Fuzzy Probability Distributions in Bayesian Analysis

Reinhard Viertl and Owat Sunanta
Department of Statistics and Probability Theory
Vienna University of Technology, Vienna, Austria
Corresponding author: Reinhard Viertl, e-mail: r.viertl@tuwien.ac.at

Abstract

A-priori information in Bayesian analysis in form of precise probability distributions is a topic of critical discussion. Therefore, a more general form of a-priori distribution is necessary. Using the concept of fuzzy valued functions, the so-called fuzzy probability densities are suitable models for soft a-priori information.

Keywords: fuzzy Bayesian analysis, fuzzy probability densities, fuzzy utility, generalized integration

1. Introduction

In Bayesian analysis, precise a-priori distributions are often not available. To capture such uncertainty, a more general form of a-priori information can be expressed by using soft models. The mathematical bases for such models are fuzzy numbers and fuzzy valued functions, especially the so-called fuzzy probability densities. Based on these generalized probability densities, Bayes' theorem can be generalized. Moreover, the concepts of predictive distributions and statistical decision models can be adapted accordingly. These concepts yield more realistic approaches for capturing the uncertainty of data and a-priori information.

2. Fuzzy Probability Densities

The generalization of probability densities to soft a-priori information is possible using fuzzy numbers.

Definition 1: A fuzzy number x^* is defined by its characterizing function $\xi(\cdot)$, which is a real function of one real variable and possesses the following conditions:

(1) $\xi: \mathbb{R} \to [0;1]$

(2) The support of $\xi(\cdot)$, denoted by $\text{supp}[\xi(\cdot)]$ and defined by
$$\text{supp}[\xi(\cdot)] \coloneqq \{x \in \mathbb{R}: \xi(\cdot) > 0\},$$
is a bounded subset of $\mathbb{R}$.

(3) For all $\delta \in (0;1]$ the so-called δ-cut $C_\delta[\xi(\cdot)]$, defined by
$$C_\delta[\xi(\cdot)] \coloneqq \{x \in \mathbb{R}: \xi(\cdot) \geq \delta\} = \bigcup_{j=1}^{k_\delta} [a_{\delta,j}; b_{\delta,j}],$$
is a non-empty and finite union of compact intervals.

Remark 1: Fuzzy numbers are the best (to-date) mathematical description for capturing imprecise data. Measurements of continuous quantities are more or less imprecise, also called fuzzy. Special forms of fuzzy numbers are the fuzzy intervals.

Definition 2: A fuzzy number whose δ-cuts are compact intervals $[a_\delta; b_\delta]$ is called fuzzy interval. Functions $f^*(\cdot)$ defined on a set M, whose values $f^*(x)$ are fuzzy intervals together with their so-called δ-level functions $\underline{f}_\delta(\cdot)$ and $\overline{f}_\delta(\cdot)$, are defined in the following way:

Let $C_\delta[f^*(x)] = [a_\delta(x); b_\delta(x)]$ $\forall\delta\in(0;1]$. The lower and upper δ-level functions are classical real-valued functions defined by their values $\underline{f}_\delta(x) \coloneqq a_\delta(x)$ $\forall\, x \in M$ and $\overline{f}_\delta(x) \coloneqq b_\delta(x)$ $\forall\, x \in M$.

In generalizing probability densities to fuzzy a-priori information, the so-called fuzzy valued functions are useful. A fuzzy valued function is a function whose values are fuzzy numbers. Special fuzzy valued functions are, consequently, called fuzzy probability densities:

Definition 3: A fuzzy valued function $f^*(\cdot)$ defined on a measure space $(M, \mathcal{A}, \mu)$ whose values $f^*(x)$ are fuzzy intervals for all $x \in M$ and possesses the following conditions:

(1) $\exists g : M \to [0;\infty)$ which is a classical probability density on $(M, \mathcal{A}, \mu)$ where $\underline{f}_1(x) \leq g(x) \leq \overline{f}_1(x)$ $\forall\, x \in M$

(2) all δ-level functions $\underline{f}_\delta(\cdot)$ and $\overline{f}_\delta(\cdot)$ are integrable functions with finite integral

is called fuzzy probability density.

