/fpsyg.2014.00371. PMCID: PMC4023022.
5. Anderson, J.A. (1984). Regression and ordered categorical variables. *Journal of the Royal Statistical Society, Series B*, **46**, 1–30.
6. Anderson, J.A. and Philips, P.R. (1981). Regression, discrimination and measurement models for ordered categorical variables. *Journal of the Royal Statistical Society, Series C*, **30**, 22–31.
7. Andreis, F., Ferrari, P.A. (2013) On a copula model with CUB margins, *QdS. Journal of Methodological Applied Statistics*, **15**, 33–51.
8. Biernacki, C. and Jacques, C. (2016). Model-based clustering of multivariate ordinal data relying on a stochastic binary search algorithm. *Statistics and Computing*, **26**, 929–943.
9. Bonnini, S., Piccolo, D., Salmaso, L. and Solmi, F. (2012). Permutation inference for a class of mixture models, *Journal Communications in Statistics - Theory and Methods*, **41**(16-17), 2879–2895.
10. Capecchi, S. (2015). Modelling the perception of conflict in working conditions. *Electronic Journal of Applied Statistics*, **8**(3), 298–311.
11. Capecchi, S., Endrizzi, I., Gasperi, F. and Piccolo, D. (2016). A multi-product approach for detecting subjects' and objects' covariates in consumer preferences. *British Food Journal*, **118**(3), 515–526.
12. Capecchi, S. and Iannario, M. (2016). Gini heterogeneity index for detecting uncertainty in ordinal data surveys. *METRON*, **74**, 223–232.
13. Capecchi, S., Iannario, M. and Simone, R. (2018). Well-being and relational goods: a model-based approach to detect significant relationships. *Social Indicators Research*, **135**(2), 729–750.
14. Capecchi, S. and Piccolo, D. (2014). Modelling the latent components of personal happiness. In: C. Perna and M. Sibillo (eds.), *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, Springer International, Springer-Verlag, pp.49–52.
15. Capecchi, S. and Piccolo, D. (2016). Investigating the determinants of job satisfaction of Italian graduates: a model-based approach. *Journal of Applied Statistics*, **43**(1), 169–179.
16. Capecchi, S. and Piccolo, D. (2017). Dealing with heterogeneity in ordinal responses. *Quality & Quantity*, **51**, 2375–2393.
17. Capecchi, S. and Simone, R. (2018). A proposal for a model-based composite indicators: experience on perceived discrimination in Europe. *Social Indicators Research*, DOI: 10.1007\s11205-018-1849-9.
18. Cappelli, C., Simone, R. and Di Iorio, F. (2017). Growing happiness: a model-based tree, in: A. Petrucci and R. Verde (eds.). "Statistics and Data Science: new challenges, new generations", *Proceedings of the 2017 Conference of the Italian Statistical Society, Florence*, 261–266. ISBN (online): 978-88-6453-521-0.
19. Carpita, M., Ciavolino, E. and Nitti, M. (2018). The MIMIC-CUB model for the prediction of the economic public opinion in Europe, *Social Indicators Research*, 1–19. DOI:10.1007\s11205-018-1885-4.

20. Colombi, R. and Giordano, S. (2016). A class of mixture models for multidimensional ordinal data, *Statistical Modelling*, **16**(4), 322–340.
21. Corduas, M. (2011a) Assessing similarity of rating distributions by Kullback-Liebler divergence. In: Fichet A. et al. (eds.), *Classification and Multivariate Analysis for Complex Data Structures, Studies in Classification, Data Analysis, and Knowledge Organization*. Springer-Verlag, Berlin, Heidelberg, pp.221–228.
22. Corduas, M. (2011b) Modelling correlated bivariate ordinal data with CUB marginals, *Quaderni di Statistica*, **13**, 109–119.
23. Corduas, M. (2015). Analyzing bivariate ordinal data with CUB margins, *Statistical Modelling*, **15**(5), 411–432.
24. Cugnata, F. and Salini, S. (2017). Comparison of alternative imputation methods for ordinal data. *Communications in Statistics. Simulation and Computation*, **46**(1), 315–330.
