ORIGINAL PAPER

Modeling random and non-random decision uncertainty in ratings data: a fuzzy beta model

Antonio Calcagnì[1] **· Luigi Lombardi**[2]

Abstract

Modeling human ratings data subject to raters' decision uncertainty is an attractive problem in applied statistics. In view of the complex interplay between emotion and decision making in rating processes, final raters' choices seldom reflect the true underlying raters' responses. Rather, they are imprecisely observed in the sense that they are subject to a non-random component of uncertainty, namely the decision uncertainty. The purpose of this article is to illustrate a statistical approach to analyse ratings data which integrates both random and non-random components of the rating process. In particular, beta fuzzy numbers are used to model raters' non-random decision uncertainty and a variable dispersion beta linear model is instead adopted to model the random counterpart of rating responses. The main idea is to quantify characteristics of latent and non-fuzzy rating responses by means of random observations subject to fuzziness. To do so, a fuzzy version of the Expectation–Maximization algorithm is adopted to both estimate model's parameters and compute their standard errors. Finally, the characteristics of the proposed fuzzy beta model are investigated by means of a simulation study as well as two case studies from behavioral and social contexts.

Keywords Fuzzy ratings data · Beta fuzzy numbers · Beta linear model · Fuzzy data analysis · Decision uncertainty · Risk-taking behaviors · Customer satisfaction

Mathematics Subject Classification 62-07 · 62J12 · 62P25

✉ Antonio Calcagnì
antonio.calcagni@unipd.it

[1] DPSS, University of Padova, via Venezia 8, 35031 Padua, Italy

[2] DIPSCO, University of Trento, Trento, Italy

1 Introduction

In social and behavioral research, satisfaction surveys, aptitude and personality testing, demographic inquiries, and life quality questionnaires are widespread tools to collect data involving subjective evaluations, agreements, and judgments. In a typical social survey, a set of questions is administered to a sample of participants and they are asked to express the extent of their agreement on bounded discrete or continuous rating scales (Aiken 1996; Miller and Salkind 2002). Traditionally, ratings data have been collected by means of pencil and paper questionnaires although more flexible implementations, such as online questionnaires, have become increasingly popular. More recently, technological advancements have fostered the development of new rating tools which offer an accurate way to trace the rating process from its beginning to the final rating outcome (Schulte-Mecklenbeck et al. 2011). Hence, unlike standard rating tools, these techniques allow researchers to collect a richer variety of participants' data, including temporal unfolding of ratings and decision uncertainty (Calcagnì and Lombardi 2014; Freeman and Ambady 2010). As several scholars have pointed out, *decision uncertainty* can be referred to a subject-specific cognitive component of the rating process designed to construct a coherent mental representation of the question being rated (Ülkümen et al. 2016). In this sense, it reflects the subjective interplay of decisional and emotional components which contribute to the final rating response (Kahneman and Tversky 1982). As such, when appropriately utilized, this source of within-subject heterogeneity can reveal more about ratings then standard crisp responses. Interestingly, this type of non-random, systematic uncertainty has been extensively studied especially with regards to its effects on rating responses (e.g., see Saal et al. 1980). Indeed, it is widely recognized that ratings data often suffer from lack of accuracy, for instance because of social desirability (Furnham 1986), faking behaviors (Lombardi et al. 2015), personality (Muthukumarana and Swartz 2014), response styles (Eid and Zickar 2007), and violations of rating rules (Iannario 2015; Preston and Colman 2000; Rabinowitz et al. 2019). These issues have not only been recognized as important by applied statisticians working with ratings data but also by several researchers working in fields like applied econometrics (e.g., see Angel et al. 2019; De Bruin et al. 2011; Zafar 2011), metrology (e.g., see Pendrill and Petersson 2016; Pendrill 2014), and risk analysis (e.g., see Slovic et al. 2004).

Modeling ratings data is a relevant problem in applied statistics. Commonly used methods to handle with discrete or continuous rating data include generalized linear models (GLMs) (McCulloch 2000), beta regression models (Ferrari and Cribari-Neto 2004; Migliorati et al. 2018; Ospina and Ferrari 2012), and combination of uniform and shifted binomial (CUBE) models (Golia 2015; Piccolo and Simone 2019; Piccolo et al. 2019). These models typically represent mean and dispersion components as a linear or non-linear function of some external covariates, which are intended to explain the observed heterogeneity of the ratings data. Although some of these methods also allow for disentangling individual indecision and heterogeneity of responses induced by the presence of

subgroups (e.g., CUBE), they are mainly intended to work with data represented as a crisp collection of responses and do not account for non-random decision components of rating process. The same applies with more general approaches to analyse within-subject heterogeneity such as random-effects and errors-in-variables models (Feng et al. 2018) which do not deal with non-random components of uncertainty. As a consequence, decision uncertainty underlying participant's rating process is not formally represented in these models.

In this contribution we propose a novel method for analysing continuous bounded ratings data that are characterized by non-random and systematic decision uncertainty. In particular, we propose a variable dispersion beta linear model which is generalized to cope with data contaminated by subjective uncertainty. We represent decision uncertainty in the framework of fuzzy data modeling, where crisp ratings data are equipped with non-random systematic uncertainty via normalized set functions (Couso and Dubois 2014; Kruse and Meyer 1987). In this setting, maximum likelihood estimation and inference are carried out through the Expectation–Maximization algorithm adapted for the case of fuzzy data (Denœux 2011; Su et al. 2015). It should be stressed that using beta linear models allows for flexibility in modeling and analysing continuous ratings data, while still retaining simplicity in estimates model's parameters (Algamal 2019; Canterle and Bayer 2019; Zeileis et al. 2010). Similarly, despite the fact that many formal theories have been proposed to deal with subjective uncertainty (e.g., soft sets, rough sets. See Lin and Cercone 2012; Liu 2010), fuzzy set theory offers a good compromise in terms of accuracy and computational costs and benefit from a long tradition of works in statistics (e.g., see Buckley 2006; Chukhrova and Johannssen 2018; Couso and Dubois 2014; Gebhardt et al. 1998; González-Rodríguez et al. 2006).

The remainder of the article is organized as follows. Section 2 briefly describes the basic characteristics of fuzzy data together with their interpretation in terms of decision uncertainty. Section 3 exposes the variable dispersion fuzzy beta model, parameters estimation, and model evaluation. Section 4 reports results of a short simulation study performed to evaluate the finite sample properties of the fuzzy beta model and the consequences of neglecting decision uncertainty on parameters estimation. Section 5 describes an application of the new approach to two case studies involving ratings data from risk-taking behaviors (application 1) and customer satisfaction (application 2). Finally, Sect. 6 concludes the article providing final remarks and suggestions for future extensions. All the materials like datasets and R-scripts used throughout the article are available to download at https://github.com/antcalcagni/fuzzyratingbeta.

2 Data representation

In this section we briefly review some concepts and terminology related to fuzzy numbers, fuzzy probability, and fuzzy ratings data.

2.1 Fuzzy numbers and fuzzy probability

A fuzzy subset $\tilde{A}$ of a universal set $\mathcal{A}$ is defined by means of its characteristic function $\xi_{\tilde{A}} \colon \mathcal{A} \to [0, 1]$. It can be easily described as a collection of crisp subsets called α-sets, i.e. $\tilde{A}_\alpha = \{x \in \mathcal{A} \colon \xi_{\tilde{A}}(x) > \alpha\}$ with $\alpha \in (0, 1]$. If the α-sets of $\tilde{A}$ are all convex sets then $\tilde{A}$ is a convex fuzzy set. The support of $\tilde{A}$ is $\tilde{A}_0 = \{x \in \mathcal{A} \colon \xi_{\tilde{A}}(x) > 0\}$ and the core is the set of all its maximal points, i.e. $\tilde{A}_c = \{x \in \mathcal{A} \colon \xi_{\tilde{A}}(x) = \max_{y \in \mathcal{A}} \xi_{\tilde{A}}(y)\}$. In the case $\max_{x \in \mathcal{A}} \xi_{\tilde{A}}(x) = 1$ then $\tilde{A}$ is a normal fuzzy set. If $\tilde{A}$ is a normal and convex subset of $\mathbb{R}$ then $\tilde{A}$ is a fuzzy number (Buckley 2006). Broadly speaking, fuzzy sets can be conceived as subsets of $\mathbb{R}$ where their Boolean characteristic function $\xi_A(x)$, $\forall x \in A \subset \mathbb{R}$, has been generalized to the real interval $[0, 1]$. The class of all normal fuzzy numbers is denoted by $\mathcal{F}(\mathbb{R})$. Fuzzy numbers can conveniently be represented using parametric models, through which $\xi_{\tilde{A}}$ is represented by means of few real parameters. Hence, we can define families of fuzzy numbers indexed by some scalars, such as m (mode) and s (spread/precision), which include a number of shapes like triangular, trapezoidal, gaussian, and exponential (Buckley 2006). Relevant classes of parametric fuzzy numbers are the so-called LR-fuzzy numbers (Dubois and Prade 1978) and their generalizations (Calcagnì et al. 2014; Dombi and Jónás 2018). In this setting, a set of operators and algebras have also been defined for fuzzy numbers, which extend traditional calculus to fuzzy numbers as well (Chwastyk and Kosiński 2013). A broader class that encompasses a wide range of fuzzy numbers is the so-called *beta fuzzy number* (Alimi 2003; Baklouti et al. 2018; Stein 1985):

$$
\begin{aligned}
\xi_{\tilde{A}}(x) &= \left(\frac{x - x_l}{m - x_l}\right)^a \left(\frac{x_u - x}{x_u - m}\right)^b \cdot \mathbb{1}_{(x_l, x_u)}(x) \\
m &= \frac{a x_u + b x_l}{a + b}
\end{aligned}
\tag{1}
$$

where $x_l, x_u, a, b \in \mathbb{R}$, with x_l and x_u being the lower and upper bounds of the set, and m the mode of the fuzzy set. This type of fuzzy numbers uses Beta functions to approximate many regular shapes such as triangular, trapezoidal or Gaussian. Likewise for LR-fuzzy numbers, beta fuzzy numbers can be defined in terms of mode $m \in \mathbb{R}$ and spread/precision $s \in \mathbb{R}^+$ parameters. In particular, let $x_l = 0$ and $x_u = 1$ without loss of generality. Then Eq. (1) can be re-arranged as follows:

$$
\begin{aligned}
\xi_{\tilde{A}}(x) &= \frac{1}{C} x^{a-1}(1 - x)^{b-1} \\
a &= 1 + ms \\
b &= 1 + s(1 - m)
\end{aligned}
\tag{2}
$$

with C being a constant ensuring $\xi_{\tilde{A}}$ is still a normal fuzzy set:

$$
C = \left(\frac{a - 1}{a + b - 2}\right)^{a-1} \cdot \left(1 - \frac{a - 1}{a + b - 2}\right)^{b-1}
$$

Interestingly, the re-parameterized Beta fuzzy number resembles the shape of Beta density distribution written using the PERT representation (Vose 2008). It should be noted that, like triangular or trapezoidal fuzzy numbers, also beta fuzzy numbers can be adopted to represent rating data variables. Indeed, the beta-like shape is particularly flexible to modeling some features of bounded rating data, such as asymmetry and tail flatness (e.g., see Migliorati et al. 2018).

