Approximated Gibbs sampling for continuous fuzzy numbers

Antonio Calcagnì
University of Padova

Przemyslaw Grzegorzewski
Warsaw University of Technology

Introduction

Fuzzy data are ubiquitous in many research contexts, including social and behavioral sciences.

Rating data are a typical example of fuzzy data, as the process of measuring human attitudes, motivations, or beliefs involve a certain **degree of uncertainty** and fuzziness.

Fuzzy data are also common in **classification-based problem**, such as when precise data are classified into imprecise categories (e.g., images or scenes classification, content analysis, human-based assessments).

Introduction

In these cases, **statistical models** have to cope with fuzzy data and appropriate methods need to be used in order to make **inference** appropriately.

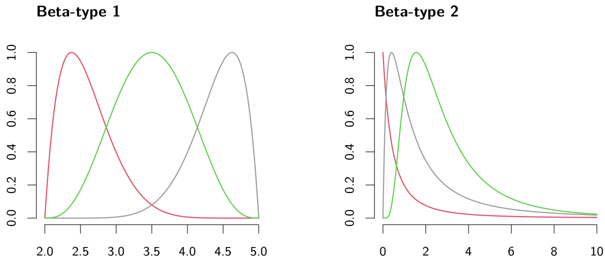
Several methods have been proposed over the years, most of them based on **generalization of likelihood theory** to fuzzy samples [1, 2].

However, statistical **estimators** often suffer from **excessive variance** (i.e., larger standard errors) especially when **epistemic fuzzy data** are considered [3].

Goal:

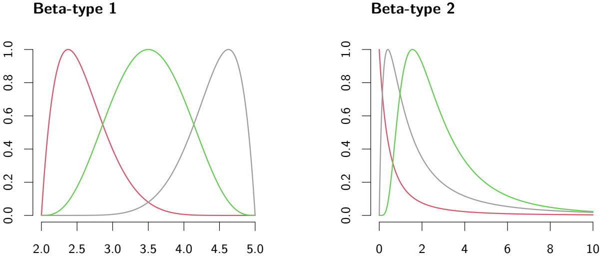
- Define a probabilistic schema to mimic the sampling process underlying epistemic fuzzy data
- Use this mechanism to make inference on the parameters of statistical models
- Plug this mechanism into statistical estimators to reduce variance of estimated parameters

The proposed solution uses Beta-type fuzzy numbers as a general template for representing **continuous** and **unimodal** fuzzy numbers.



A conditional sampling schema

The Beta-type fuzzy numbers



- Flexible and parsimonious as they require **two parameters** only (m: mode; s: precision)
- Deal with variables supported on **bounded** or **semi-infinite** intervals (as those commonly used in social and behavioral research)
- Generalize **triangular** fuzzy numbers as well

A conditional sampling schema

Statement of the problem

Let $Y_1, \ldots, Y_n$ be n independent continuous r.vs. and $\tilde{\mathbf{y}} = (\tilde{y}_1, \ldots, \tilde{y}_n)$ a sample of fuzzy observations. The vector $\tilde{\boldsymbol{y}}$ is a **blurred** version of $\mathbf{y}$ because of *post-sampling* or epistemic uncertainty-based processes.

The interest lies in studying $f_{Y_1,\ldots,Y_n}(\mathbf{y}; \boldsymbol{\theta}_{\boldsymbol{y}})$ with the purpose of making inference on $\boldsymbol{\theta}_{\boldsymbol{y}}$ given the fuzzy sample $\tilde{\boldsymbol{y}}$.

Each fuzzy observation $\tilde{y}_i$ consists of mode and precision $\{m_i, s_i\}$ of a Beta-type fuzzy number.

