A Hierarchical Representation of Fuzziness in Fuzzy Data Analysis

Antonio Calcagnì, *University of Padova*

Przemyslaw Grzegorzewski, *Warsaw University of Technology*

EUSFLAT 2025, July 21

Introduction

Fuzzy data in action

It is widely recognized that statistical analyses benefit from using **fuzzy numbers** to handle real situations involving **post-sampling** or **epistemic uncertainty**.

This is quite evident in **social science** research, which frequently suffers from imprecise measurement [Cao et al., 2024].

Yet fuzzy data also arise in the **life sciences**, for instance in RNA-seq analyses where the read-to-gene alignment problem produces multireads (**fuzzy counts**) [Consiglio et al., 2016, Mencar and Pedrycz, 2020].

Introduction

Modeling fuzzy data

Our problem can be well-settled within the **Tanaka-Okuda approach** to fuzzy data analysis [Tanaka et al., 1977, Gebhardt et al., 1998].

Introduction

Modeling fuzzy data

Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (S, \mathcal{S})$ be a $\mathcal{A}$-$\mathcal{S}$-measurable function. The induced distribution $\mathbb{P}_X$ on $(S, \mathcal{S})$ is assumed to belong to a *parametric* family $\{\mathbb{P}_\theta : \theta \in \Theta\}$.

The sample $X_1, \ldots, X_n$ is assumed to be blurred into the **fuzzy sample**

$$\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_n),$$

with $\tilde{x}_i$ being a fuzzy subset of S characterized by a Borel-measurable membership function $\xi_{\tilde{x}_i} : S \to [0, 1]$. Here, $\tilde{\mathcal{S}}$ is a fuzzy cover of S or a **fuzzy information system** in Tanaka's sense.

The statistical problem here is to identify $\hat{\theta} \in \Theta$ such that $\mathbb{P}_{\hat{\theta}}$ describes the distribution of $\mathbf{x}$ based on $\tilde{\mathbf{x}}$. This is a type of **filtering** or **de-blurring** problem.

Fuzziness as coarsening

Is fuzziness a form of coarsening?

It has been argued that fuzziness can be interpreted as a form of coarsening, such as **grouping** [Gebhardt et al., 1998] or **interval censoring** [Nguyen and Wu, 2006, Denœux, 2011].

Moreover, a *likelihood-based interpretation* of fuzzy data has been proposed as a generalization of the assumption that data are coarsened at random (**CAR**) [Cattaneo, 2017].

Fuzziness as coarsening

Is fuzziness a form of coarsening?

It has been argued that fuzziness can be interpreted as a form of coarsening, such as **grouping** [Gebhardt et al., 1998] or **interval censoring** [Nguyen and Wu, 2006, Denœux, 2011].

Moreover, a *likelihood-based interpretation* of fuzzy data has been proposed as a generalization of the assumption that data are coarsened at random (**CAR**) [Cattaneo, 2017].

Note that CAR implies **ignorability**: the mechanism generating fuzziness can be ignored. However, under the Tanaka-Okuda conditions, we argue that **fuzziness is not ignorable**.

Fuzziness as coarsening

No ignorability for fuzziness

Argument 1 [Gill and Grünwald, 2008]

Consider a non-empty finite set S and a collection $\mathcal{S}_* \subseteq \mathcal{P}(S) \setminus \{\emptyset\}$.
A *coarsening mechanism* is a mapping $\phi : S \to \mathcal{S}_*$ such that for any realization $x \in S$ of X, we have $x \in \phi(x)$.

Here, rather than measuring x, the observers measures a coarsened version of it, the set $A \in \mathcal{S}_*$ containing x.

The coarsening mechanism is characterized by the conditional probability of observing A given x, namely $\mathbb{P}[\phi(x) = A | X = x]$.

Fuzziness as coarsening

No ignorability for fuzziness

Argument 1 [Gill and Grünwald, 2008]

In general, ϕ models a CAR mechanism iff:

$$(1)\ \mathbb{P}[\phi(x) = A \mid X = x] = \mathbb{P}[\phi(x) = A \mid X = x'], \quad \forall x, x' \in A \qquad \text{(CAR condition)}$$

$$(2)\ \sum_{A \in \mathcal{S}_*} \mathbb{P}[\phi(x) = A \mid X = x] = 1, \quad \forall x \in S \qquad \text{(Normalization)}$$

Fuzziness as coarsening

No ignorability for fuzziness

Argument 1 [Gill and Grünwald, 2008]

$\mathcal{S}_*$ supports a CAR mechanism if the system

$$\mathbf{Mz} = \mathbf{1}_n,$$

has a unique non-negative solution, with $\mathbf{M}$ being the incidence matrix associated with $\mathcal{S}_*$. This provides an *operative test* for the CAR assumption.