When a probability space is defined over the reals, the probability of a fuzzy set $\mathbb{P}(\tilde{A})$ can also be defined. Over the years, there have been various attempts to define the probability of a fuzzy set in terms of expected value of its membership function (Zadeh 1968), conditional probability of prior information (Coletti and Scozzafava 2004; Singpurwalla and Booker 2004), imprecise probability (Augustin et al. 2014), fuzzy numbers (Hesamian and Shams 2017), and likelihood induced by random events (Cattaneo 2017). Following the findings of Denœux (2011), in this contribution we adopt Zadeh's definition of fuzzy probability (Zadeh 1968). In particular, let $(\mathbb{R}, \mathcal{B}, \mathbb{P})$ be a probability space. Then, $\mathbb{P}(\tilde{A})$ is defined as follows:

$$\mathbb{P}(\tilde{A}) = \int_{\mathbb{R}} \xi_{\tilde{A}}(x) \mathrm{d}\mathbb{P} \tag{3}$$

with $\xi_{\tilde{A}}$ being Borel measurable. In this context, two fuzzy sets $\tilde{A}$ and $\tilde{B}$ are said *independent* w.r.t. to $\mathbb{P}$, if $\mathbb{P}(\tilde{A}\tilde{B}) = \mathbb{P}(\tilde{A}) \cdot \mathbb{P}(\tilde{B})$, with the fuzzy product being defined as $\xi_{\tilde{A}\tilde{B}}(x) = \xi_{\tilde{A}}(x) \cdot \xi_{\tilde{B}}(x)$ (Zadeh 1968). The *conditional probability* of two independent fuzzy events is

$$\mathbb{P}(\tilde{A}|\tilde{B}) = \frac{\mathbb{P}(\tilde{A}, \tilde{B})}{\mathbb{P}(\tilde{B})} = \frac{\int \xi_{\tilde{A}}(x)\xi_{\tilde{B}}(x)\mathrm{d}\mathbb{P}(x)}{\int \xi_{\tilde{B}}(z)\mathrm{d}\mathbb{P}(z)} \tag{4}$$

with $\mathbb{P}(\tilde{B}) > 0$. Note that one can also obtain the conditional probability between a crisp set A and a fuzzy set $\tilde{B}$ as a special case of Eq. (4). If a discrete or continuous random variable Y is defined over $\mathcal{B}$, then fuzzy probability can be generalized accordingly. For instance, denoting with $f_Y(y;\boldsymbol{\theta})$ the probability density of Y, then $\mathbb{P}(\tilde{A}) = \int_{\mathcal{Y}} \xi_{\tilde{A}}(y) f_Y(y;\boldsymbol{\theta}) \, \mathrm{d}y$, with $\mathcal{Y}$ being the support of Y. Similarly, when a sample of n independent observations $\mathbf{y} = (y_1, \ldots, y_n)$ from $Y_1, \ldots, Y_n$ is available, the likelihood of the sample $\mathbf{y}$ can be generalized as follows:

$$\mathcal{L}(\boldsymbol{\theta};\mathbf{y}) = \prod_{i=1}^{n} \int_{\mathcal{Y}} \xi_{\tilde{y}_i}(y) f_{Y_i}(y;\boldsymbol{\theta}) \, \mathrm{d}y \tag{5}$$

where the definition $\xi_{\tilde{\mathbf{y}}} = \prod_{i=1}^{n} \xi_{\tilde{y}_i}(y)$ has been used for the joint fuzzy set (Gebhardt et al. 1998). Further details about fuzzy generalization of likelihood functions, fuzzy random variables, and fuzzy probability space can be found in Cattaneo (2017), Couso et al. (2014), Denœux (2011), Gebhardt et al. (1998), Gil et al. (2006).

We interpret fuzzy data in the context of random variables following the epistemic viewpoint on fuzzy set theory (Couso and Dubois 2014). In particular, for a fuzzy set $\tilde{A}$, $\xi_{\tilde{A}}(Y = y)$ is interpreted as the *possibility* that the crisp event $Y = y$ has to occur. Indeed, $\xi_{\tilde{A}}(Y = y) \in (0, 1)$ can be conceived as a graded plausibility about the occurrence of the event $Y = y$, with $\xi_{\tilde{A}}(Y = y) = 1$ indicating the fact that $Y = y$

is fully possible. By contrast, $\xi_{\widetilde{A}}(Y = y) = 0$ indicates that $Y = y$ is not possible at all. Hence, fuzzy sets can intuitively be viewed as *graded constraints* on crisp random variables. In this way, the randomness due to the data generation process and the fuzziness due to observer's state of knowledge can be analysed simultaneously by means of a common statistical representation. As a remark, note that in this setting $\xi_{\widetilde{A}}(Y = y)$ is thought as being the consequence of a two-step generation process, in which a realization y is drawn from Y first and then a fuzzy set $\xi_{\widetilde{A}}$ is used to encapsulate the uncertainty about $Y = y$ in terms of possibility distribution. Hence, only the first stage is a random experiment whereas the second stage is a non-random fuzzification of the outcomes being realized.

2.2 Fuzzy scaling and fuzzy data

Since the seminal work of Hesketh et al. (1988), fuzzy sets have been extensively used in the context of ratings data (for a review, see Calcagnì and Lombardi 2014; Lubiano et al. 2016; Calcagnì et al. 2021). Although several formats have been proposed for implementing fuzzy ratings tools (e.g., conversion scales, direct rating scales, implicit rating scales), all of them share the same idea that ratings data cannot be coerced into crisp numbers without a certain loss of information. Except for the case of dichotomous ratings, polytomous responses often show some degrees of imprecision and fuzziness, which is essentially due to participants' rating processes (Kahneman and Tversky 1982). In general, ratings data can be enriched by including information related to participants' response process and fuzzy conversion or fuzzy rating systems can be adopted to this purpose. In what follows, we will briefly review both the approaches to fuzzy rating.

Fuzzy conversion scales aim at turning standard crisp ratings into fuzzy numbers through the adoption of user-defined fuzzy systems (e.g., see Vonglao 2017) or statistically-oriented procedures (e.g., see Yu and Wu 2009). In general, a typical implementation of a user-defined fuzzy conversion scale, a fuzzy system relates a space of crisp responses (input) to a space of fuzzy numbers (output). Next implication rules mapping both input and output sources are finally established. For instance, in the simplest case of a 5-point scale, the crisp response $Y = 2$ activates the fuzzy sets of the output space via an IF–THEN implication rule. On the basis of the implemented fuzzy system (e.g., Mamdani, Sugeno), the output can result in the activation of one or two fuzzy sets and the final fuzzy response—which is given by the intersection of the activated fuzzy sets—will represent a more certain or less certain response, respectively.

Fuzzy rating scales adopt computerized interfaces through which the rater's response process is directly mapped to fuzzy numbers (e.g., see Hesketh et al. 1988; Calcagnì and Lombardi 2014; de Sáa et al. 2014). For instance, in the direct rating based implementation (de Sáa et al. 2014), raters are asked to respond by drawing fuzzy sets according to their perceived uncertainty. In this case, the rating procedure proceeds as follows. First, raters draw an interval on a pseudo-continuous graphical scale, which represents the set of admissible responses compatible with the assessment of the item being rated. Then, a degree of confidence is expressed by drawing another interval

around the previous interval response. Finally, both the intervals are interpolated to form the final fuzzy responses (e.g., trapezoidal or triangular). In a similar way, fuzzy indirect scales adopt a computerized graphical interface to get responses from raters. However, in this case, they are not asked to directly draw their responses; rather, fuzzy sets are build from a set of implicit measures associated to the final crisp response (e.g., response time). For instance, the DYFRAT scale (Calcagnì and Lombardi 2014) uses a set of biometric measures associated with the cognitive response process (i.e., response time, computer-mouse trajectories) to derive the rater's fuzzy response. In particular, raters are presented with a pseudo-circular scale with M levels and—similarly to the case of crisp Likert-type scales—they are asked to choose which of these levels is the most appropriate to represent their response for a given item being rated. Meanwhile, the system records the streaming coordinates of the computer mouse cursor (at a fixed sampling rate) and the elapsed time since the beginning of the response trajectory. Finally, both information are used to define a fuzzy response according to the following rationale: (1) computer-mouse trajectories are projected onto a linear scale and the histogram of linearized radians is used to derive the fuzzy set (e.g., triangular, beta) in terms of its support and core, (2) response times are used to intensify or reduce the fuzziness of the sets by means of linguistic quantifiers (e.g., the longer the response time, the higher the fuzziness of the set).

In the context of fuzzy scaling, fuzzy numbers are used as a formal representation for rating responses involving individual-based judgments, attitudes, and opinions. Although fuzzy conversion and rating scales differ in the way they derive fuzzy responses, both aim at providing a model for the fuzziness or imprecision which is present in the rating process $Y \sim f_Y(y; \boldsymbol{\theta})$. As for the more general case of LR fuzzy numbers, the parameters of a beta fuzzy number can be linked to the rating process as follows. First, consider a continuous rating scale bounded on a subset (y_l, y_u) of reals. Then, m represents the most plausible final rating choice $\xi_{\widetilde{Y}}(Y = y) = 1$, s is the precision of m such that smaller values indicate larger levels of hesitation in the rating choice, and $\xi_{\widetilde{Y}}$ conveys the overall decision uncertainty in terms of fuzziness (the larger the fuzziness, the highest the decision uncertainty). Note that, ideally, if there was no subjective uncertainty, then the fuzziness would tend to zero and true rating realizations would be precisely observed (i.e., $Y = y = m$). In this case, there would not need to represent ratings as fuzzy data.

3 Variable dispersion beta model for fuzzy ratings

In this section we illustrate our proposal to analyse continuous bounded ratings data in situations with decision uncertainty. Hereafter, ratings data will be considered scaled into the real subset $(0, 1)$ without loss of generality.

3.1 Model

Let $\mathbf{y} = (y_1, \ldots, y_i, \ldots, y_n) \in (0, 1)^n$ be a sample of n observations from Beta distributed independent random variables $Y_1, \ldots, Y_i, \ldots, Y_n$ with density:

$$f_{\mathbf{Y}}(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\phi}) = \prod_{i=1}^{n} \frac{\Gamma(\phi_i)}{\Gamma(\phi_i\mu_i)\Gamma(\phi_i - \mu_i\phi_i)} \, y_i^{(\mu_i\phi_i-1)}(1-y_i)^{(\phi_i-\mu_i\phi_i-1)} \qquad (6)$$

with $\boldsymbol{\mu} \in (0,1)^n$ being the $n \times 1$ vector of location parameters and $\boldsymbol{\phi} \in (0,\infty)^n$ the $n \times 1$ vector of precision parameters (Ferrari and Cribari-Neto 2004). The sequence $Y_1, \dots, Y_i, \dots, Y_n$ models the ratings for each of the n participants, with the convention that $\{0,1\}$ represent the lower and upper bounds of the rating domain, respectively. In order to account for heterogeneity and non-constant variance in rating responses, location and dispersion parameters can be non-linearly re-written using monotonic and twice differentiable link functions, mapping the support $(0,1)$ into $\mathbb{R}$, as follows:

$$g_1(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} \quad \text{and} \quad g_2(\boldsymbol{\phi}) = \mathbf{Z}\boldsymbol{\gamma} \qquad (7)$$

where $\mathbf{X}$ and $\mathbf{Z}$ are $n \times J$ and $n \times H$ matrices of known continuous or categorical covariates, with $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ being vectors of appropriate order containing unknown parameters. The functions $g_1(\cdot)$ and $g_2(\cdot)$ can be chosen among a variety of link functions (e.g., logit, probit, log. See McCulloch 2000). Two typical choices are the logit and logarithm functions, which yield to:

$$\boldsymbol{\mu} = \left(1 + \exp(\mathbf{X}\boldsymbol{\beta})\right)^{-1} \quad \text{and} \quad \boldsymbol{\phi} = \exp(\mathbf{Z}\boldsymbol{\gamma}) \qquad (8)$$

Under Eq. (8), the log-likelihood function for the variable dispersion beta model is:

$$\begin{aligned}
l(\boldsymbol{\theta};\mathbf{y}) = \sum_{i=1}^{n} \; & \ln\Gamma\left(\exp(\mathbf{z}_i\boldsymbol{\gamma})\right) - \ln\Gamma\left(\frac{\exp(\mathbf{z}_i\boldsymbol{\gamma})}{1+\exp(-\mathbf{x}_i\boldsymbol{\beta})}\right) \\
& - \ln\Gamma\left(\exp(\mathbf{z}_i\boldsymbol{\gamma}) - \frac{\exp(\mathbf{z}_i\boldsymbol{\gamma})}{1+\exp(-\mathbf{x}_i\boldsymbol{\beta})}\right) \\
& + \ln y_i\left(\frac{\exp(\mathbf{z}_i\boldsymbol{\gamma})}{1+\exp(-\mathbf{x}_i\boldsymbol{\beta})} - 1\right) \\
& + \ln(1-y_i)\left(\exp(\mathbf{z}_i\boldsymbol{\gamma}) - \frac{\exp(\mathbf{z}_i\boldsymbol{\gamma})}{1+\exp(-\mathbf{x}_i\boldsymbol{\beta})} - 1\right)
\end{aligned} \qquad (9)$$

with $\Gamma(\cdot)$ being the Euler gamma function and $\mathbf{t} = \left(\ln\mathbf{y}, \ln(\mathbf{1}-\mathbf{y})\right)$ the sufficient statistics for the inference on $\boldsymbol{\theta} = (\boldsymbol{\beta},\boldsymbol{\gamma})$.