25. Deldossi, L. and Paroli, R. (2015). Bayesian variable selection in a class of mixture models for ordinal data: a comparative study, *Journal of Statistical Computation and Simulation*, **85**(10), 1926–1944.
26. Deldossi, L. and Zappa, D. (2014). A Novel Approach to Evaluate Repeatability and Reproducibility for Ordinal Data, *Communications in Statistics - Theory and Methods*, **43**(4), 851–866.
27. D'Elia, A. (2003). Modelling ranks using the inverse hypergeometric distribution, *Statistical Modelling*, **3**, 65–78.
28. D'Elia, A. and Piccolo, D. (2005). A mixture model for preference data analysis. *Computational Statistics & Data Analysis*, **49**, 917–934.
29. Di Iorio, F. and Iannario, M. (2012). Residual diagnostics for interpreting CUB models. *STATISTICA*, **LXXII**, 163–172.
30. Di Nardo, E. and Simone, R. (2016). CUB models: a preliminary fuzzy approach to heterogeneity. *Proceedings of the 48th Scientific Meeting of the Italian Statistical Society*, Eds. M. Pratesi and C. Perna (ISBN: 9788861970618), pp. 1–10.
31. Fasola, S. and Sciandra, M. (2015). New Flexible Probability Distributions for Ranking Data. In: *Advances in Statistical Models for Data Analysis*. Springer, Springer-Verlag, pp. 117–124.
32. Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions, *Journal of Applied Statistics*, **31**, 799–815.
33. Golia, S. (2015). On the interpretation of the uncertainty parameter in CUB models, *Electronic Journal of Applied Statistical Analysis*, **8**(3), 312–328.
34. Gottard, A., Iannario, M. and Piccolo, D. (2014). Varying uncertainty in CUB models. *Advances in Data Analysis and Classifications*, **10**, 225–244.
35. Grilli, L., Iannario, M., Piccolo, D. and Rampichini, C. (2014). Latent class CUB models. *Advances in Data Analysis and Classifications*, **8**, 105–119.
36. Iannario, M. (2010). On the identifiability of a mixture model for ordinal data. *Metron*, **LXVIII**, 87–94.
37. Iannario, M. (2012a). Modelling *shelter* choices in a class of mixture models for ordinal responses. *Statistical Methods and Applications*, **21**, 1–22.
38. Iannario, M. (2012b). Hierarchical CUB models for ordinal variables. *Communications in Statistics. Theory and Methods*, **41**, 3110–3125.
39. Iannario, M. (2013). Modelling uncertainty and overdispersion in ordinal data. *Communications in Statistics. Theory and Methods*, **43**, 771–786.
40. Iannario, M., Manisera, M., Piccolo, D. and Zuccolotto, P. (2012). Sensory analysis in the food industry as a tool for marketing decisions. *Advances in Data Analysis and Classification*, **6**(4), 303–321.
41. Iannario, M., Manisera, M., Piccolo, D. and Zuccolotto, P. (2018). Ordinal data models for No-opinion responses in attitude surveys. *Sociological Methods & Research*, **6**(4), 1–27. DOI: 10.1177/0049124118769081.
42. Iannario, M., Monti, A.C., Piccolo, D. and Ronchetti, E. (2017). Robust inference for ordinal response models. *Electronic Journal of Statistics*, **11**, 3407–3445.
43. Iannario, M. and Piccolo, D. (2016a). A generalized framework for modelling ordinal data, *Statistical Methods and Applications*, **25**, 163–189.

44. Iannario, M. and Piccolo, D. (2016b). A comprehensive framework of regression models for ordinal data, *METRON*, **74**, 233–252.
45. Iannario, M., Piccolo, D. and Simone, R. (2018). CUB: a class of mixture models for ordinal data. R package version 1.1.2, http://CRAN.R-project.org/package=CUB.