In this case,

$$\hat{\mathbb{P}}[\phi(x) = A_j | X \in A_j] = \hat{z}_j, \quad \text{where} \quad j \in \{1, \ldots, |\mathcal{S}_*|\}.$$

Fuzziness as coarsening

No ignorability for fuzziness

Argument 1 [Gill and Grünwald, 2008]

Now, if $\tilde{\mathcal{S}}_*$ constitutes a collection of fuzzy subsets of S (i.e., a fuzzy cover or partition), as in the Tanaka–Okuda condition, then ϕ **is no longer CAR**.

Fuzziness as coarsening

No ignorability for fuzziness

Argument 1 [Gill and Grünwald, 2008]

Now, if $\tilde{\mathcal{S}}_*$ constitutes a collection of fuzzy subsets of S (i.e., a fuzzy cover or partition), as in the Tanaka–Okuda condition, then ϕ **is no longer CAR**.

Intuitively, since $\xi_{\tilde{A}}(x)$ varies over $x \in \tilde{A}$, realizations are **no longer exchangeable** within A, unlike in the crisp case.

Argument 1 [Gill and Grünwald, 2008]

Let $\widetilde{\mathbf{M}} = (\xi_{\tilde{A}_j}(x_i))_{ij}$ denote the fuzzy incidence matrix. Then, the solutions to the associated linear system no longer satisfy x-independence (condition 1); that is,

$$\hat{\mathbb{P}}[\phi(x) = \tilde{A}_j \mid X \in \tilde{A}_j] \neq \hat{\mathbb{P}}[\phi(x) = \tilde{A}_j \mid X = x] = \xi_{\tilde{A}_j}(x)\hat{z}_j,$$
$$\text{where} \quad j \in \{1, \ldots, |\mathcal{S}_*|\}.$$

Indeed,

$$\xi_{\tilde{A}_j}(x)\hat{z}_j \neq \xi_{\tilde{A}_j}(x')\hat{z}_j,$$

unless, in the trivial case, $\widetilde{\mathbf{M}} = \mathbf{1}c$ for some $c \in [0,1]$. However, the system becomes non-identifiable in this case.

Fuzziness as coarsening

No ignorability for fuzziness

Argument 2 [Kaymak et al., 2003]

Consider a *probabilistic fuzzy system* with crisp antecedents S (equipped with a probability distribution $\mathbb{P}_\theta$) and fuzzy consequents $\tilde{\mathcal{S}}_*$.

The input-output connecting rule ϕ is evaluated via the conditional probability of $\tilde{A}_j \in \tilde{\mathcal{S}}_*$ given $x \in S$, i.e.

$$\begin{aligned}\mathbb{P}[\phi(x) = \tilde{A}_j | X = x] &= \mathbb{P}[\tilde{A}_j \cap \{x\}]\ (\mathbb{P}_\theta[X = x])^{-1} \\ &= \xi_{\tilde{A}_j}(x).\end{aligned}$$

Still, unless $\xi_{\tilde{A}_j}(x)$ is constant over $x \in \tilde{A}_j$, the coarsening probability depends on the latent realization x.

Fuzziness as coarsening

Fuzziness requires CNAR

Arguments 1 and 2 point to a coarsening mechanism that cannot be ignored (**CNAR**: Coarsening Not At Random).

Fuzziness as coarsening

Fuzziness requires CNAR

Arguments 1 and 2 point to a coarsening mechanism that cannot be ignored (**CNAR**: Coarsening Not At Random).

As in MNAR problems [Molenberghs and Verbeke, 2005], a similar factorization arises in this context:

$$\mathbb{P}_\theta(\mathbf{x}, \tilde{\mathbf{x}} \mid \ldots) = \underbrace{\mathbb{P}_\theta(\tilde{\mathbf{x}} \mid \mathbf{x}, \ldots)}_{\substack{\text{coarsening} \\ \text{mechanism}}} \underbrace{\mathbb{P}_\theta(\mathbf{x} \mid \ldots)}_{\substack{\text{measurement} \\ \text{distribution}}}.$$

Fuzziness as coarsening

Fuzziness requires CNAR

Arguments 1 and 2 point to a coarsening mechanism that cannot be ignored (**CNAR**: Coarsening Not At Random).