In light of the data representation adopted in this work, decision uncertainty is treated as a systematic and non-random component which occurs after the sampling process $\mathbf{y} \sim f_{\mathbf{Y}}(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\phi})$ has been realized. This leads to a situation where the sample $\mathbf{y}$ cannot be precisely observed and a collection of fuzzy data $\tilde{\mathbf{y}}$ is instead available. When fuzzy data are represented as beta fuzzy numbers, then

$$\tilde{\mathbf{y}} = (\mathbf{m},\mathbf{s}) = \left((m_1,s_1), \dots, (m_i,s_i), \dots, (m_n,s_n)\right)$$

with $\mathbf{m}$ and $\mathbf{s}$ being $n \times 1$ vectors of modes and precisions/spreads for the fuzzy observations. Turning Eqs. (3) into (6), the joint density of $\tilde{\mathbf{y}}$ can be written as:

$$f_{\widetilde{\mathbf{Y}}}(\mathbf{y};\boldsymbol{\mu},\boldsymbol{\phi}) = \prod_{i=1}^{n} \int_{\mathcal{Y}} \xi_{\widetilde{y}_i}(y) \, d\mathbb{P}_{\theta}\big(Y_i = y\big)$$

$$= \prod_{i=1}^{n} \int_{\mathcal{Y}} \xi_{\widetilde{y}_i}(y) \, \frac{\gamma(\phi_i)y^{(\mu_i\phi_i-1)}(1-y)^{(\phi_i-\mu_i\phi_i-1)}}{\gamma(\phi_i\mu_i)\gamma(\phi_i-\mu_i\phi_i)} \, dy \tag{10}$$

where the joint fuzzy set $\xi_{\widetilde{\mathbf{y}}} = \prod_{i=1}^{n} \xi_{\widetilde{y}_i}(y)$ has been factorized in terms of product. The log-likelihood of the model under fuzzy observations can analogously be obtained using Eq. (10). Note that, in this representation the vectors $\mathbf{m}$ and $\mathbf{s}$ of the fuzzy sets enter the model as observed quantities whereas the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ still remain non-fuzzy quantities.

3.2 Parameter estimation

To provide estimates for $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$ in the context of fuzzy ratings data, one can maximize the log-likelihood $l(\boldsymbol{\theta};\tilde{\mathbf{y}})$ function, which is obtained by Eq. (10). This would require an iterative procedure, alternating between the numerical computation of the integral and the maximization of the function. However, to avoid the problem of approximating integrals in Eq. (10) and have a way to compute standard errors consistently, we will use a variant of the Expectation–Maximization algorithm generalized for the case of fuzzy data (Denœux 2011). As for the standard EM algorithm, the fuzzy-EM version at the kth iteration alternates between the E-step, which involves the computation of the expected complete-data log-likelihood using $\boldsymbol{\theta}^{(k-1)}$, and the M-step, which instead maximizes the expected complete-data log-likelihood w.r.t. to $\boldsymbol{\theta}^{(k)}$. These steps generate a non-decreasing sequence of lower bounds for the maximization of the observed-data log-likelihood $l(\boldsymbol{\theta};\tilde{\mathbf{y}})$ (for formal details, see Denœux 2011). In the fuzzy-EM variant, the complete-data log-likelihood is that obtained if $\mathbf{y}$ was precisely observed (see Eq. 9). Therefore, given the $(k-1)$th estimates $\boldsymbol{\theta}' = \boldsymbol{\theta}^{(k-1)}$ the E-step for the kth iteration of the algorithm consists in the computation of the following quantity:

$$\mathcal{Q}\big(\boldsymbol{\theta},\boldsymbol{\theta}'\big) = \mathbb{E}_{\theta'}\big[l(\boldsymbol{\theta};\mathbf{y})\big|\tilde{\mathbf{y}}\big]$$

$$= C + \sum_{i=1}^{n} \mathbb{E}_{\theta'}\big[\ln Y_i|\tilde{y}_i\big]\left(\frac{\exp\big(\mathbf{z}_i\boldsymbol{\gamma}'\big)}{1+\exp\big(-\mathbf{x}_i\boldsymbol{\beta}'\big)} - 1\right)$$

$$+ \sum_{i=1}^{n} \mathbb{E}_{\theta'}\big[\ln(1-Y_i)|\tilde{y}_i\big]\left(\exp\big(\mathbf{z}_i\boldsymbol{\gamma}'\big) - \frac{\exp\big(\mathbf{z}_i\boldsymbol{\gamma}'\big)}{1+\exp\big(-\mathbf{x}_i\boldsymbol{\beta}'\big)} - 1\right) \tag{11}$$

where C contains all the terms of Eq. (9) that do not involve the random quantities to be filtered. To compute the conditional expectations of the E-step, note that $Y_i|\tilde{y}_i$ is a random variable conditioned on fuzzy events and its density can be obtained using Eq. (4) simplified for the case where A is a crisp event:

$$f_{Y_i|\tilde{y}_i}(y;\mu_i,\phi_i) = \frac{\xi_{\tilde{y}_i}(y)f_{Y_i}(y;\mu_i,\phi_i)}{\int_{\mathcal{Y}}\xi_{\tilde{y}_i}(x)f_{Y_i}(x;\mu_i,\phi_i)\,dx}$$
$$\propto y_i^{(\eta_i-1)}(1-y_i)^{(\nu_i-1)}$$

(12)

Under the case of beta fuzzy numbers and up to a normalization constant, the conditional density $f_{Y_i|\tilde{y}_i}(y;\mu_i,\phi_i)$ corresponds to a beta density with parameters given as a function of fuzzy data and crisp parameters of the complete-data density:

$$\eta_i = \mu_i\phi_i + s_im_i + 1$$
$$\nu_i = \phi_i(1-\mu_i) + s_i(1-m_i) + 1$$

(13)

Then, the first expectation $\mathbb{E}_{\theta'}\left[\ln Y_i|\tilde{y}_i\right]$ can be approximated via Taylor expansion around $\mathbb{E}_{\theta'}\left[Y_i|\tilde{y}_i\right]$ as follows:

$$\mathbb{E}_{\theta'}\left[\ln Y_i|\tilde{y}_i\right] \stackrel{\sim}{=} \ln \mathbb{E}_{\theta'}\left[Y_i|\tilde{y}_i\right] - \frac{1}{2}\frac{\mathbb{V}\mathrm{ar}_{\theta'}\left[Y_i|\tilde{y}_i\right]}{\left(\mathbb{E}_{\theta'}\left[Y_i|\tilde{y}_i\right]\right)^2}$$
$$\stackrel{\sim}{=} \ln\left(\frac{\eta_i}{\eta_i+\nu_i}\right) - \frac{\nu_i}{2\eta_i(1+\eta_i+\nu_i)}$$

(14)

Similarly, the second expectation $\mathbb{E}_{\theta'}\left[\ln(1-Y_i)|\tilde{y}_i\right]$ is obtained by symmetry of the beta function:

$$\mathbb{E}_{\theta'}\left[\ln(1-Y_i)|\tilde{y}_i\right] \stackrel{\sim}{=} \ln \mathbb{E}_{\theta'}\left[(1-Y_i)|\tilde{y}_i\right] - \frac{1}{2}\frac{\mathbb{V}\mathrm{ar}_{\theta'}\left[(1-Y_i)|\tilde{y}_i\right]}{\left(\mathbb{E}_{\theta'}\left[(1-Y_i)|\tilde{y}_i\right]\right)^2}$$
$$\stackrel{\sim}{=} \ln\left(\frac{\nu_i}{\eta_i+\nu_i}\right) - \frac{\eta_i}{2\nu_i(1+\eta_i+\nu_i)}$$

(15)

Once expected values are computed, the M-step of the algorithm involves the maximization of $\mathcal{Q}(\theta,\theta')$ with respect to the elements of θ and can be performed by plugging-in Eqs. (14)–(15) into the log-likelihood Eq. (11). The simultaneous score equations for M-step are as follows:

$$\mathcal{U}_\beta = \frac{\partial\mathcal{Q}(\theta,\theta')}{\partial\mu}\frac{\partial\mu}{\partial\beta} = \mathbf{0}_J$$
$$= \mathbf{X}^T\left[\phi\left(\mathbf{y}_{(1)}^* - \mathbf{y}_{(2)}^* - \Psi(\phi\mu) + \Psi(\phi(1-\mu))\right)\right]\mu^2 = \mathbf{0}_J$$

(16)

$$\mathcal{U}_\gamma = \frac{\partial\mathcal{Q}(\theta,\theta')}{\partial\phi}\frac{\partial\phi}{\partial\gamma} = \mathbf{0}_H$$
$$= \mathbf{Z}^T\left[\phi\Psi(\phi) - \Psi(\phi-\phi\mu)\phi + \Psi(\phi-\phi\mu)\phi\mu + \mathbf{y}_{(1)}^*\mu\phi\right.$$
$$\left. - \phi\Psi(\mu\phi)\mu - \mathbf{y}_{(2)}^*\phi(\mu-1)\right] = \mathbf{0}_H$$

(17)

where μ and ϕ are defined as in Eq. (8) whereas the terms

$$\mathbf{y}^*_{(1)} = \left(\mathbb{E}_{\theta'}\left[\ln Y_1 | \tilde{y}_1\right], \ldots, \mathbb{E}_{\theta'}\left[\ln Y_n | \tilde{y}_n\right] \right)$$

$$\mathbf{y}^*_{(2)} = \left(\mathbb{E}_{\theta'}\left[\ln(1 - Y_1) | \tilde{y}_1\right], \ldots, \mathbb{E}_{\theta'}\left[\ln(1 - Y_n) | \tilde{y}_n\right] \right)$$

denote the filtered data which are computed using Eqs. (14)–(15). Finally, $\theta^{(k)}$ is obtained by computing the roots of Eqs. (16)–(17) numerically, for instance using Gauss-Newton method.

