A Bayesian beta linear model for fuzzy rating responses

Antonio Calcagnì,
Massimiliano Pastore, Gianmarco Altoè, Livio Finos

University of Padova

SIS | 2022
51st Scientific Meeting
of the Italian Statistical Society

Università
degli Studi
di Padova

Rating data are common when measuring human-based characteristics where attitudes, motivations, satisfaction, or beliefs are quantified using rating scales.

Rating data are common when measuring human-based characteristics where attitudes, motivations, satisfaction, or beliefs are quantified using rating scales.

A typical example is that of rating the question:

- *I am satisfied with my life* -

using the graded scale:

Strongly disagree — Disagree — 0 — Agree — Strongly agree

Rating data are common when measuring human-based characteristics where attitudes, motivations, satisfaction, or beliefs are quantified using rating scales.

A typical example is that of rating the question:

- *I am satisfied with my current work* -

using the graded scale:

Strongly disagree — Disagree — 0 — Agree — Strongly agree

As they involve human raters, rating data are often affected by **fuzziness** because of the **decision uncertainty** that affects the **response process**.

Several methods might be adopted to quantify fuzziness (**fuzzy scaling**):

- direct fuzzy rating [3]
- implicit fuzzy rating [1]
- deterministic crisp-to-fuzzy conversion systems [7]
- statistically-oriented crisp-to-fuzzy conversion systems [8]

Besides their differences, all these approaches aim at quantifying the fuzziness present in rating data.

Università degli Studi di Padova

Recently, **fuzzy-IRTree** has been proposed as a new methodology to represent as much information as possible from the rating response process [2].

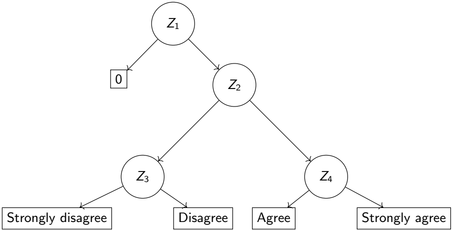
Università degli Studi di Padova

Recently, **fuzzy-IRTree** has been proposed as a new methodology to represent as much information as possible from the rating response process [2].

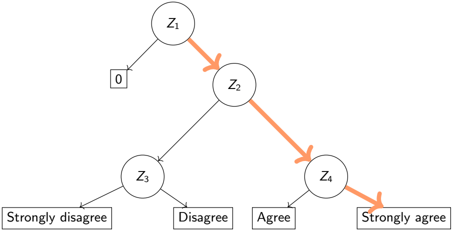
Key idea: The entire response process can be modeled stage-wise by means of an *Item Response Theory tree*:

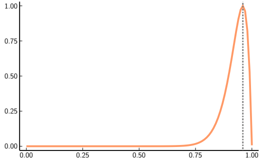
Binary trees + Rasch psychometric model

IRTree-based fuzzy numbers



Fuzziness arises as a result of the **transitions probabilities** estimated by the IRTree (i.e., *the easier the transition, the lesser the fuzziness*).

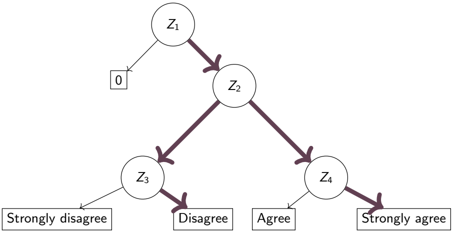




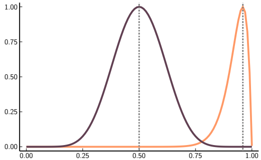
Fuzziness arises as a result of the **transitions probabilities** estimated by the IRTree (i.e., *the easier the transition, the lesser the fuzziness*).

An example of rating response with **low degree of fuzziness**.

Università degli Studi di Padova



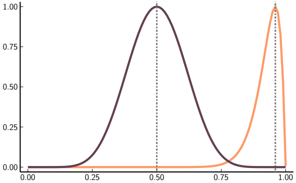
Fuzziness arises as a result of the **transitions probabilities** estimated by the IRTree (i.e., *the easier the transition, the lesser the fuzziness*).



An example of rating response with **low degree of fuzziness**.

An example of rating response with **high degree of fuzziness**.

UNIVERSITÀ DEGLI STUDI DI PADOVA



Fuzzy numbers are derived from the **estimated parameters** of the IRTree model. Further details in [2].