The idea is to use a **conditional schema** linking the parameters of fuzzy numbers (i.e., mode m and precision s) to $f_{Y_1,\ldots,Y_n}(\mathbf{y};\boldsymbol{\theta}_{\mathbf{y}})$:

The idea is to use a **conditional schema** linking the parameters of fuzzy numbers (i.e., mode m and precision s) to $f_{Y_1,\ldots,Y_n}(\mathbf{y};\boldsymbol{\theta}_{\boldsymbol{y}})$:

$$
\begin{aligned}
y_i &\sim f_Y(y;\boldsymbol{\theta}_{\boldsymbol{y}}) \\
s_i &\sim f_S(s;\boldsymbol{\theta}_{\boldsymbol{s}}) \\
m_i|y_i, s_i &\sim f_{M|S,Y}(m;\omega(y,s))
\end{aligned}
$$

$$\boxed{y_i \sim f_Y(y; \boldsymbol{\theta_y})}$$
$$s_i \sim f_S(s; \boldsymbol{\theta_s})$$
$$m_i | y_i, s_i \sim f_{M|S,Y}(m; \omega(y, s))$$

Rv governing the stochastic (**non-fuzzy**) sampling process. The parameters can be expressed as a function of external covariates $\boldsymbol{\theta_y} = g^{-1}(\mathbf{X}\beta)$ as for GLMs.

It depends on the specific problem one is dealing with (e.g., Beta distribution, Logistic distribution, Weibull distribution).

$$y_i \sim f_Y(y; \boldsymbol{\theta}_y)$$

$$\boxed{s_i \sim \mathcal{G}a(s; \alpha_s, \beta_s)}$$

$$m_i | y_i, s_i \sim f_{M|S,Y}(m; \omega(y,s))$$

Gamma distribution with $\alpha_s > 0$ and $\beta_s > 0$ modeling the precision (or spread) of the fuzzy number. In the simplest case, $s_i \perp\!\!\!\perp y_i$ although it can be generalized to cope with cases where s_i depends on y_i or external covariates.

$$y_i \sim f_Y(y; \boldsymbol{\theta_y})$$
$$s_i \sim \mathcal{G}a(s; \alpha_s, \beta_s)$$
$$\boxed{m_i | y_i, s_i \sim f_{M|S,Y}(m; \omega(y, s))}$$

Rv for the mode of the fuzzy number as a function of the true unobserved outcome y_i and the spread s_i.

A conditional sampling schema

$$y_i \sim f_Y(y;\boldsymbol{\theta_y})$$
$$s_i \sim \mathcal{G}a(s;\alpha_s,\beta_s)$$
$$\boxed{m_i|y_i,s_i \sim \mathcal{B}e_{4P}(m;s_iy_i,s_i-s_iy_1,lb,ub)}$$

Rv governing the mode of the fuzzy number as a function of the true unobserved outcome y_i and the spread s_i.

Case 1: $y \in (lb,ub) \subset \mathbb{R}$ the **four-parameter Beta distribution** is used.

$$y_i \sim f_Y(y;\boldsymbol{\theta_y})$$
$$s_i \sim \mathcal{G}a(s;\alpha_s,\beta_s)$$
$$\boxed{m_i|y_i,s_i \sim \mathcal{B}e_P(m;y_i+y_is_i,s_i+2)}$$

Rv governing the mode of the fuzzy number as a function of the true unobserved outcome y_i and the fuzziness s_i.

Case 2: $y \in (0,+\infty)$ the **Beta prime distribution** is instead used.