3.3 Standard errors, diagnostics, and inference

Standard errors for $\hat{\beta}$ and $\hat{\gamma}$ can be computed using the observed information matrix $\mathcal{I}_{\hat{\theta}}$ from the Hessian of the maximum likelihood estimates $\hat{\theta}$ obtained solving Eqs. (16)–(17) for β and γ. In the context of EM algorithm, $\mathcal{I}_{\hat{\theta}}$ can be approximated using the empirical observed information matrix (Meilijson 1989), as follows:

$$\mathcal{I}_{\hat{\theta}} \; \tilde{=} \; \mathcal{I}^e_{\hat{\theta}} = \sum_{i=1}^{n} \mathcal{U}^{(i)}_{\hat{\theta}} \left(\mathcal{U}^{(i)}_{\hat{\theta}} \right)^T \tag{18}$$

where $\mathcal{U}^{(i)}_{\hat{\theta}} = [\mathcal{U}^{(i)}_{\hat{\beta}}; \mathcal{U}^{(i)}_{\hat{\gamma}}]$ is the score vector for the ith observation calculated at $\hat{\beta}$ and $\hat{\gamma}$. The standard errors are then calculated as usual:

$$\sigma_{\hat{\theta}} = \left(\mathrm{diag}\, \{(\mathcal{I}^e_{\hat{\theta}})^{-1}\} \right)^{\frac{1}{2}} \tag{19}$$

Note that this approximation avoids the computation of Hessian of the complete-data log-likelihood as it solely uses the score equations where the unobserved vector $\mathbf{y}$ is replaced by $\mathbf{y}^*_{(1)}$ and $\mathbf{y}^*_{(2)}$. Alternatively, standard errors can also be obtained via non-parametric bootstrap (McLachlan and Peel 2004). An important quantity commonly used to assess the quality of the estimated model is that involving standardized residuals which, for the case of fuzzy data, can be generalized as follows:

$$r_i = \frac{(m_i - \hat{\mu}_i)(1 - \xi_{\tilde{y}_i}(\hat{\mu}_i))}{\hat{\mu}_i(1 - \hat{\mu}_i)/(1 + \hat{\phi}_i)} \tag{20}$$

where $\hat{\mu}_i = (1 + \exp(\mathbf{x}_i\hat{\beta}))^{-1}$ and $\hat{\phi} = \exp(\mathbf{z}_i\hat{\gamma})$ whereas $\xi_{\tilde{y}_i}(\hat{\mu}_i)$ is the fuzzy membership computed for the predicted quantity $\hat{\mu}_i$. In general, diagnostics for the model can be performed by plotting, for instance, $r_1, \ldots, r_n$ against the indices of the observations in order to check for particular trends or patterns in the predicted data. Similarly to generalized and beta linear models, also the overall fit of the estimated fuzzy beta model with respect to the observed fuzzy data can be assessed by means of pseudo-R^2 indices, which generalize the standard residuals-based R^2 indices (Ferrari and Cribari-Neto 2004; Veall and Zimmermann 1994). In this context, we resorted in applying a likelihood-ratio based pseudo-R^2 index (Aldrich and Nelson 1984; Veall and Zimmermann 1994) which, like for the McFadden's pseudo-R^2 (McFadden et al. 1973), contrasts the likelihood value $l(\hat{\theta};\mathbf{y})_{\mathcal{M}_1}$ of the fuzzy beta linear model against the likelihood value of the null model $l(\hat{\theta};\mathbf{y})_{\mathcal{M}_0}$ (i.e., a fuzzy beta linear model with no predictors for μ and ϕ). The index is as follows:

$$\text{pseudo-}R^2 = -\frac{\omega(1-\lambda)}{(\omega+n)\lambda}$$

where $\omega = 2\left(l(\hat{\boldsymbol{\theta}};\mathbf{y})_{\mathcal{M}_1} - l(\hat{\boldsymbol{\theta}};\mathbf{y})_{\mathcal{M}_0}\right)$ and $\lambda = \frac{1}{n}l(\hat{\boldsymbol{\theta}};\mathbf{y})_{\mathcal{M}_0}$. The likelihood-ratio based pseudo-R^2 is normalized in [0, 1] and approximates the relationship between the likelihood-ratio statistic and the R^2 index in the linear regression setting (Veall and Zimmermann 1994).

Finally, likewise for the non-fuzzy case, inference on $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ can be performed using maximum-likelihood (ML) theory (Denœux 2011; McLachlan and Peel 2004) and, consequently, hypothesis testing on model's parameters can be performed using fuzzy version of likelihood ratio test (e.g., see Berkachy and Donzé 2019; Najafi et al. 2010). In this context, as for the maximum-likelihood theory under the EM procedure, inferential results are based on the asymptotic properties of ML-based estimators. For further details, we refer the reader to Denœux (2011), McLachlan and Peel (2004), and Berkachy and Donzé (2019).

4 Simulation study

The aim of this study is twofold. First, we will evaluate the performances of EM estimators for location and precision parameters for the fuzzy beta linear model. Second, we will assess whether standard methods, such as fixed/random-effects beta linear models, perform as good as the proposed method if applied on defuzzified data. Although the EM algorithm for fuzzy data has been validated elsewhere (e.g., see De Bruin et al. 2011; Su et al. 2015), in the present study we have preferred to evaluate the performances of the fuzzy-EM procedure to further provide converging results. The whole simulation procedure has been performed on a (remote) HPC machine based on 16 cpu Intel Xeon CPU E5-2630L v3 1.80 GHz, 16×4 GB Ram whereas computations and analyses have been done in the R framework for statistical analyses.

Design The design involved three factors, namely (1) $n \in \{50, 100, 250, 500\}$, (2) $J \in \{2, 4\}$, and (3) $H \in \{1, 3\}$, which were varied in a complete factorial design, producing $K = 4 \times 2 \times 2 = 16$ possible combinations. For each combination, $B = 5000$ samples were generated yielding to $5000 \times 16 = 80,000$ new data as well as an equivalent number of parameters. The true parameters of the model were fixed as follows:

$$\boldsymbol{\beta}^0 = \{(-0.5, -0.81), (-0.5, -0.81, 0.7, 1.15)\}$$
$$\boldsymbol{\gamma}^0 = \{(4.8), (4.8, -1.5, 1.03)\}$$

Procedure Let n_l, j_k, h_m be distinct levels of factors n, J, H. Then, fuzzy data were generating according to the following procedure which mimics the hierarchical process underlying rating under decision uncertainty:

(a) $\mathbf{X}_{n_l \times j_k} = [\mathbf{1}_{n_l}, \mathbf{X}_{n_l \times j_{k-1}}]$ and $\mathbf{Z}_{n_l \times h_m} = [\mathbf{1}_{n_l}, \mathbf{Z}_{n_l \times h_{m-1}}]$, where $\mathbf{X}_{n_l \times j_{k-1}}$ and $\mathbf{Z}_{n_l \times h_{m-1}}$ were drawn from $\mathcal{U}nif(1, 5)$

(b) location and precision terms were computed as follows:

$$\boldsymbol{\mu}_{n_l \times 1} = \text{logit}^{-1}\left\{ 1 + \exp(\mathbf{X}_{n_l \times j_k} \boldsymbol{\beta}^0_{j_k \times 1}) \right\}$$

$$\boldsymbol{\phi}_{n_l \times 1} = \exp(\mathbf{Z}_{n_l \times h_m} \boldsymbol{\gamma}^0_{h_m \times 1})$$

(c) crisp data underlying fuzzy observations were generated according to the variable dispersion beta linear model $y_i \sim \mathcal{B}eta(\mu_i \phi_i, \phi_i - \phi_i \mu_i)$, $i = 1, \dots, n_l$

(d) fuzzy data were generated by making $\mathbf{y}_{n_l}$ imprecise via a two-step data-generation process (Quost and Denoeux 2016; Su et al. 2014). First, spread components were generated as $\mathbf{s}_{n_l} \sim \mathcal{G}amma(1.025, 0.001)$. Second, modes were generated by $m_i \sim \mathcal{B}eta(y_i s_i, s_i - s_i y_i)$, $i = 1, \dots, n_l$

(e) parameters $\boldsymbol{\beta}_{j_k}$ and $\boldsymbol{\gamma}_{h_m}$ were estimated using four methods:

- fEM: expectation–maximization estimators for the fuzzy case.
- dML: maximum-likelihood estimators on two type of defuzzified data computed using the *centroid method* $y_i^* = \int_0^1 y\, \xi_{\widetilde{y}_i}(y)\, \mathrm{d}y$ and the *first-maximum method* $y_i^* = \sup_{y \in (0,1)} \left\{ \xi_{\widetilde{y}_i}(y) \right\}$ for $i = 1, \dots, n_l$. The ML procedure implemented in the R library `betareg` has been used in this case (Zeileis et al. 2010).
- dREML: restricted maximum-likelihood estimators on defuzzified data obtained by treating fuzzy sets $\xi_{\widetilde{y}_1}, \dots, \xi_{\widetilde{y}_n}$ as random effects. In particular, for each observation $i = 1, \dots, n_l$, extremes of α-sets $y_i^* = [\inf(\mathbf{y}_i^c), \sup(\mathbf{y}_i^c)]$ were used, with $\mathbf{y}_i^c = \{ y \in (0, 1) : \xi_{\widetilde{y}_i}(y) \geq \alpha \}$ being the set obtained by cutting the ith fuzzy data at $\alpha = [0.01, 0.35, 0.7]$. In this case, estimates have been performed using the R library `glmmTMB` for random-effects beta linear models (Brooks et al. 2017).

Outcome measures For each condition of the simulation design, sample results were evaluated using *bias* of estimates and *root mean square error*.

Results Tables 1 and 2 show the results of the simulation study with regards to averaged bias and root mean square error. For the sake of clarity, results for the cases $J = 2$ and $J = 4$ were reported separately. To better interpret the results, it should be noted that conditions with $H = 1$ represent simplest cases where the precision term is held constant (the linear term for $\boldsymbol{\phi}$ is a simple intercept model). By contrast, conditions with $H = 3$ represent those situations showing some levels of heterogeneity in the response variable (in this case the linear term for $\boldsymbol{\phi}$ contains two slopes). We first consider the conditions with $J = 2$. With respect to the parameters $\hat{\boldsymbol{\beta}}$, all the methods showed negligible bias in estimating the location terms of the beta linear model both in the cases with $H = 1$ and $H = 3$. However, unlike dML and dREML, the fEM solution achieved lowest RMSE. Considering the parameters $\hat{\boldsymbol{\gamma}}$, dML and dREML algorithms showed worse performances when compared to fEM both in the case of low ($H = 1$) and high ($H = 3$) model complexity. In particular, they showed larger bias in estimating the precision terms of the beta linear model, with bias being higher with increasing model complexity ($H = 3$). A similar pattern was

Table 1 Monte Carlo study: average bias and average root mean square errors for the arrays of parameters $\hat{\beta}$ and $\hat{\gamma}$ (case $J = 2$)

n, H	fEM		dML (mean)		dML (mode)		dREML	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
$\hat{\beta}$								
$n = 50$								
$H = 1$	− 0.001	0.198	− 0.017	0.251	0.002	0.276	− 0.032	0.459
$H = 3$	0.000	0.355	− 0.007	0.387	− 0.006	0.401	− 0.063	0.626
$n = 100$								
$H = 1$	− 0.001	0.120	− 0.010	0.181	0.002	0.177	− 0.043	0.412
$H = 3$	− 0.001	0.186	− 0.007	0.220	− 0.003	0.200	− 0.051	0.433
$n = 250$								
$H = 1$	0.000	0.073	− 0.006	0.142	0.001	0.118	− 0.042	0.341
$H = 3$	− 0.002	0.126	− 0.008	0.173	− 0.003	0.140	− 0.054	0.331
$n = 500$								
$H = 1$	− 0.001	0.050	− 0.006	0.132	0.000	0.086	− 0.035	0.289
$H = 3$	0.000	0.084	− 0.003	0.134	− 0.002	0.095	− 0.042	0.255
$\hat{\gamma}$								
$H = 1$	0.069	0.039	− 0.351	0.104	− 0.531	0.138	− 1.415	0.320
$H = 3$	− 0.025	0.425	− 0.177	0.540	− 0.241	0.714	− 0.059	1.077
$n = 100$								
$H = 1$	0.048	0.033	− 0.408	0.102	− 0.608	0.145	− 1.204	0.270
$H = 3$	− 0.013	0.613	− 0.152	0.766	− 0.141	0.923	− 0.077	1.292
$n = 250$								
$H = 1$	0.039	0.025	− 0.435	0.098	− 0.665	0.150	− 1.113	0.241
$H = 3$	− 0.031	0.655	− 0.180	0.835	− 0.190	1.177	− 0.148	1.723
$n = 500$								
$H = 1$	0.028	0.018	− 0.447	0.097	− 0.681	0.149	− 1.138	0.241
$H = 3$	− 0.020	0.743	− 0.147	1.369	− 0.163	1.215	− 0.195	1.628

also observed for RMSE. In particular, when compared to fEM, dML and dREML showed larger values, with severity increasing over model complexity. Interestingly, although the condition with $H = 3$ represents a more complex situation with respect to parameters estimation, fEM outperformed the other methods. Results for the conditions with $J = 4$ largely resemble those obtained with $H = 1$. Also in this case, the fEM algorithm showed better performances over dML and dREML. Finally, for each method we also computed overall indices of over/under-estimation r_β and r_γ, as the ratio between the number of positive and negative bias, and the overall percentage of over-estimation p_β and p_γ. Table 3 reports the results along with the overall RMSE for both the arrays of parameters. In general, fEM and dML (mode) showed negligible overestimation for β^0 whereas dML (mean) and dREML tended to overestimate the true arrays of parameters. On the contrary, when estimating γ^0, fEM tended to overestimate the true population parameters whereas dML and dREML