Several choices are available to this purpose. **Beta fuzzy numbers** are a good candidate in terms of flexibility and simplicity:

$$\xi_{\widetilde{y}_{ij}}(y) = \frac{1}{C}\; y^{m_{ij}s_{ij}}(1-y)^{s_{ij}-s_{ij}m_{ij}} \qquad m_{ij} \in (0,1) \quad s_{ij} \in \mathbb{R}^+$$

Università degli Studi di Padova

For a non-fuzzy collection of i.i.d. $(0,1)$-realizations $\mathbf{y} = (y_1, \ldots, y_n)$, the **Beta density** is as follows:

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

where

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

where $\boldsymbol{\mu} \in (0,1)^n$ is the $n \times 1$ vector of location parameters and $\boldsymbol{\phi} \in (0,\infty)^n$ the $n \times 1$ vector of precision parameters, which have been linearly expanded to account for covariates $\mathbf{X}_{n\times p}$ and $\mathbf{Z}_{n\times q}$.

Università degli Studi di Padova

Model parameters: $\boldsymbol{\theta} \in \{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in \mathbb{R}^p \times \mathbb{R}^q$

To sample from the posterior density of the model parameters $f(\boldsymbol{\beta}, \boldsymbol{\gamma})$, an **adaptive Metropolis-Hastings** algorithm has been used [5], with the **transition densities** of the MCMCs being defined as:

$$q(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t-1)}) = \mathcal{N}(;\boldsymbol{\theta}^{(t-1)}, \boldsymbol{\Sigma}^{(t)})$$

The acceptance ratio of the sampler is:

$$\alpha^{(t)} = \frac{\tilde{\mathcal{L}}(\boldsymbol{\theta}^{(t)}; \mathbf{m}, \mathbf{s})\ q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}^{(t)})\ f(\boldsymbol{\theta}^{(t)})}{\tilde{\mathcal{L}}(\boldsymbol{\theta}^{(t-1)}; \mathbf{m}, \mathbf{s})\ q(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t-1)})\ f(\boldsymbol{\theta}^{(t-1)})}$$

where $f(\boldsymbol{\theta})$ is a prior density ascribed to the model parameters.

UNIVERSITÀ DEGLI STUDI DI PADOVA

Since non-fuzzy realizations **y** are thought as unobserved random quantities, in this case the likelihood function is as follows [4]:

$$\tilde{\mathcal{L}}(\boldsymbol{\theta}^{(t)};\mathbf{m},\mathbf{s}) = \prod_{i=1}^{n} \int_0^1 \xi_{\widetilde{y}_i}(y;m_i,s_i)\frac{\Gamma(\phi_i)y^{(\mu_i\phi_i-1)}(1-y)^{(\phi_i-\mu_i\phi_i-1)}}{\Gamma(\phi_i\mu_i)\Gamma(\phi_i-\mu_i\phi_i)}\,dy$$

where

$$\boldsymbol{\xi}_{\tilde{\mathbf{y}}}(\mathbf{y};\mathbf{m},\mathbf{s}) = \prod_{i=1}^{n}\xi_{\widetilde{y}_i}(y;m_i,s_i)$$

follows from the **non-interactive** assumption of fuzzy observations.

UNIVERSITÀ DEGLI STUDI DI PADOVA

Aim: Investigate predictors of self-report (fuzzy) sexual intimacy [6].

Sample: $n = 450$ participants from Flanders (73% female, mean age 32.9 years, mean relationship length 7.68 years).

Predictors: (i) perceived *partner responsiveness*, (ii) sexual *desire*, (iii) avoidant *attachment score*.

7-point rating scales from 1 ("definitely not") to 7 ("yes, definitely").

Università degli Studi di Padova

Data analysis: Three additive Beta linear models M1-M3 have been defined to predict sexual intimacy. The models differ in terms of covariates for the term $\boldsymbol{\mu}$ whereas $\boldsymbol{\phi} = \exp(\mathbf{1}\gamma)$.

For all the models, $f(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}; \mathbf{0}, \mathbf{I}10)$ and $f(\gamma) = \mathcal{N}(\gamma; 0, 3)$.

Four parallel MCMCs have been run with 20000 samples and 5000 samples for the burn-in phase.

The final model has been chosen according to the LOO information criterion.