A conditional sampling schema

Proposed solution

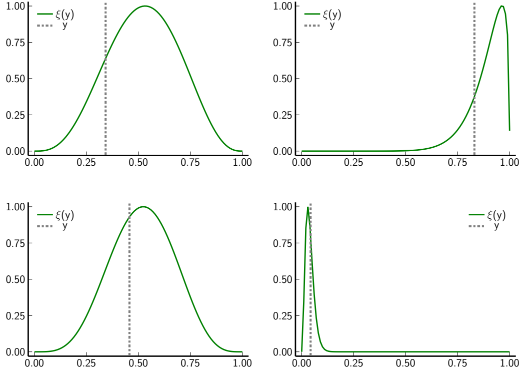
$$y_i \sim f_Y(y; \boldsymbol{\theta_y}) \tag{1}$$

$$s_i \sim \mathcal{G}a(s; \alpha_s, \beta_s) \tag{2}$$

$$m_i | s_i, y_i \sim \begin{cases} \mathcal{B}e_{4P}(m; s_i y_i, s_i - s_i y_1, lb, ub), & \text{if } y_i \in (lb, ub) \\ \mathcal{B}e_P(m; y_i + y_i s_i, s_i + 2), & \text{if } y_i \in (0, +\infty) \end{cases} \tag{3}$$

In both cases, the **fuzziness propagation** through Eq. (3) acts by letting $(m_1, \ldots, m_n)$ spread out near $\mathbb{E}[Y]$.

A conditional sampling schema
Proposed solution



Examples of a Beta-type 1 fuzzy number $\xi_{\widetilde{y}}$ masking the (true) uncorrupted realizations y

Inference about $\boldsymbol{\theta_y}$ involves a kind of **deblurring** procedure which uses $\tilde{\boldsymbol{y}}$ instead of the unobserved realizations $\mathbf{y}$.

Inference on θ_y

Approximated Gibbs sampling

The idea is to plug the hypothesized sampling schema into the estimation procedure, which naturally leads to a **Gibbs sampler**-based solution:

For $t > 1$ do:

$$\mathbf{y}^{(t)} \sim \pi(\mathbf{y}|\mathbf{m}, \mathbf{s}, \boldsymbol{\theta}_y^{(t-1)})$$

$$\boldsymbol{\theta}_{\boldsymbol{y}}^{(t)} \sim \pi(\boldsymbol{\theta}_{\boldsymbol{y}}|\mathbf{m}, \mathbf{s}, \boldsymbol{y}^{(t)})$$

For large T inference on $\boldsymbol{\theta}_{\boldsymbol{y}}$ can be performed by inspection of the posterior sequence $\left(\boldsymbol{\theta}_{\boldsymbol{y}}^{(1)}, \ldots, \boldsymbol{\theta}_{\boldsymbol{y}}^{(T)}\right)$.

Conditional posterior densities $\pi(\mathbf{y}|\ldots)$ and $\pi(\boldsymbol{\theta}_{\mathbf{y}}|\ldots)$ do not have known form under the proposed sampling schema. Then, hybrid solutions, such as **posterior approximation** or **Metropolis within Gibbs** could be used to solve the problem.

Conditional posterior densities $\pi(\mathbf{y}|\ldots)$ and $\pi(\boldsymbol{\theta}_{\mathbf{y}}|\ldots)$ do not have known form under the proposed sampling schema. Then, hybrid solutions, such as **posterior approximation** or **Metropolis within Gibbs** could be used to solve the problem.

Posterior sampling schema:

$\pi(\boldsymbol{\theta}_{\mathbf{y}}|\mathbf{y},\{\mathbf{m},\mathbf{s}\})$ via **MCMC** [5]

$\pi(\mathbf{y},\boldsymbol{\theta}_{\mathbf{y}}|\{\mathbf{m},\mathbf{s}\})$ $\nearrow$ $\searrow$

$\pi(\mathbf{y}|\boldsymbol{\theta}_{\mathbf{y}},\{\mathbf{m},\mathbf{s}\})$ via quadratic **posterior approximation** [4]

Conditional posterior densities $\pi(\mathbf{y}|\ldots)$ and $\pi(\boldsymbol{\theta}_\mathbf{y}|\ldots)$ do not have known form under the proposed sampling schema. Then, hybrid solutions, such as **posterior approximation** or **Metropolis within Gibbs** could be used to solve the problem.