Table 2 Monte Carlo study: average bias and average root mean square errors for the arrays of parameters $\hat{\beta}$ and $\hat{\gamma}$ (case $J = 4$)

n, H	fEM		dML (mean)		dML (mode)		dREML	
	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
$\hat{\beta}$								
$n = 50$								
$H = 1$	0.001	0.401	0.000	0.439	0.001	0.523	0.000	0.521
$H = 3$	0.001	0.631	0.010	0.642	0.006	0.695	0.019	0.892
$n = 100$								
$H = 1$	0.000	0.341	0.010	0.367	− 0.001	0.469	0.038	0.518
$H = 3$	0.000	0.433	0.000	0.458	0.002	0.512	0.002	0.695
$n = 250$								
$H = 1$	0.000	0.348	− 0.001	0.455	0.000	0.539	0.001	0.760
$H = 3$	0.003	0.637	0.009	0.644	0.011	0.903	0.034	1.001
$n = 500$								
$H = 1$	0.006	0.562	− 0.003	0.545	0.001	0.669	0.000	0.974
$H = 3$	0.001	0.461	0.006	0.557	0.004	0.612	0.011	0.756
$\hat{\gamma}$								
$n = 50$								
$H = 1$	0.090	0.040	− 0.308	0.094	− 0.480	0.139	− 1.279	0.294
$H = 3$	0.044	0.601	− 0.066	0.794	− 0.035	0.908	0.020	1.264
$n = 100$								
$H = 1$	0.075	0.034	− 0.374	0.096	− 0.582	0.142	− 1.198	0.268
$H = 3$	0.051	0.732	0.038	1.260	− 0.009	1.350	− 0.143	1.593
$n = 250$								
$H = 1$	0.041	0.024	− 0.421	0.094	− 0.643	0.149	− 1.140	0.246
$H = 3$	− 0.035	0.753	− 0.187	1.110	− 0.200	1.198	− 0.163	1.749
$n = 500$								
$H = 1$	0.030	0.018	− 0.435	0.094	− 0.684	0.151	− 1.158	0.246
$H = 3$	− 0.015	1.016	− 0.132	1.073	− 0.158	1.533	− 0.211	2.556

tended to underestimate. On the whole, fEM showed less variable and less biased estimates of β^0 and γ^0 than the other procedures adopted for defuzzified data.

5 Applications

In this section we will illustrate the application of the fuzzy beta model to two empirical studies involving fuzzy ratings data collected using two types of fuzzy rating scales. In particular, the first application concerns the analysis of risk-taking behavioral data collected by means of an indirect fuzzy rating approach (Calcagnì and Lombardi 2014). By contrast, the second application is about the analysis of customer satisfaction data collected by means of a direct fuzzy rating approach

Table 3 Monte Carlo study: overall ratios r between over and under-estimation, percentage p of over-estimation, and overall root mean square error

	fEM			dML (mean)			dML (mode)			dREML		
	r	p	$rmse_{ov}$	r	p	$rmse_{ov}$	r	p	$rmse_{ov}$	r	p	$rmse_{ov}$
$\hat{\beta}$	1.002	50.1	0.313	1.244	55.3	0.358	1.008	50.2	0.401	1.161	53.4	0.579
$\hat{\gamma}$	1.388	44.0	0.361	0.362	77.8	0.533	0.351	79.1	0.636	0.390	78.4	0.938

Note that all the indices were computed over $B = 5000$ replications

(de Sáa et al. 2014). Note that the applications are chosen to review two of the most common methods for fuzzy scaling, namely fuzzy indirect (case study 1) and direct rating scales (case study 2).

5.1 Case study 1: Reckless-driving and risk-taking behavior in young drivers

Reckless-driving among young people is one of the major cause of mortality and injuries worldwide (Toroyan 2015). It is a very complex phenomenon involving a number of human and non-human factors, such as personality, cognitive styles, social context (e.g., family, peer group), infrastructures (e.g., roads, light), and cultures (e.g., see Bıçaksız and Özkan 2016; McNally and Bradley 2014; Scott-Parker et al. 2015; Taubman-Ben-Ari 2010). Several studies have recognized the role of subjective factors like sensation-seeking, normlessness, anxiety, aggressiveness, driving attitudes in determining risky behaviors (Bıçaksız and Özkan 2016). Researchers have also assessed the contribution of parenting styles and peer relations to young drivers' intention to take risks (Taubman-Ben-Ari 2010, 2014). Because of its characteristics, assessing reckless-driving behaviors is a typical situation where self-reported measures can show some levels of decision uncertainty, which cannot appropriately be analysed using final crisp responses only. In this application, we consider a set of models where reckless-driving behaviors (`rdb`) were linearly predicted by the use of substances (`drugs`), driving anger (`anger`), and family climate (`fcrs`). In particular, we hypothesized that both the use of substances and driving anger would linearly increase the self-reported reckless-driving behaviors whereas family climate would instead acts by decreasing the amount of risky behaviors. Moreover, we also assessed whether the dispersion of fuzzy ratings data varied as a function of participants' characteristics such as gender (`sex`) and frequency of driving (`driving_frequency`).

Data and measures A questionnaire survey was carried out on $n = 69$ young drivers in Trentino region (north-est of Italy). Of these, 31% were women with mean age of 18.27 years (SD $= 0.56$). All participants were young drivers with an average of driving experience of 12 months since receipt of their driver's license. About 73% of them drove frequently during the week, 26% drove once a week. The survey consisted of 24 items from three self-reported questionnaires: (1) the Reckless Driving Behavior Scale (RDB) (McNally and Bradley 2014) used to assess those behaviors that increase the probability of a vehicle crash due to driving under the influence of

substances (drugs), extreme motorsport behaviors (extreme), and speeding/steering behaviors (positioning); (2) a short version of the Driving Anger Scale (DAS) (Deffenbacher et al. 1994), adopted to assess driving angers provoked by someone else's behaviors like slow driving and discourtesy; (3) a simplified version of the Family Climate Road Safety (FCRS) questionnaire (Taubman-Ben-Ari and Katz-Ben-Ami 2013), adopted to evaluate the role of parents in teens' safe driving, especially with regards to communication, monitoring, and parents' messages. Questionnaires were administered using DYFRAT (Calcagnì and Lombardi 2014), a computerized fuzzy rating scale which adopts the mouse-tracking methodology (Freeman and Ambady 2010) as a tool to implicitly quantify rating processes. For each item of the survey, participants were asked to respond using a pseudo-circular rating scale with six-points (RDB), five-points (DAS), and four-points (FCRS) anchors, respectively. Participants gave their responses by mouse-clicking the chosen level of the scale. Meanwhile, we recorded the streaming $x-y$ coordinates of the computer-mouse needed to reach the chosen anchor. According to the DYFRAT methodology, the linearized histograms of the collected radians were used to compute beta fuzzy sets for each item and for each participant (fuzzy sets were obtained by means of an heuristic optimization procedure which converts histograms into fuzzy sets). Finally, modes $\mathbf{m}_1, \ldots, \mathbf{m}_n$ and precisions $\mathbf{s}_1, \ldots, \mathbf{s}_n$ were used to represent final responses and uncertainties involved during the rating process. Figure 1 shows an example of observed beta fuzzy numbers for the RDB response variable.

Data analysis and results The fuzzy response variable rdb was computed by aggregating the fuzzy variables extreme and positioning of the RDB questionnaire in terms of mean (Hanss 2005). Similarly, fcrs was obtained by aggregating the crisp responses for the variables monitoring, messages, and communications from the FCRS questionnaire. To simplify interpretation of the results, variables drugs and fcrs were made categorical using median split on their crisp responses. This yielded to two new dichotomous variables, namely drugs (non-use/use) and fcrs (bad/satisfactory family climate). By contrast, the variable anger entered the model as a fuzzy variable in terms of mode and precision components. First, we run a model (*model 1*) where dispersion $\boldsymbol{\phi}$ was held fixed for all participants whereas the mean $\boldsymbol{\mu}$ was modeled using the fuzzy variable anger and the categorical variables fcrs and drugs. Table 4
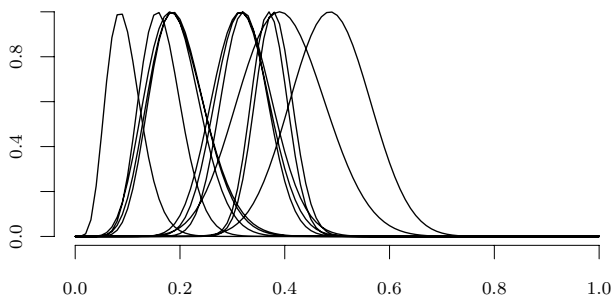

Fig. 1 Case study 1: Beta fuzzy responses on the RDB variable for a subsample of participants

shows the final estimates along with their standard errors computed using the empirical observed information matrix (see Eqs. 18–19) and Pearson's residuals. The results suggest that rdb increased as a function of mode components of anger ($\hat{\beta}_1 = 1.351$, $\sigma_{\hat{\beta}_1} = 0.356$) whereas its precision/spread components did not affect the response variable ($\hat{\beta}_2 = 0.001$, $\sigma_{\hat{\beta}_2} = 0.002$). Moreover, participants using drugs showed higher levels of rdb ($\hat{\beta}_4 = 0.458$, $\sigma_{\hat{\beta}_4} = 0.187$) when compared to those who did not use substances. As expected, participants with satisfactory family climate showed lower levels of rdb ($\hat{\beta}_3 = -0.192$, $\sigma_{\hat{\beta}_3} = 0.440$) when compared to participants with a bad family climate. To account for heterogeneity in the response variable rdb, two further models were estimated, one which included the dichotomous variable sex, and the other which also included driving_frequency. In order to evaluate models improvements, both the models were compared in terms of fuzzy likelihood-ratio test (Berkachy and Donzé 2019). Table 4 reports the results for model 2 and model 3. With regards to model 2, the likelihood-ratio test computed against model 1 reveled that sex improved the fit of the model ($\chi^2_{7-6=1} = 5.001$, $p = 0.025$, $\mathrm{AIC}_{\mathrm{model1}} = -163.38$, $\mathrm{AIC}_{\mathrm{model2}} = -166.38$), with men showing lower levels of heterogeneity in the response variable ($\gamma_1 = -1.199$, $\sigma_{\hat{\gamma}_1} = 0.613$) when compared to women. To further analyse the variability in reckless-driving behaviors, we asked whether this varied as a function of participants' frequency of driving. To assess this hypothesis, model 3 included driving_frequency as an additional term in the precision equation. The likelihood-ratio test conducted against model 2 revealed that driving frequency did not improve the fit of the model ($\chi^2_{8-7=1} = 1.626$, $p = 0.2021$, $\mathrm{AIC}_{\mathrm{model2}} = -166.38$, $\mathrm{AIC}_{\mathrm{model3}} = -166.01$). The results were in line with the literature (e.g., see: Bıçaksız and Özkan 2016) and suggested that self-reported reckless-driving behaviors were positively associated to driving anger ($\hat{\beta}_1 = 1.485$, $\sigma_{\hat{\beta}_1} = 0.378$) and substance use ($\hat{\beta}_4 = 0.518$, $\sigma_{\hat{\beta}_4} = 0.235$). By contrast, family climate acted as a protective factor with satisfactory family climate being negatively associated to risky behaviors ($\hat{\beta}_3 = -0.126$, $\sigma_{\hat{\beta}_3} = 0.157$). Interestingly, variability of self-reported responses varied as a function of gender, with female drivers showing more variable responses than male drivers ($\hat{\gamma}_1 = -1.999$, $\sigma_{\hat{\gamma}_1} = 0.613$). Figure 2 plots the predicted curves against the observed fuzzy data as a function of both continuous (panels A-B) and categorical (colors in both panels) predictors (note that only estimated modes $\hat{\boldsymbol{\mu}}$ were plotted for the sake of simplicity). To further investigate the role of fuzziness to predict self-reported reckless-driving behaviors (RDB), we contrasted the results of the final model (model 2) against the same model adapted on mean-based defuzzified response data (i.e., data without the components $\mathbf{s}_1, \ldots, \mathbf{s}_n$ for the decision uncertainty). Table 5 shows the final estimates computed using maximum-likelihood as implemented in standard beta regression model (Ferrari and Cribari-Neto 2004). The coefficients $\hat{\boldsymbol{\beta}}$ for the mean $\hat{\boldsymbol{\mu}}$ component were nearly similar to those obtained for the fuzzy case. By contrast, the coefficients $\hat{\boldsymbol{\gamma}}$ for the precisions were lower if compared to the fuzzy case. This is in line with the results of the simulation study. However, standard errors of the estimates for defuzzified data were smaller than those obtained for the fuzzy case, especially with regards to precision coefficients $\hat{\boldsymbol{\gamma}}$. Overall, this is