Results

According to the Gelman and Rubin's convergence diagnostics, all the chains reached the convergence (i.e., $\hat{R} = 1.00$).

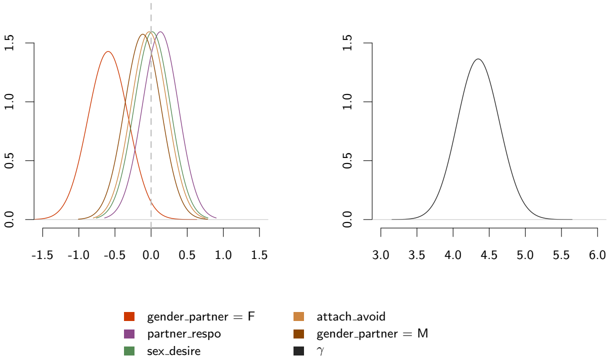
Models comparison:

Model	Covariates	LOO
M1	`partner_respo`, `sex_desire`	873.80
M2	`partner_respo`, `sex_desire`, `attach_avoid`	863.50
M3	`partner_respo`, `sex_desire`, `attach_avoid`, `gender_partner`	**857.00**

Università degli Studi di Padova

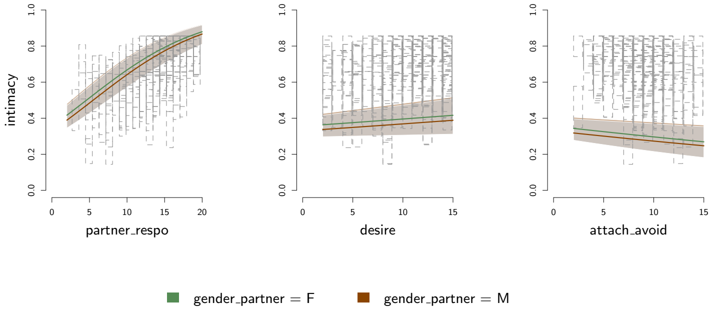
Results

Posterior parameters densities:



Results

Fitted vs. observed fuzzy data (rectangular sections):



Conclusions

- When coupled with standard probability theory, fuzzy numbers can deal with different sources of uncertainty in the same statistical model (i.e., randomness and imprecision).

- With regards to ratings data, fuzzy numbers provide flexible formal representations which might be used to integrate several information from the rating process (e.g., final response, decision uncertainty).

Università degli Studi di Padova

[1] Calcagnì, A., and Lombardi, L.
Dynamic fuzzy rating tracker (dyfrat): a novel methodology for modeling real-time dynamic cognitive processes in rating scales.
Applied soft computing 24 (2014), 948–961.

[2] Calcagnì, A., Cao, N., Rubaltelli, E., and Lombardi, L.
A psychometric modeling approach to fuzzy rating data.
Fuzzy Sets and Systems (2022), 1–24 (In Press).

[3] de Sáa, S. d. l. R., Gil, M. Á., González-Rodríguez, G., López, M. T., and Lubiano, M. A.
Fuzzy rating scale-based questionnaires and their statistical analysis.
IEEE Transactions on Fuzzy Systems 23, 1 (2014), 111–126.

[4] Gil, M., Corral, N., and Gil, P.
The minimum inaccuracy estimates in $\chi 2$ tests for goodness of fit with fuzzy observations.
Journal of Statistical Planning and Inference 19, 1 (1988), 95–115.

[5] Haario, H., Saksman, E., and Tamminen, J.
An adaptive metropolis algorithm.
Bernoulli (2001), 223–242.

[6] Van Lankveld, J. J., Dewitte, M., Verboon, P., and van Hooren, S. A.
Associations of intimacy, partner responsiveness, and attachment-related emotional needs with sexual desire.
Frontiers in Psychology 12 (2021).

[7] Vonglao, P.
Application of fuzzy logic to improve the likert scale to measure latent variables.
Kasetsart Journal of Social Sciences 38, 3 (2017), 337–344.

[8] Yu, S.-C., and Wu, B.
Fuzzy item response model: a new approach to generate membership function to score psychological measurement.
Quality and Quantity 43, 3 (2009), 381.

Università degli Studi di Padova

https://psicostat.dpss.psy.unipd.it/

antonio.calcagni@unipd.it

Università degli Studi di Padova