Posterior sampling schema:

$$\pi(\boldsymbol{\theta}_\mathbf{y}|\mathbf{y},\{\mathbf{m},\mathbf{s}\}) \quad \text{via } \textbf{MCMC} \text{ [5]}$$

$$\pi(\mathbf{y},\boldsymbol{\theta}_\mathbf{y}|\{\mathbf{m},\mathbf{s}\}) \nearrow \searrow$$

$$\boxed{\pi(\mathbf{y}|\boldsymbol{\theta}_\mathbf{y},\{\mathbf{m},\mathbf{s}\})} \quad \text{via quadratic } \textbf{posterior approximation} \text{ [4]}$$

$\boxed{\textbf{Case 1}: y \in (lb, ub)}$

$$\ln \pi(y_i|\boldsymbol{\theta}_{\mathbf{y}}, \ldots) \propto -\ln \Gamma(y_i^* s_i) - \ln \Gamma(s_i - s_i y_i^*) + s_i y_i^* \ln\left(\frac{m_i - lb}{ub - m_i}\right) \;+\; \ln f_Y(y; \boldsymbol{\theta}_{\mathbf{y}})$$

$$\stackrel{\sim}{=} \ln \mathcal{B}e_{4P}(y; \lambda\sigma, \sigma - \sigma\lambda, lb, ub)$$

$y_i^* = (y_i - lb)/(ub - lb)$

$\boxed{\textbf{Case 1} : y \in (lb, ub)}$

$$\ln \pi(y_i|\boldsymbol{\theta_y},\ldots) \propto \overbrace{-\ln\Gamma(y_i^* s_i) - \ln\Gamma(s_i - s_i y_i^*) + s_i y_i^* \ln\left(\frac{m_i - lb}{ub - m_i}\right)}^{h(y;m,s,lb,ub)} + \ln f_Y(y;\boldsymbol{\theta_y})$$
$$\stackrel{\sim}{=} \ln \mathcal{B}e_{4P}(y; \lambda\sigma, \sigma - \sigma\lambda, lb, ub)$$

$\{\lambda, \sigma\} \in (lb, ub) \times \mathbb{R}^+$:

$$\frac{\partial^k}{\partial y^k} \ln \mathcal{B}eta_{4P}(y; \lambda\sigma, \sigma - \sigma\lambda, lb, ub) = \frac{\partial^k}{\partial y^k}\Big(h(y; m, s, lb, ub) + \ln f_Y(y; \boldsymbol{\theta_y})\Big)$$
$$k = 1, 2$$

$$\boxed{\textbf{Case 2}: y \in (lb, +\infty)}$$

$$\begin{aligned}\ln \pi(y_i|\boldsymbol{\theta_y},\ldots) &\propto \ln \mathrm{B}(y_i+s_i, s_i+2)^{-1} + \ln\left(\frac{m_i}{m_i+1}\right)(y_i + s_i y_i) + \ln m_i + 2\ln(1+m_i) + \\ &+ \ln f_Y(y;\boldsymbol{\theta_y}) \\ &\stackrel{\sim}{=} \ln \mathcal{B}e_P(y;\lambda+\lambda\sigma, \sigma+2)\end{aligned}$$

$\boxed{\textbf{Case 2} : y \in (lb, +\infty)}$

$$\ln \pi(y_i|\boldsymbol{\theta_y},\ldots) \propto \overbrace{\ln \mathrm{B}(y_i + s_i, s_i + 2)^{-1} + \ln\left(\frac{m_i}{m_i+1}\right)(y_i + s_i y_i) + \ln m_i + 2\ln(1+m_i)}^{g(y;m,s)} +$$

$$+ \ln f_Y(y; \boldsymbol{\theta_y})$$

$$\stackrel{\sim}{=} \ln \mathcal{B}e_P(y; \lambda + \lambda\sigma, \sigma + 2)$$