As in MNAR problems [Molenberghs and Verbeke, 2005], a similar factorization arises in this context:

$$\mathbb{P}_\theta(\mathbf{x}, \tilde{\mathbf{x}} \mid \ldots) = \underbrace{\mathbb{P}_\theta(\tilde{\mathbf{x}} \mid \mathbf{x}, \ldots)}_{\text{coarsening mechanism}} \underbrace{\mathbb{P}_\theta(\mathbf{x} \mid \ldots)}_{\text{measurement distribution}}.$$

▶ To specify the coarsening mechanism, we propose using a parametric **hierarchical model**.

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

To fix ideas, consider a collection of bounded Beta-type fuzzy numbers

$$\tilde{\mathbf{X}} = ((m_1, s_1), \ldots, (m_n, s_n)),$$

parametrized using mode $m \in \mathbb{R}$ and precision $s \in \mathbb{R}^+$.

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

Under the CNAR assumption, the fuzziness mechanism can be specified as follows [Calcagnì et al., 2025]:

$$\begin{aligned} f(\{\mathbf{m},\mathbf{s}\} \mid \mathbf{x},\boldsymbol{\theta}) &= f(\mathbf{m}, \mid \mathbf{s},\mathbf{x},\boldsymbol{\theta}) f(\mathbf{s} \mid \mathbf{x},\boldsymbol{\theta}) f(\mathbf{x} \mid \boldsymbol{\theta}) \\ &= f(\mathbf{m}, \mid \mathbf{s},\mathbf{x}) \underbrace{f(\mathbf{s} \mid \boldsymbol{\theta}_s) f(\mathbf{x} \mid \boldsymbol{\theta}_x)}_{S_i \perp\!\!\!\perp X_i}, \end{aligned}$$

where

$$\mathrm{Sup}(X_i) \subseteq \mathrm{Sup}(M_i),$$

$$\mathbb{E}[M_i] = \mathbb{E}[X_i],$$

$$\mathbb{V}\mathrm{ar}[M_i] = g(\mathbb{V}\mathrm{ar}[X_i], \mathbb{E}[X_i], c), \text{ where } c > 0.$$

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

A particular *instance* of hierarchical model is the following

$$\begin{aligned} x_i &\sim f_X(x;\boldsymbol{\theta}_x),\\ s_i &\sim \mathcal{G}(s;\alpha_s,\beta_s),\\ m_i|s_i,x_i &\sim \mathcal{B}e_{4P}(m;s_ix_i,s_i-s_ix_i,lb,ub), \end{aligned}$$

where $f_X(x;\boldsymbol{\theta}_x)$ is the measurement model with $g(\mathbb{E}[X_i]) = \mathbf{z}_i\boldsymbol{\beta}$ to account for external covariates.

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

To check the effects of *coarsening mispecification*, consider a simple **application** of the hierarchical model on a $n = 318$ sample of Beta-type fuzzy numbers ([Calcagnì et al., 2025], Section 6.4).

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

▷ Models specification:

CNAR	CAR
$f_{X_i}(x;\boldsymbol{\theta}) = \mathcal{B}e_{(0,1)}(x;\mu\phi,\phi-\phi\mu)$	–
$m_i \mid s_i, x_i \sim \mathcal{B}e_{4P}(m; s_i x_i, s_i - s_i x_i, 0, 1)$	$m_i \mid s_i \sim \mathcal{B}e_{4P}(m; s_i\mu, s_i - s_i\mu, 0, 1)$

where $\boldsymbol{\theta} = \{\phi, \mu\} \in \mathbb{R}_+ \times (0,1)$ in both cases.