Table 4 Case study 1: Variable dispersion fuzzy beta models for reckless-driving behaviors

	fEM	
	Estimate	SE
Model 1		
Residuals quantiles: $Q1$: -6.789, Med: -0.162, $Q3$: 6.431		
$\hat{\mu}$		
β_0 (Intercept)	-2.140	0.282
anger (m)	1.351	0.356
anger (s)	0.001	0.002
fcrs (bad vs. satisfactory)	-0.192	0.144
drugs (non-use vs. use)	0.458	0.187
$\hat{\phi}$		
γ_0 (Intercept)	3.330	0.301
Model 2		
Residuals quantiles: $Q1$: -8.525, Med: -0.114, $Q3$: 4.860		
$\hat{\mu}$		
β_0 (Intercept)	-2.312	0.292
anger (m)	1.485	0.378
anger (s)	0.001	0.002
fcrs (bad vs. satisfactory)	-0.126	0.157
drugs (non-use vs. use)	0.518	0.235
$\hat{\phi}$		
γ_0 (Intercept)	4.069	0.472
sex (female vs. male)	-1.199	0.613
Model 3		
Residuals quantiles: $Q1$: -7.154, Med: -0.067, $Q3$: 4.253		
$\hat{\mu}$		
β_0 (Intercept)	-2.362	0.289
anger (m)	1.551	0.397
anger (s)	0.001	0.002
fcrs (bad vs. satisfactory)	-0.133	0.159
drugs (non-use vs. use)	0.567	0.252
$\hat{\phi}$		
γ_0 (Intercept)	3.972	0.485
sex (female vs. male)	-1.302	0.644
driving_frequency (always vs. weekend)	0.712	0.793

Note that categorical variables were codified using dummy coding with the following reference levels: fcrs (ref.: bad), drugs (ref.: non-use), and driving frequency (ref.: always). The parameters μ and ϕ were linked to the response variable using logit($\cdot$) and log($\cdot$) link functions, respectively
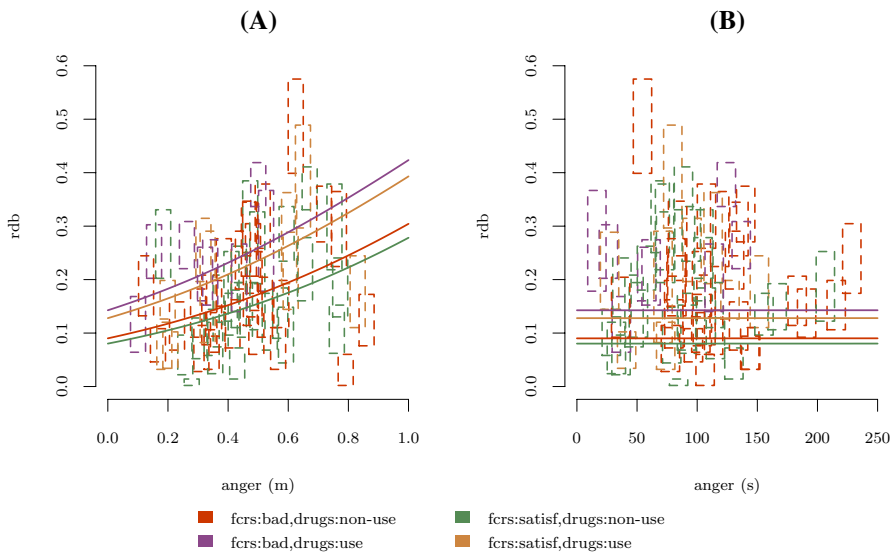
coherent with the estimation approach used for fuzzy data as, in this case, standard errors were computed using score Eqs. (16)–(17) which are in turn based on filtered data $(\mathbf{y}_1^*, \ldots, \mathbf{y}_n^*)$.

Table 5 Case study 1: Variable dispersion crisp beta model for reckless-driving behaviors adapted on defuzzified data

	dML	
	Estimate	SE
Model 2		
Residuals quantiles: $Q1: -0.3973$, Med: 0.172, $Q3: 0.636$		
$\hat{\boldsymbol{\mu}}$		
β_0 (Intercept)	-2.199	0.217
anger (m)	1.284	0.347
anger (s)	0.000	0.001
fcrs (bad vs. satisfactory)	-0.105	0.120
drugs (non-use vs. use)	0.530	0.145
$\hat{\boldsymbol{\phi}}$		
γ_0 (Intercept)	3.746	0.251
sex (female vs. male)	-1.082	0.330

Note that categorical variables were codified as for the fuzzy beta linear models (see Table 4). The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\phi}$ were linked to the response variable using logit($\cdot$) and log($\cdot$) link functions, respectively

(A)   **(B)**

fcrs:bad,drugs:non-use   fcrs:satisf,drugs:non-use
fcrs:bad,drugs:use   fcrs:satisf,drugs:use

Fig. 2 Case study 1: Observed fuzzy data for RDB as a function of fcrs and drugs categorical predictors (colors in both panels) and the two predictors for fuzzy anger (**a, b**). Fitted curves correspond to model 2 in Table 4. Note that rectangles represent α-cuts of the observed fuzzy data with $\alpha = 0.5$, i.e. $\mathbf{y}_i^\alpha = \left[\min \left(\{ y \in [0,1] : \xi_{\tilde{y}_i}(y) > 0.5 \} \right), \max \left(\{ y \in [0,1] : \xi_{\tilde{y}_i}(y) > 0.5 \} \right) \right]$

5.2 Case study 2: Service quality in restaurant industry

Service quality is an important factor to assess the performances of a company and it is of a great interest for restaurants services as well. Several research have been conducted to understand the overall effect of perceived service quality on customer satisfaction which, in turn, leads to positive consumption behaviors like revisiting and recommending the restaurant (Almohaimmeed 2017; Namkung and Jang 2008). A number of variables influencing dining experience have been suggested, such as food quality, menu variety, food presentation, quality of staff service, internal/external environments (Ha and Jang 2010). Last but not least, variables like prices and restaurant type (e.g., fast-food, fine restaurants) seem to be also involved in the heterogeneity of restaurant quality (Almohaimmeed 2017). In this application, we will consider a simple model for restaurant service quality where perceived quality of food (food) and staff's perception of being courteous (employees) were used to predict perceived service quality (service_quality). In addition, we also evaluated the extent to which heterogeneity in the response variable can be accounted by price levels (prices) and restaurant type (type).

Data and measures Data were originally collected by de Sáa et al. (2014) and refers to a survey of 14 items administered to a sample of $n = 70$ customers of different age, background, and occupation. The questionnaire included two of the most important factors of restaurant quality, namely food/beverage and service quality. We considered only complete cases, i.e. cases with no missing values. This yielded to a subset of $n = 49$ customers (31% women) with modal age between 25 and 34 years. Informal restaurants were about 67% of the total, 16% of them were fine, 10% fast-food, and 7% were self-service restaurants. About 69% of restaurants showed prices levels on average (about 15 Euro) whereas 31% of them reported higher prices. Ratings measures were collected by the authors using a Likert-type computerized fuzzy rating scale (de Sáa et al. 2014). In this case, participants provided their responses by means of a two-step direct rating procedure. First, they were asked to draw on the graphical pseudo-continuous scale the core of the set containing the most plausible rating values. Then, conditioned on the previous choice, they were asked to draw the support of the set, which instead represents the most compatible set for the rating values. Finally, both the intervals were interpolated to form a trapezoidal fuzzy response. For the purposes of this application, trapezoidal fuzzy numbers were converted into beta fuzzy numbers adopting a procedure minimizing the information content of the original fuzzy sets. In general, several transformation procedures are available to this purpose (Nasibov and Peker 2008). Here, for the sake of simplicity, we resorted to adopt the simplest approximation based on the minimization of the area under the curve between two fuzzy sets. In particular, the parameters $\{m, s\} \in (0, 1) \times \mathbb{R}^+$ of the beta fuzzy set $\xi_{\widetilde{B}}(y; m, s)$ that approximates a trapezoidal fuzzy set $\xi_{\widetilde{A}}(y; a, b, c, d)$ with real parameters $a < b < c < d$ were found by minimizing the function $\delta(m, s) = |\int_{A_0} \xi_{\widetilde{A}}(y; a, b, c, d) \mathrm{d}y - \int_{B_0} \xi_{\widetilde{B}}(y; m, s) dy|$ w.r.t. m and s. Similarly to the first application, modes $\mathbf{m}_1, \ldots, \mathbf{m}_n$ of the beta fuzzy numbers represent final participants' responses whereas their precisions $\mathbf{s}_1, \ldots, \mathbf{s}_n$ model the decision uncertainty occurred during the rating process. Figure 3 shows an

example of transformed beta fuzzy numbers along with the associated observed trapezoidal fuzzy sets for the `service_quality` response variable.