$$\{\lambda, \sigma\} \in (0, +\infty) \times \mathbb{R}^+ :$$

$$\frac{\partial^k}{\partial y^k} \ln Be_P(y; \lambda + \lambda\sigma, \sigma + 2) = \frac{\partial^k}{\partial y^k}\Big(g(y; m, s) + \ln f_Y(y; \boldsymbol{\theta_y})\Big)$$

$$k = 1, 2$$

Aim: Assessing the quadratic posterior approximation of $\pi(y|\boldsymbol{\theta}_{\boldsymbol{y}},\ldots)$ via $\mathcal{B}e_{4P}$ and $\mathcal{B}e_P$ distributions for both bounded and left-bounded cases.

Methods: The derivative-based density approximation (DA) is contrasted against the Adaptive Rejection Sampling (ARS) algorithm.

Measures:

- Total variation distance: $d_{TV} = \frac{1}{2}\int \left|\tilde{\pi}(y|\ldots) - \pi(y|\ldots)\right| dy$
- Computation time

$\boxed{\textbf{Case 1} : y \in (lb, ub)}$

$f_Y(y; \boldsymbol{\theta_y}) \stackrel{\text{def}}{=} \mathcal{LG}Norm(y; \mu, \phi)$ *Logit Normal distribution*

$lb = 0, ub = 1$

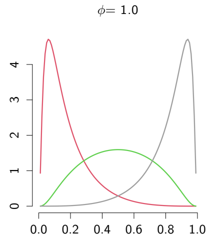
$\mu \in \{-1.85, 0, 1.85\}$

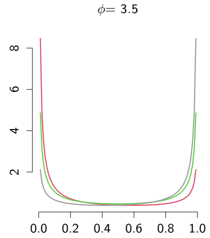
$\phi \in \{1.0, 3.5\}$

$s \sim \mathcal{G}a(s; 45.0, 45.0/\mu_s)$

$\mu_s \in \{5.0, 25.0, 50.0\}$

$n = 2000$ replicates for each combination



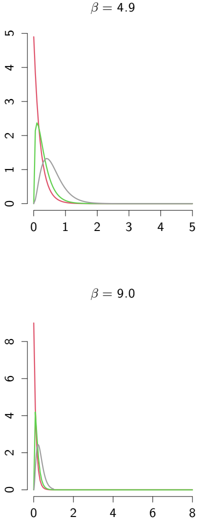


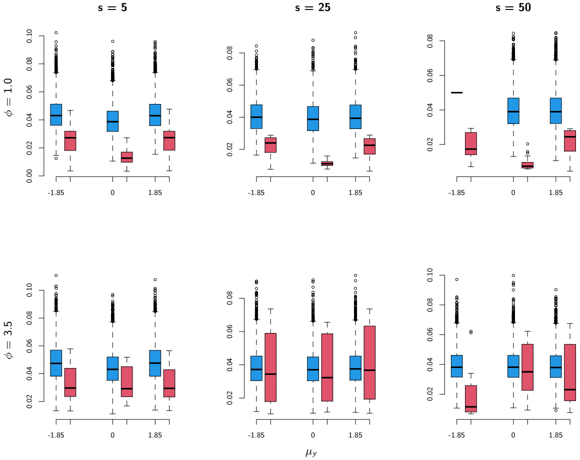
$$\boxed{\textbf{Case 2} : y \in (0, +\infty)}$$

$f_Y(y; \boldsymbol{\theta_y}) \stackrel{\text{def}}{=} \mathcal{G}a(y; \alpha, \beta)$ *Gamma distribution*
$\alpha \in \{1.0, 1.5, 3.0\}$
$\beta \in \{4.9.0, 9.0\}$