Remark 2: The δ-cuts $C_\delta[f^*(x)] = [a_\delta(x); b_\delta(x)]$ are defining families $\mathbb{D}_\delta$ of classical probability densities in the following way:

$\mathbb{D}_\delta \coloneqq \{h: h$ is a classical density on $(M, \mathcal{A}, \mu)$ obeying $a_\delta(x) \leq h(x) \leq b_\delta(x)\ \forall\, x \in M\}$

Based on fuzzy probability densities, generalized probabilities $P^*(A)$ of events $A \in \mathcal{A}$ are defined in the following way:

A family of compact intervals $B_\delta = [c_\delta; d_\delta]$, $\delta\in(0;1]$ is given by

$$\left.\begin{array}{l} c_\delta \coloneqq \inf\{\int_A h(x)\, d\mu(x);\, h\in \mathbb{D}_\delta\} \\ d_\delta \coloneqq \sup\{\int_A h(x)\, d\mu(x);\, h\in \mathbb{D}_\delta\} \end{array}\right. \quad \forall\delta\in(0;1].$$

By the so-called construction lemma for characterizing functions of fuzzy numbers [see Viertl (2011)], the characterizing function $\xi(\cdot)$ of $P^*(A)$ is given by

$\xi(x) = \sup\{\delta.\mathbb{1}_{[c_\delta;\, d_\delta]}(x)\colon \delta\in[0;1]\}$ $\forall x\in\mathbb{R}$, where $\mathbb{1}_{[c_\delta;\, d_\delta]}(\cdot)$ denotes the indicator function of the set $[c_\delta; d_\delta]$, and $[c_0; d_0] \coloneqq \mathbb{R}$.

3. Fuzzy Data

Real observations of continuous stochastic quantities X are not precise numbers or vectors, whereas the measurement results are more or less non-precise, also called fuzzy. The best (to-date) mathematical description of such observations is by means of fuzzy numbers $x_1^*, \ldots, x_n^*$ with corresponding characterizing functions $\xi_1(\cdot), \ldots, \xi_n(\cdot)$.

Remark 3: The fuzziness of observations x_i^* resolves the problem in standard continuous stochastic models where observed data have zero probability.

For a fuzzy observation x_i^* from a density $f(\cdot)$ the probability of x_i^* is given by

$$Prob(x_i^*) = \int_{\mathbb{R}} \xi_i(x) f(x) d\mu(x) > 0.$$

This makes it possible to define the likelihood function of fuzzy data $x_1^*, \dots, x_n^*$ in a natural way:

$$l(\theta; x_1^*, \dots, x_n^*) = \prod_{i=1}^{n} Prob(x_i^*) = \prod_{i=1}^{n} \int_{\mathbb{R}} \xi_i(x) f(x|\theta) d\mu(x) \qquad \forall \theta \in \Theta$$

4. Generalized Bayes' Theorem

The standard Bayes' theorem has to be generalized to handle fuzzy a-priori densities $\pi^*(\cdot)$ on the parameter space $\ominus$ and fuzzy data $x_1^*, \dots, x_n^*$ of parametric stochastic model $X \sim f(\cdot|\theta)$; $\theta \in \ominus$. This is possible by using the δ-level functions $\underline{\pi}_\delta(\cdot)$ and $\overline{\pi}_\delta(\cdot)$ of $\pi^*(\cdot)$ along with defining the δ-level functions $\underline{\pi}_\delta(\cdot|x_1^*, \dots, x_n^*)$ and $\overline{\pi}_\delta(\cdot|x_1^*, \dots, x_n^*)$ of the fuzzy a-posteriori density in the following way:

$$\overline{\pi}_\delta(\theta|x_1^*, \dots, x_n^*) = \frac{\overline{\pi}_\delta(\theta). l(\theta; x_1^*, \dots, x_n^*)}{\int_\ominus \frac{\underline{\pi}_\delta(\theta) + \overline{\pi}_\delta(\theta)}{2} l(\theta; x_1^*, \dots, x_n^*) d\theta}$$

and $\forall \delta \in (0;1]$

$$\underline{\pi}_\delta(\theta|x_1^*, \dots, x_n^*) = \frac{\underline{\pi}_\delta(\theta). l(\theta; x_1^*, \dots, x_n^*)}{\int_\ominus \frac{\underline{\pi}_\delta(\theta) + \overline{\pi}_\delta(\theta)}{2} l(\theta; x_1^*, \dots, x_n^*) d\theta}$$

Remark 4: The averaging $\frac{\underline{\pi}_\delta(\theta) + \overline{\pi}_\delta(\theta)}{2}$ is necessary in order to keep the sequential updating of standard Bayes' theorem.