Data analysis and results To take into account decision uncertainty for predictors, in this case fuzzy variables `food` and `employees` were defuzzified using the centroid method before entering the model. A first model was run including `food` and `employees` as predictors for the mean and `type` and `price` for precision components. Table 6 shows the final estimates along with their standard errors. As expected, `service_quality` increased as a function of `food` ($\hat{\beta}_1 = 1.870$, $\sigma_{\hat{\beta}_1} = 0.856$) and `employees` ($\hat{\beta}_2 = 1.501$, $\sigma_{\hat{\beta}_2} = 0.718$). This is in line with previous studies on restaurant quality, suggesting that a higher perceived quality of food and staff predicts the perceived quality of restaurant services. In addition, the dispersion component of the model varied as a function of `type`, with informal restaurants showing higher heterogeneity in `service_quality` ($\gamma_1 = 1.084$, $\sigma_{\hat{\gamma}_1} = 1.055$) then self-service restaurants. The same applied for fine restaurants ($\gamma_3 = 0.865$, $\sigma_{\hat{\gamma}_3} = 1.158$) whereas fast-food showed decreasing levels of variability in `service_quality` ($\gamma_2 = -0.232$, $\sigma_{\hat{\gamma}_2} = 1.124$) when compared to self-service restaurants. Interestingly, variability in `service_quality` is positively associated to `price` ($\gamma_4 = 1.802$, $\sigma_{\hat{\gamma}_4} = 1.239$), with high-priced restaurants being more homogeneous in terms of perceived quality. This indicates that `service_quality` was not homogeneous over normal-priced restaurants and further covariates like internal/external atmospherics or image of restaurant might instead be needed to further account for such differences (Almohaimmeed 2017; Ha and Jang 2010). Figure 4 plots the predicted curves against the observed fuzzy data as a function of both continuous predictors (note that only estimated modes $\hat{\boldsymbol{\mu}}$ were plotted for the sake of simplicity). As for the previous case study, we compared the fuzzy beta model with the beta regression model adapted on mean-based defuzzified data. Table 7 shows the final estimates computed using maximum-likelihood as implemented in standard beta regression model (Ferrari and Cribari-Neto 2004). In line with the results from the simulation study, the estimated coefficients $\hat{\boldsymbol{\gamma}}$ for the precision component $\hat{\boldsymbol{\phi}}$ were lower than those for the fuzzy case and showed smaller standard errors. As stated previously, this can be interpreted in light of the estimation method used for the fuzzy case.
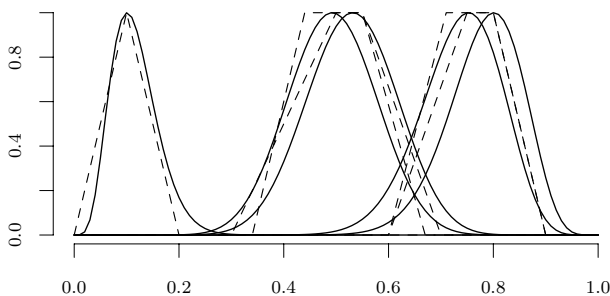

Fig. 3 Case study 2: Beta fuzzy responses (straight black lines) and trapezoidal fuzzy responses (dashed black lines) on the `service_quality` variable for a subsample of participants

Table 6 Case study 2: Variable dispersion fuzzy beta models for service quality in restaurant industry

	fEM	
	Estimate	SE
Model 1		
Residuals quantiles: $Q1: -5.061$, Med : 0.030, $Q3 : 7.102$		
$\hat{\boldsymbol{\mu}}$		
β_0 (Intercept)	− 1.509	0.622
food	1.804	0.875
employees	1.696	0.844
$\hat{\boldsymbol{\phi}}$		
γ_0 (Intercept)	1.047	2.923
type (self-service vs. informal)	2.267	1.779
type (self-service vs. fast-food)	− 0.232	1.124
type (self-service vs. fine)	0.865	1.158
price (high vs. on average)	1.802	1.239

Note that categorical variables were codified using dummy coding with the following reference levels: type (ref.: self-service), price (ref.: high). The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\phi}$ were linked to the response variable using logit($\cdot$) and log($\cdot$) link functions, respectively

Table 7 Case study 2: Variable dispersion crisp beta model for reckless-driving behaviors adapted on defuzzified data

	dML	
	Estimate	SE
Model 1		
Residuals quantiles: $Q1: -0.738$, Med: -0.098, $Q3 : 0.801$		
$\hat{\boldsymbol{\mu}}$		
β_0 (Intercept)	− 1.316	0.358
food	1.597	0.405
employees	1.484	0.383
$\hat{\boldsymbol{\phi}}$		
γ_0 (Intercept)	1.924	0.894
type (self-service vs. informal)	0.924	0.843
type (self-service vs. fast-food)	− 0.394	0.987
type (self-service vs. fine)	0.01	0.953
price (high vs. on average)	0.842	0.431

Note that categorical variables were codified as for the fuzzy beta linear models (see Table 6). The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\phi}$ were linked to the response variable using logit($\cdot$) and log($\cdot$) link functions, respectively

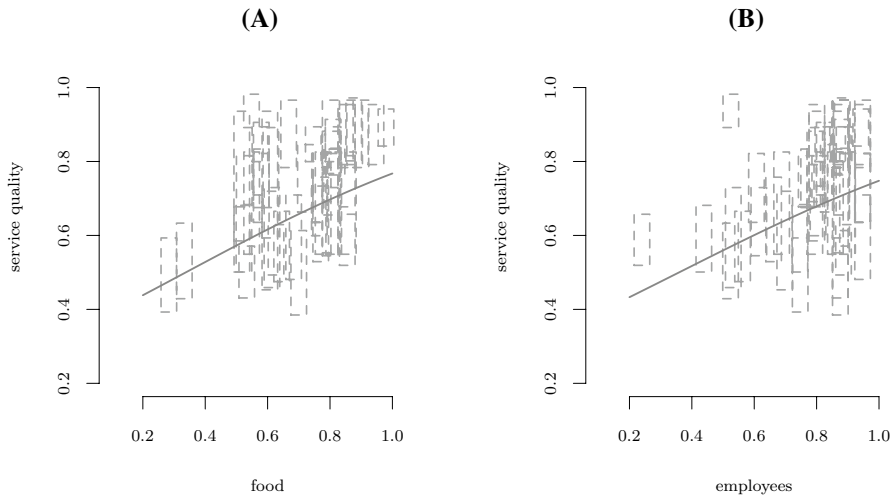**(A)**   **(B)**

Fig. 4 Case study 2: Observed fuzzy data for `service_quality` as a function of the two continuous predictors `food` (**a**) and `employees` (**b**). Fitted curves correspond to model 1 in Table 6. Note that rectangles represent α-cuts of the observed fuzzy data with $\alpha = 0.5$, i.e. $\mathbf{y}_i^\alpha = \left[\min\left(\{y \in [0,1] : \xi_{\tilde{y}_i}(y) > 0.5\}\right), \max\left(\{y \in [0,1] : \xi_{\tilde{y}_i}(y) > 0.5\}\right)\right]$

6 Conclusions

In this article we developed a statistical approach to deal with bounded continuous ratings data in the case of non-random uncertainty. In particular, beta fuzzy numbers were adopted to represent ratings data subject to decision uncertainty and the beta regression framework was used to model the random counterpart of the overall rating process. The fuzzy component of the data was then used to estimate the latent and non-fuzzy characteristics of the raters population. Parameters estimation was performed by maximum likelihood using a version of the Expectation–Maximization algorithm generalized for the case of fuzzy data. A simulation study and two real applications were used to highlight the characteristics of the proposed approach. The simulation study revealed that the fuzzy beta linear model showed more accurate results over a set of standard methods which can be applied in the case of fuzzy data. The applications showed how the proposed method can be adopted in real cases involving ratings data represented as fuzzy numbers.

A nice advantage of the proposed approach is its simplicity and flexibility in dealing with fuzzy data. Indeed, as it encapsulates decision rating uncertainty directly in $\tilde{\mathbf{y}}$, the fuzzy beta linear model does not require the extension of its parametric structure to account for modes $\mathbf{m}$ and precisions $\mathbf{s}$ of beta fuzzy data. Indeed, in the current proposal, the model's parameters $\theta = \{\beta, \gamma\}$ are not represented as fuzzy numbers. Consequently, parameters estimation and inference can still be performed using the asymptotic properties of maximum likelihood theory. In this setting, the fuzzy beta model recovers the parameters of the underlying rating process $Y \sim f_Y(y;\theta)$ by filtering the imprecise data in terms of expectation $\mathbb{E}_\theta\left[Y|\tilde{y}\right]$. However, although this constitutes an important advance, it should be noticed that fuzziness of the data is

not propagated to the output of the model. Indeed, as a consequence of the averaging approach upon which the fuzzy-EM is based, the predictions $\hat{\boldsymbol{y}}$ of the fuzzy beta model are non-fuzzy and they are made at the level of the underlying crisp rating mechanism $f_Y(y;\boldsymbol{\theta})$. This may limit the use of this approach in some circumstances, for example when researchers are interested in forecasting fuzziness based on current fuzzy data, or rather when components of fuzzy data play a different role in predicting the variable response (e.g., modes and precisions/spreads of the outcome variable interact with those of the explanatory variables in some way). In all these cases, different statistical approaches may instead be preferred, like those based on full-likelihood approaches for imprecise data (e.g., see Denoeux 2014; Kanjanatarakul et al. 2016) or min–max based estimation methods (e.g., see Guillaume and Dubois 2020).

Various possible extensions of our approach can be considered in future works. For instance, fuzzy beta model involving fuzzy data with different shapes simultaneously (e.g., beta, triangular, trapezoidal) would offer a way to deal with more complex scenarios. Another future generalization which might be interesting to investigate is the case where fuzziness in ratings data is coupled with random uncertainty which varies as a function of subgroups in the model (like for CUBE models, see Piccolo and Simone 2019). This would offer the opportunity to further decompose the overall uncertainty in ratings responses in terms of different and possibly interacting components underlying participants' rating processes. Finally, an attractive extension of the current approach would consider the case of multivariate fuzzy beta models where joint fuzzy sets do not obey to the product rule (see Eq. 4). In this case, a fuzzy copula representation may instead be used to formally represent the joint fuzziness information (e.g., see Ranjbar and Hesamian 2017).

References

Aiken, L.R.: Rating Scales and Checklists: Evaluating Behavior, Personality, and Attitudes. Wiley, New York (1996)

Aldrich, J.H., Nelson, F.D.: Linear Probability, Logit, and Probit Models, p. 45. Sage, Beverly Hills (1984)

Algamal, Z.Y.: A particle swarm optimization method for variable selection in beta regression model. Electron. J. Appl. Stat. Anal. **12**(2), 508–519 (2019)

Alimi, A.M.: Beta neuro-fuzzy systems. TASK Q. J. Special Issue Neural Netw. **7**(1), 23–41 (2003)

Almohaimmeed, B.M.: Restaurant quality and customer satisfaction. Int. Rev. Manag. Market. **7**(3), 42–49 (2017)

Angel, S., Disslbacher, F., Humer, S., Schnetzer, M.: What did you really earn last year? Explaining measurement error in survey income data. J. R. Stat. Soc. Ser. A (Stat. Soc.) **182**(4), 1411–1437 (2019)

Augustin, T., Coolen, F.P., De Cooman, G., Troffaes, M.C. (eds.) Introduction to Imprecise Probabilities. Wiley (2014)

Baklouti, N., Abraham, A., Alimi, A.M.: A beta basis function interval type-2 fuzzy neural network for time series applications. Eng. Appl. Artif. Intell. **71**, 259–274 (2018)

Berkachy, R., Donzé, L.: Fuzzy confidence interval estimation by likelihood ratio. In: 2019 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (EUSFLAT 2019), Atlantis Press (2019)

Bıçaksız, P., Özkan, T.: Impulsivity and driver behaviors, offences and accident involvement: a systematic review. Transp. Res. Part F Traffic Psychol. Behav. **38**, 194–223 (2016)

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M., Bolker, B.M.: glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. **9**(2), 378–400 (2017)

Buckley, J.J.: Fuzzy Probability and Statistics. Springer, Berlin (2006)

Calcagnì, A., Lombardi, L.: Dynamic fuzzy rating tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales. Appl. Soft Comput. **24**, 948–961 (2014)

Calcagnì, A., Lombardi, L., Pascali, E.: Non-convex fuzzy data and fuzzy statistics: a first descriptive approach to data analysis. Soft Comput. **18**(8), 1575–1588 (2014)

Calcagnì, A., Cao, N., Rubaltelli, E., Lombardi, L.: A psychometric modeling approach to fuzzy rating data. arxiv:2102.04273 [stat.AP] (2021)

Canterle, D.R., Bayer, F.M.: Variable dispersion beta regressions with parametric link functions. Stat. Pap. **60**(5), 1541–1567 (2019)

Cattaneo, M.E.: The likelihood interpretation as the foundation of fuzzy set theory. Int. J. Approx. Reason. **90**, 333–340 (2017)

Chukhrova, N., Johannssen, A.: Randomized versus non-randomized hypergeometric hypothesis testing with crisp and fuzzy hypotheses. Stat. Pap. **61**, 2605–2641 (2018)

Chwastyk, A., Kosiński, W.: Fuzzy calculus with applications. Math. Appl. **41**, 1 (2013)

Coletti, G., Scozzafava, R.: Conditional probability, fuzzy sets, and possibility: a unifying view. Fuzzy Sets Syst. **144**(1), 227–249 (2004)

Couso, I., Dubois, D.: Statistical reasoning with set-valued information: ontic vs. epistemic views. Int. J. Approx. Reason. **55**(7), 1502–1518 (2014)