$s \sim \mathcal{G}a(s; 45.0, 45.0/\mu_s)$
$\mu_s \in \{5.0, 25.0, 50.0\}$

$n = 2000$ replicates for each combination

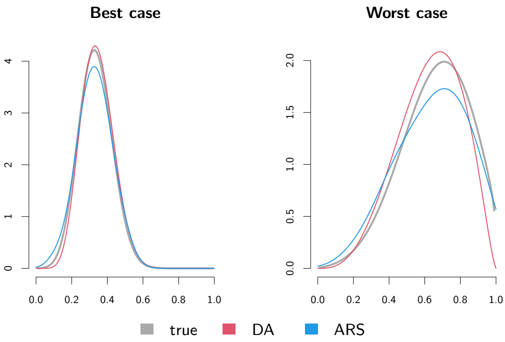




	DA	ARS
Average Accuracy	0.97	0.95
Average Log-Time	-6.97	0.67

Simulation study

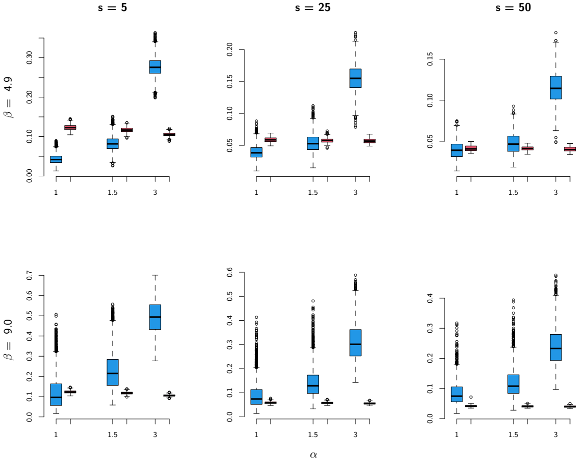
Results - Case 1



	DA	ARS
Average Accuracy	0.97	0.95
Average Log-Time	-6.97	0.67

Simulation study

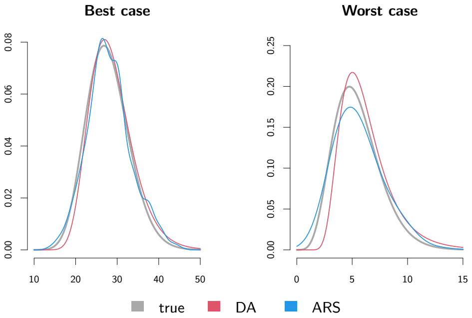
Results - Case 2



	DA	ARS
Average Accuracy	0.93	0.85
Average Log-Time	-7.27	1.15

Simulation study

Results - Case 2



	DA	ARS
Average Accuracy	0.93	0.85
Average Log-Time	-7.27	1.15

Conclusions

- A general framework for **data analysis** by taking the advantages of a **general sampling schema** for the fuzziness propagation over the outcomes of $f_Y(y; \boldsymbol{\theta_y})$
- Results are still **preliminary**: Further simulations coupling DA with MCMC are currently underway
- Generalizations to non-convex and **trapezoidal fuzzy numbers** need also to be considered

[1] Coppi, R., Gil, M. A., and Kiers, H. A.
The fuzzy approach to statistical analysis.
Computational statistics & data analysis 51, 1 (2006), 1–14.

[2] Denœux, T.
Maximum likelihood estimation from fuzzy data using the em algorithm.
Fuzzy sets and systems 183, 1 (2011), 72–91.

[3] Grzegorzewski, P., and Goławska, J.
In search of a precise estimator based on imprecise data.
In *19th World Congress of the International Fuzzy Systems Association (IFSA), 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and 11th International Summer School on Aggregation Operators (AGOP)* (2021), Atlantis Press, pp. 530–537.

[4] Miller, J. W.
Fast and accurate approximation of the full conditional for gamma shape parameters.
Journal of Computational and Graphical Statistics 28, 2 (2019), 476–480.

[5] Zhou, H., and Huang, X.
Bayesian beta regression for bounded responses with unknown supports.
Computational Statistics & Data Analysis 167 (2022), 107345.

antonio.calcagni@unipd.it