46. Iannario M. and Simone R. (2017). Mixture models for rating data: the method of moments via Gröebner basis. *Journal of Algebraic Statistics*, **8**(2), 1–28.
47. Kateri, M. (2014). *Contingency Table Analysis: Methods and Implementations Using R*. New York: Birkäuser/Springer.
48. Lang, J.B. (1996). Maximum Likelihood Methods for a Generalized Class of Log-Linear Models, *The Annals of Statistics*, **24**(2), 726–752.
49. Kuhn, T. (1962). *The Structure of Scientific Revolutions* (fourth edition 2012). University of Chicago Press, Chicago.
50. McCullagh, P. (1980). Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society, Series B*, **42**, 109–142.
51. McCullagh, P. and Nelder, J.A. (1989). *Generalized linear models*, 2^{nd} edition. Chapman & Hall, London.
52. McFadden, K. (1978). Modeling the choice of residential location, in: Karlqvist, A. et al. *Spatial Interaction Theory and Residential Location*. Amsterdam: North-Holland, pp.75–76.
53. Manisera, M. and Zuccolotto, P. (2014a). Modeling rating data with Nonlinear CUB models. *Computational Statistics & Data Analysis*, **78**, 100–118.
54. Manisera, M. and Zuccolotto, P. (2014b). Modelling "Don't know" responses in rating scales. *Pattern Recognition Letters*, **45**, 226–234.
55. Morrison, D.G. (1979). Purchase intentions and purchase behavior. *Journal of Marketing*, **43**, 65–74.
56. Oberski, D.L. and Vermunt, J.K. (2015). The relationship between CUB and loglinear models with latent variables. *Electronic Journal of Applied Statistical Analysis*, **8**(3), 374–383.
57. Peterson, B. and Harrell, F.E. (1990). Partial proportional odds models for ordinal response variables. *Applied Statistics*, **39**, 205–217.
58. Piccolo, D. (2003). On the moments of a mixture of uniform and shifted binomial random variables. *Quaderni di Statistica*, **5**, 85–104.
59. Piccolo, D. (2006). Observed information matrix in MUB models. *Quaderni di Statistica*, **8**, 33–78.
60. Piccolo, D. (2015). Inferential issues on CUBE models with covariates. *Communications in Statistics. Theory and Methods*, **44**, 5023–5036.
61. Piccolo, D. and D'Elia, A. (2008). A new approach for modelling consumers' preferences. *Food Quality and Preference*, **19**, 247–259.
62. Piccolo, D., Simone R. and Iannario, M. (2018). Cumulative and CUB models for rating data: a comparative analysis, *International Statistical Review*, forthcoming.
63. Simone R. and Tutz G. (2018). Modelling uncertainty and response styles in ordinal data, *Statistica Neerlandica*, doi: 10.1111/stan.12129
64. Simone, R., Tutz, G. and Iannario, M. (2018). Subjective heterogeneity and uncertainty for ordinal repeated measurements, *preliminary report*.
65. Taverne, C. and Lambert, P. (2014). Inflated Discrete Beta Regression Models for Likert and Discrete Rating Scale Outcomes, *arXiv:1405.4637v1*, 19 May 2014.
66. Tutz, G. (2012). *Regression for Categorical Data*. Cambridge: Cambridge University Press.
67. Tutz, G. and Berger, M. (2017). Separating location and dispersion in ordinal regression models, *Econometrics and Statistics*, **2**, 131–148.
68. Tutz, G. and Schauberger, G. (2010). Visualization of categorical response models: from data glyphs to parameter glyphs. *Journal of Computational and Graphical Statistics*, **22**, 156–177.
69. Tutz G., Schneider M., Iannario M. and Piccolo D. (2017). Mixture Models for Ordinal Responses to Account for Uncertainty of Choice. *Advances in Data Analysis and Classification*, **11**(2), 281–305.
70. Ursino, M. and Gasparini, M. (2018). A new parsimonious model for ordinal longitudinal data with application to subjective evaluations of a gastrointestinal disease, *Statistical Methods in Medical Research*, **27**(5), 1376–1393.