Couso, I., Dubois, D., Sánchez, L.: Random sets and random fuzzy sets as ill-perceived random variables. In: Springer-Briefs in Computational Intelligence. Springer, Berlin (2014)

de Sáa S.R., Gil M.Á., González-Rodríguez, G., López, M.T., Lubiano, M.A.: Fuzzy rating scale-based questionnaires and their statistical analysis. IEEE Trans. Fuzzy Syst. **23**(1), 111–126 (2014)

De Bruin, W.B., Manski, C.F., Topa, G., Van Der Klaauw, W.: Measuring consumer uncertainty about future inflation. J. Appl. Econ. **26**(3), 454–478 (2011)

Deffenbacher, J.L., Oetting, E.R., Lynch, R.S.: Development of a driving anger scale. Psychol. Rep. **74**(1), 83–91 (1994)

Denœux, T.: Maximum likelihood estimation from fuzzy data using the EM algorithm. Fuzzy Sets Syst. **183**(1), 72–91 (2011)

Denoeux, T.: Likelihood-based belief function: justification and some extensions to low-quality data. Int. J. Approx. Reason. **55**(7), 1535–1547 (2014)

Dombi, J., Jónás, T.: Flexible fuzzy numbers for likert scale-based evaluations. In: International Workshop Soft Computing Applications, pp. 81–101. Springer, Berlin (2018)

Dubois, D., Prade, H.: Operations on fuzzy numbers. Int. J. Syst. Sci. **9**(6), 613–626 (1978)

Eid, M., Zickar, M.J.: Detecting response styles and faking in personality and organizational assessments by mixed Rasch models. In: Multivariate and Mixture Distribution Rasch Models, pp. 255–270. Springer, Berlin (2007)

Feng, Z., Zhang, J., Chen, Q.: Statistical inference for linear regression models with additive distortion measurement errors. Stat. Pap. **61**, 2483–2509 (2018)

Ferrari, S., Cribari-Neto, F.: Beta regression for modelling rates and proportions. J. Appl. Stat. **31**(7), 799–815 (2004)

Freeman, J.B., Ambady, N.: Mousetracker: software for studying real-time mental processing using a computer mouse-tracking method. Behav. Res. Methods **42**(1), 226–241 (2010)

Furnham, A.: Response bias, social desirability and dissimulation. Pers. Individ. Differ. **7**(3), 385–400 (1986)

Gebhardt, J., Gil, M.A., Kruse, R.: Fuzzy set-theoretic methods in statistics. In: Fuzzy Sets in Decision Analysis, Operations Research and Statistics, pp. 311–347. Springer, Berlin (1998)

Gil, M.A., López-Díaz, M., Ralescu, D.A.: Overview on the development of fuzzy random variables. Fuzzy Sets Syst. **157**(19), 2546–2557 (2006)

Golia, S.: On the interpretation of the uncertainty parameter in CUB models. Electron. J. Appl. Stat. Anal. **8**(3), 312–328 (2015)

González-Rodríguez, G., Colubi, A., Gil, M.Á.: A fuzzy representation of random variables: an operational tool in exploratory analysis and hypothesis testing. Comput. Stat. Data Anal. **51**(1), 163–176 (2006)

Guillaume, R., Dubois, D.: A min–max regret approach to maximum likelihood inference under incomplete data. Int. J. Approx. Reason. **121**, 135–149 (2020)

Ha, J., Jang, S.S.: Effects of service quality and food quality: the moderating role of atmospherics in an ethnic restaurant segment. Int. J. Hosp. Manag. **29**(3), 520–529 (2010)

Hanss, M.: Applied Fuzzy Arithmetic. Springer, Berlin (2005)

Hesamian, G., Shams, M.: A note on fuzzy probability of a fuzzy event. Int. J. Intell. Syst. **32**(7), 676–685 (2017)

Hesketh, T., Pryor, R., Hesketh, B.: An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences. Int. J. Man-Mach. Stud. **29**(1), 21–35 (1988)

Iannario, M.: Modelling scale effects and uncertainty in rating surveys. Electron. J. Appl. Stat. Anal. **8**(3), 329–345 (2015)

Kahneman, D., Tversky, A.: Variants of uncertainty. Cognition **11**(2), 143–157 (1982)

Kanjanatarakul, O., Denoeux, T., Sriboonchitta, S.: Prediction of future observations using belief functions: a likelihood-based approach. Int. J. Approx. Reason. **72**, 71–94 (2016)

Kruse, R., Meyer, K.D.: Statistics with Vague Data, vol. 6. Springer, Berlin (1987)

Lin, T.Y., Cercone, N.: Rough Sets and Data Mining: Analysis of Imprecise Data. Springer, Berlin (2012)

Liu, B.: Uncertainty theory. In: Uncertainty Theory, pp. 1–79. Springer, Berlin (2010)

Lombardi, L., Pastore, M., Nucci, M., Bobbio, A.: SGR modeling of correlational effects in fake good self-report measures. Methodol. Comput. Appl. Probab. **17**(4), 1037–1055 (2015)

Lubiano, M.A., de Sáa SR, Montenegro M, Sinova B, Gil MÁ: Descriptive analysis of responses to items in questionnaires. Why not using a fuzzy rating scale? Inf. Sci. **360**, 131–148 (2016)

McCulloch, C.E.: Generalized linear models. J. Am. Stat. Assoc. **95**(452), 1320–1324 (2000)

McFadden, D. Conditional logit analysis of qualitative choice be. In: Zarembka, P. (eds.) Frontiers in Econometrics, pp. 105–142. Academic Press, New York (1973)

McLachlan, G.J., Peel, D.: Finite Mixture Models. Wiley, New York (2004)

McNally, B., Bradley, G.L.: Re-conceptualising the reckless driving behaviour of young drivers. Accid. Anal. Prev. **70**, 245–257 (2014)

Meilijson, I.: A fast improvement to the EM algorithm on its own terms. J. R. Stat. Soc. Ser. B (Methodol.) **51**(1), 127–138 (1989)

Migliorati, S., Di Brisco, A.M., Ongaro, A., et al.: A new regression model for bounded responses. Bayesian Anal. **13**(3), 845–872 (2018)

Miller, D.C., Salkind, N.J.: Handbook of Research Design and Social Measurement. Sage, Beverly Hills (2002)

Muthukumarana, S., Swartz, T.B.: Bayesian analysis of ordinal survey data using the Dirichlet process to account for respondent personality traits. Commun. Stat.-Simul. Comput. **43**(1), 82–98 (2014)

Najafi, Z., Taheri, S.M., Mashinchi, M.: Likelihood ratio test based on fuzzy data. Int. J. Intell. Technol. Appl. Stat. **3**(3), 285–301 (2010)

Namkung, Y., Jang, S.S.: Are highly satisfied restaurant customers really different? A quality perception perspective. Int. J. Contemp. Hosp. Manag. **20**(2), 142–155 (2008)

Nasibov, E.N., Peker, S.: On the nearest parametric approximation of a fuzzy number. Fuzzy Sets Syst. **159**(11), 1365–1375 (2008)

Ospina, R., Ferrari, S.L.: A general class of zero-or-one inflated beta regression models. Comput. Stat. Data Anal. **56**(6), 1609–1623 (2012)

Pendrill, L., Petersson, N.: Metrology of human-based and other qualitative measurements. Meas. Sci. Technol. **27**(9), 094,003 (2016)

Pendrill, L.R.: Using measurement uncertainty in decision-making and conformity assessment. Metrologia **51**(4), S206 (2014)

Piccolo, D., Simone, R.: The class of CUB models: statistical foundations, inferential issues and empirical evidence. Stat. Methods Appl. **28**(3), 389–435 (2019)

Piccolo, D., Simone, R., Iannario, M.: Cumulative and CUB models for rating data: a comparative analysis. Int. Stat. Rev. **87**(2), 207–236 (2019)

Preston, C.C., Colman, A.M.: Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychol. **104**(1), 1–15 (2000)

Quost, B., Denoeux, T.: Clustering and classification of fuzzy data using the fuzzy EM algorithm. Fuzzy Sets Syst. **286**, 134–156 (2016)

Rabinowitz, J., Schooler, N.R., Brown, B., Dalsgaard, M., Engelhardt, N., Friedberger, G., Kinon, B.J., Lee, D., Ockun, F., Mahableshwarkar, A., et al.: Consistency checks to improve measurement with the Montgomery–Asberg depression rating scale (MADRS). J. Affect. Disord. **256**, 143–147 (2019)

Ranjbar, V., Hesamian, G.: Copula function for fuzzy random variables: applications in measuring association between two fuzzy random variables. In: Statistical Papers, pp. 1–20 (2017)

Saal, F.E., Downey, R.G., Lahey, M.A.: Rating the ratings: assessing the psychometrics quality of rating data. Psychol. Bull. **88**(2), 413–428 (1980)

Schulte-Mecklenbeck, M., Kühberger, A., Johnson, J.G.: A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide. Psychology Press (2011)

Scott-Parker, B., King, M.J., Watson, B.: The psychosocial purpose of driving and its relationship with the risky driving behaviour of young novice drivers. Transp. Res. Part F Traffic Psychol. Behav. **33**, 16–26 (2015)

Singpurwalla, N.D., Booker, J.M.: Membership functions and probability measures of fuzzy sets. J. Am. Stat. Assoc. **99**(467), 867–877 (2004)

Slovic, P., Finucane, M.L., Peters, E., MacGregor, D.G.: Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. Risk Anal. Int. J. **24**(2), 311–322 (2004)

Stein, W.E.: Fuzzy probability vectors. Fuzzy Sets Syst. **15**(3), 263–267 (1985)

Su, Z., Zheng, S., Wang, P.: Likelihood-based multivariate fuzzy model with linear inequality constraints. J. Intell. Fuzzy Syst. **27**(5), 2191–2209 (2014)

Su, Z.G., Wang, P.H., Li, Y.G., Zhou, Z.K.: Parameter estimation from interval-valued data using the expectation–maximization algorithm. J. Stat. Comput. Simul. **85**(2), 320–338 (2015)

Taubman-Ben-Ari, O.: Attitudes toward accompanied driving: the views of teens and their parents. Transp. Res. Part F Traffic Psychol. Behav. **13**(4), 269–276 (2010)

Taubman-Ben-Ari, O.: How are meaning in life and family aspects associated with teen driving behaviors? Transp. Res. Part F Traffic Psychol. Behav. **24**, 92–102 (2014)

Taubman-Ben-Ari, O., Katz-Ben-Ami, L.: Family climate for road safety: a new concept and measure. Accid. Anal. Prev. **54**, 1–14 (2013)

Toroyan, T.: Global Status Report on Road Safety. World Health Organization, Department of Violence and Injury Prevention and Disability, Supporting a decade of action Geneva (2015)

Ülkümen, G., Fox, C.R., Malle, B.F.: Two dimensions of subjective uncertainty: clues from natural language. J. Exp. Psychol. Gen. **145**(10), 1280 (2016)

Veall, M.R., Zimmermann, K.F.: Evaluating pseudo-r 2's for binary probit models. Qual. Quant. **28**(2), 151–164 (1994)

Vonglao, P.: Application of fuzzy logic to improve the likert scale to measure latent variables. Kasetsart J. Soc. Sci. **38**(3), 337–344 (2017)

Vose, D.: Risk Analysis: A Quantitative Guide. Wiley, New York (2008)

Yu, S.C., Wu, B.: Fuzzy item response model: a new approach to generate membership function to score psychological measurement. Qual. Quant. **43**(3), 381 (2009)

Zadeh, L.A.: Probability measures of fuzzy events. J. Math. Anal. Appl. **23**(2), 421–427 (1968)

Zafar, B.: Can subjective expectations data be used in choice models? Evidence on cognitive biases. J. Appl. Econ. **26**(3), 520–544 (2011)

Zeileis, A., Cribari-Neto, F., Grün, B., Kos-midis, I.: Beta regression in R. J. Stat. Softw. **34**(2), 1–24 (2010)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.