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

▷ Parameter estimation:

MCMC with $4 \times 4e3$ samples (burn-in: $1e3$ samples)

▷ Model performance:

Posterior Predictive Checks [Gelman et al., 2008]:

$\pi(\hat{\mathbf{m}}, \hat{\mathbf{s}} \mid .., \boldsymbol{\theta})$ vs. $\{\mathbf{m}, \mathbf{s}\}$

$\pi(\sup(\hat{\tilde{\mathbf{x}}}), \mid .., \boldsymbol{\theta})$ vs. $\sup(\tilde{\mathbf{x}})$

$\pi(\text{kaufman}(\hat{\tilde{\mathbf{x}}}), \mid .., \boldsymbol{\theta})$ vs. $\text{kaufman}(\tilde{\mathbf{x}})$

Measures:

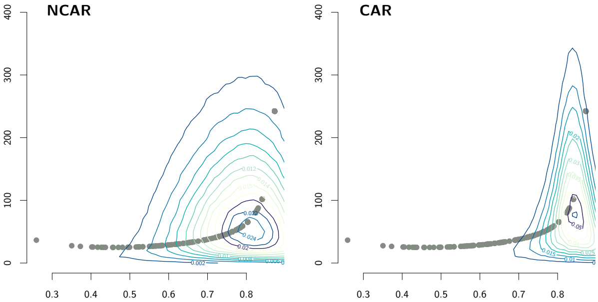
coverage (the higher, the better)

transformed Bayesian p-value (the lower, the better)

A hierarchical model for fuzziness

An application with Beta fuzzy numbers

- ▷ Results:



	CNAR	CAR
	support	
covg	0.962	0.956
bayPv	0.034	0.069
	kaufman	
covg	0.730	0.651
bayPv	0.035	0.071

Conclusions

- Fuzziness can be seen as a form of coarsening, but standard CAR assumptions imply x-independent coarsening probabilities
- In the fuzzy case, membership functions $\xi_{\tilde{A}}$ introduce x-dependence into the coarsening probabilities, violating CAR
- This implies that fuzziness needs to be treated as CNAR
- Hierarchical models can then be used to explicitly specify the CNAR mechanism

[Calcagnì et al., 2025] Calcagnì, A., Grzegorzewski, P., and Romaniuk, M. (2025).
Bayesianize fuzziness in the statistical analysis of fuzzy data.
International Journal of Approximate Reasoning, page 109495.

[Cao et al., 2024] Cao, N., Finos, L., Lombardi, L., and Calcagnì, A. (2024).
A novel CFA+EFA model to detect aberrant respondents.
Journal of the Royal Statistical Society Series C: Applied Statistics, 73(5):1283–1309.

[Cattaneo, 2017] Cattaneo, M. E. (2017).
The likelihood interpretation as the foundation of fuzzy set theory.
International Journal of Approximate Reasoning, 90:333–340.

[Consiglio et al., 2016] Consiglio, A., Mencar, C., Grillo, G., Marzano, F., Caratozzolo, M. F., and Liuni, S. (2016).
A fuzzy method for RNA-Seq differential expression analysis in presence of multireads.
BMC bioinformatics, 17:95–110.

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

[Gebhardt et al., 1998] Gebhardt, J., Gil, M. A., and Kruse, R. (1998).
Fuzzy set-theoretic methods in statistics.
In *Fuzzy sets in decision analysis, operations research and statistics*, pages 311–347. Springer.

[Gelman et al., 2008] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2008).
Bayesian data analysis (second edition).

[Gill and Grünwald, 2008] Gill, R. D. and Grünwald, P. D. (2008).
An algorithmic and a geometric characterization of coarsening at random.
The Annals of Statistics, 36(5):2409–2422.

[Kaymak et al., 2003] Kaymak, U., Van Den Bergh, W.-M., and van den Berg, J. (2003).
A fuzzy additive reasoning scheme for probabilistic mamdani fuzzy systems.
In *The 12th IEEE International Conference on Fuzzy Systems, 2003. FUZZ'03.*, volume 1, pages 331–336. IEEE.

[Mencar and Pedrycz, 2020] Mencar, C. and Pedrycz, W. (2020).
Granular counting of uncertain data.
Fuzzy Sets and Systems, 387:108–126.

[Molenberghs and Verbeke, 2005] Molenberghs, G. and Verbeke, G. (2005).
Models for discrete longitudinal data.
Springer.

[Nguyen and Wu, 2006] Nguyen, H. T. and Wu, B. (2006).
Random and fuzzy sets in coarse data analysis.
Computational statistics & data analysis, 51(1):70–85.

[Tanaka et al., 1977] Tanaka, H., Okuda, T., and Asai, K. (1977).
On decision-making in fuzzy environment fuzzy information and decision making.
The international journal of production research, 15(6):623–635.

antonio.calcagni@unipd.it
https://unipd.link/acalcagni