5. Fuzzy Predictive Densities

Classical predictive densities $p(\cdot|D)$ for stochastic models $X \sim f(\cdot|\theta)$, where $\theta \in \ominus$, based on data D are defined by using the a-posteriori density $(\cdot|\theta)$ on the parameter space $\ominus$:

$$p(x|D) := \int_\ominus f(x|\theta) \pi(\theta|D)\, d\theta \quad \forall x \in M_x$$

In case of fuzzy a-posteriori density $\pi^*(\cdot|D^*)$, the integral has to be generalized. There are different possibilities for doing so. The most natural generalization seems to be the following:

Let $\mathbb{D}_\delta$ be the set of classical probability densities $h(\cdot)$ on the parameter space Θ, fulfilling $\underline{\pi}_\delta(\theta) \leq h(\theta) \leq \overline{\pi}_\delta(\theta)$ $\forall\theta\in\Theta$, where $\underline{\pi}_\delta(\cdot)$ and $\overline{\pi}_\delta(\cdot)$ are the lower and upper δ-level functions of $\pi^*(\cdot|D^*)$ respectively.

Definition 4: The fuzzy predictive density $p^*(\cdot|D^*)$ is defined by its values $p^*(x|D^*)$ which are fuzzy intervals for all $x \in M_X$. The characterizing function $\psi_x(\cdot)$ of $p^*(x|D^*)$ is obtained from the generation lemma:

Let the generating family of intervals $[c_\delta ; d_\delta]$ be defined by

$$c_\delta := \inf\{\int_\Theta f(x|\theta)h(\theta)d\theta : h\in\mathbb{D}_\delta\}$$
$$\forall\delta\in(0;1].$$
$$d_\delta := \sup\{\int_\Theta f(x|\theta)h(\theta)d\theta : h\in\mathbb{D}_\delta\}$$

The characterizing function $\psi_x(\cdot)$ of $p^*(x|D^*)$ is obtained from

$$\psi_x(y) = \sup\{\delta.\mathbb{1}_{[c_\delta ; d_\delta]}(y) : \delta\in[0;1]\} \quad \forall y\in\mathbb{R}, \text{ where } [c_0 ; d_0] := \mathbb{R}.$$

Remark 5: The fuzzy valued function $p^*(\cdot|D^*)$ is a fuzzy probability density on the space M_X of possible values of X.

6. Decisions based on Fuzzy Information

Classical statistical decision theory is based on classical probability distributions on the state space Θ and on real valued utility functions $U(\cdot,\cdot)$. Both assumptions are not always justified, especially when utility values (θ, d) for decisions d are uncertain. Therefore a more realistic approach is to consider fuzzy valued utility functions $U^*(\cdot,\cdot)$. In this situation the concept of expected utility $\mathbb{E}_P U(\tilde{\theta}, d)$ needs to be generalized. Here, P denotes the probability distribution on the state space Θ. For continuous state space Θ, the expected utility of a decision d in the classical framework is given by

$$\mathbb{E}_P U(\tilde{\theta}, d) = \int_\Theta U(\theta, d)\, dP(\theta).$$

In case of fuzzy utility functions $U^*(\cdot,\cdot)$ whose values are fuzzy intervals and fuzzy probability density $p^*(\cdot)$ on the state space, the integral has to be generalized as:

$$\mathbb{E}_{P^*} U^*(\tilde{\theta}, d) = \int_\Theta U^*(\theta, d)\, p^*(\theta) d\theta$$

This generalized integral is defined in the following way:

Based on the family $\mathbb{D}_\delta$ from section 5, the generating family of intervals $[c_\delta ; d_\delta]$ for the fuzzy interval $\mathbb{E}_{P^*} U^*(\tilde{\theta}, d)$ is defined by

$$c_\delta := \inf\{\int_\Theta \underline{U}_\delta(\theta, d)h(\theta)d\theta : h\in\mathbb{D}_\delta\}$$
$$\forall\delta\in(0;1].$$
$$d_\delta := \sup\{\int_\Theta \overline{U}_\delta(\theta, d)h(\theta)d\theta : h\in\mathbb{D}_\delta\}$$

The characterizing function of the generalized (fuzzy) expected utility of the decision d is obtained from the construction lemma.

7. Final Remark

In classical decision analysis, the expected utilities are real numbers. In this case, it is rather simple to compare the expected utilities for different decisions. On the other hand, in a more realistic situation with fuzzy utilities, it is quite difficult to compare such expected utilities. One possibility is to use defuzzification in order to compare the expected utilities. In realistic situations, it is often very complicated to rank different decisions. This is a challenging field of further research.

References

Viertl, R. (2011) *Statistical Methods for Fuzzy Data*, Wiley, Chichester.

Viertl, R., Yeganeh, S.M. (2013) "Fuzzy Probability Distributions in Reliability Analysis, Fuzzy HPD-regions, and Fuzzy Predictive Distributions," in *Towards Advanced Data Analysis by Combining Soft Computing and Statistics*, C. Borgelt et al. (Eds.), Springer, Berlin.