Fuzzy probability distributions in Bayesian reliability analysis

R. Viertl and S.M. Yeganeh

Forschungsbericht SM-2012-1
Jänner 2012

Kontakt: P.Filzmoser@tuwien.ac.at
Abstract In reliability analysis there are different kinds of uncertainty present: Variability, imprecision of lifetimes, model uncertainty concerning probability distributions, and uncertainty of a-priori information in Bayesian analysis. For the description of imprecise lifetimes so-called fuzzy numbers are suitable. In order to model the uncertainty of a-priori information fuzzy probability distributions are the most up to date mathematical structure.

1 Introduction

The variability of lifetimes $T$ of similar units is described since a long time by probability distributions. More recently imprecise lifetime data are modeled by so-called fuzzy numbers (compare [1] and [2]). These fuzzy numbers are special fuzzy subsets of the set of real numbers $\mathbb{R}$ whose membership functions $\xi(\cdot)$ are obeying the following:

1. $\xi : \mathbb{R} \rightarrow [0; 1]$.
2. $\forall \delta \in (0; 1]$ the so-called $\delta$-cut $C_\delta[\xi(\cdot)]$ defined by $C_\delta[\xi(\cdot)] := \{x \in \mathbb{R} : \xi(x) \geq \delta\} \neq \emptyset$, and all $\delta$-cuts are finite unions of compact intervals.
3. $\text{supp}[\xi(\cdot)]$ is contained in a compact interval.

Functions $\xi(\cdot)$ fulfilling 1. - 3. are called characterizing functions.

If all $\delta$-cuts of a fuzzy number are compact intervals, then this fuzzy number is called fuzzy interval. The set of all fuzzy intervals is denoted by $\mathcal{F}_I(\mathbb{R})$.

A generalization of probability densities which is necessary in connection with Bayesian inference for fuzzy data, are so-called fuzzy probability densities on measurable spaces $(\mathcal{M}, \mathcal{A}, \mu)$. 

Department of Statistics and Probability Theory, Vienna University of Technology, A- 1040 Vienna, Austria r.viertl@tuwien.ac.at · shohreh.mir@gmail.com
A fuzzy probability density \( f^* (\cdot) \) is a function \( f^* : \mathcal{M} \rightarrow \mathcal{F}_1([0; \infty)), \) i.e. a function whose values \( f^*(x) \) are fuzzy intervals whose supports are subsets of the non-negative numbers \([0; \infty)\) for which all so-called \( \delta \)-level functions \( \underline{f}_\delta (\cdot) \) and \( \bar{f}_\delta (\cdot) \) are integrable. These \( \delta \)-level functions are defined by their values \( \underline{f}_\delta (x) \) and \( \bar{f}_\delta (x) \) by \( C_\delta [f^*(x)] = [\underline{f}_\delta (x); \bar{f}_\delta (x)] \) for all \( \delta \in (0; 1] \) and all \( x \in \mathcal{M} \). This means all integrals

\[
\int_{\mathcal{M}} \underline{f}_\delta (x) d\mu(x) \quad \text{and} \quad \int_{\mathcal{M}} \bar{f}_\delta (x) d\mu(x)
\]

exist and are finite. Based on fuzzy probability densities so-called fuzzy probabilities of events \( A \in \mathcal{A} \) are determined in the following way:

The definition of fuzzy probabilities is based on a generating family of subsets of \( \mathbb{R} \) to define a fuzzy interval via the so-called generation lemma for characterizing functions (compare [3]). The generating intervals \([a_\delta; b_\delta] \) are defined using families of classical probability densities \( f(\cdot) \) on \((\mathcal{M}, \mathcal{A})\):

\[
\mathcal{D}_\delta := \{ f : \underline{f}_\delta (x) \leq f(x) \leq \bar{f}_\delta (x) \quad \forall x \in \mathcal{M} \}
\]

then \( a_\delta \) and \( b_\delta \) are defined by

\[
a_\delta := \inf \{ \int_{\mathcal{A}} f(x) d\mu(x) : f \in \mathcal{D}_\delta \}
\]

and

\[
b_\delta := \sup \{ \int_{\mathcal{A}} f(x) d\mu(x) : f \in \mathcal{D}_\delta \} \quad \forall \delta \in (0; 1]
\]

The fuzzy probability \( p^* (A) \) is the fuzzy interval whose characterizing function \( \eta(\cdot) \) is given by

\[
\eta(x) := \sup \{ \delta \cdot 1_{[a_\delta; b_\delta]}(x) : \delta \in [0; 1] \} \quad \forall x \in \mathbb{R},
\]

where \( 1_{[a_\delta; b_\delta]}(\cdot) \) is the indicator function of the interval \([a_\delta; b_\delta]\), and \([a_0; b_0] = \mathbb{R} \).

2 Fuzzy lifetimes

In applied reliability analysis observed lifetimes as observations of time which is a continuous quantity are more or less fuzzy. Therefore a sample consists of \( n \) fuzzy numbers \( t^*_1, \ldots, t^*_n \). The corresponding characterizing functions are denoted as \( \xi_1(\cdot), \ldots, \xi_n(\cdot) \). Based on this kind of samples the reliability function \( R(\cdot) \) can be estimated by a generalization of the empirical reliability function (ERF) \( \hat{R}_n(\cdot) \). For precise lifetimes \( t^*_1, \ldots, t^*_n \) the ERF \( \hat{R}_n(\cdot) \) is defined by its values
\[ \hat{R}_n(t) := \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{(t;\infty)}(t_i) \quad \forall t \geq 0. \quad (5) \]

In case of fuzzy lifetimes there are two possibilities to generalize \( \hat{R}_n(\cdot) \). First a smoothed version of the ERF is obtained by

\[ \hat{R}_{smn}(t) := \frac{1}{n} \sum_{i=1}^{n} \frac{\int_{t}^{\infty} \xi_i(x) dx}{\int_{0}^{\infty} \xi_i(x) dx} \quad \forall t \geq 0. \quad (6) \]

In figure 1 the characterizing functions of 4 fuzzy lifetimes and the corresponding function \( \hat{R}_{smn}(\cdot) \) are depicted. Alternatively a fuzzy valued gen-

![Fig. 1 Fuzzy lifetimes and smoothed empirical reliability function](image-url)
eralization of the ERF, called fuzzy empirical reliability function (FERF), is obtained in the following way: Let the following functions $\hat{R}_{\delta,L}(\cdot)$ and $\hat{R}_{\delta,U}(\cdot)$ be defined for all $\delta \in (0; 1]$ by

$$\begin{align*}
\hat{R}_{\delta,L} := & \frac{\# \{ t^*_i : C_\delta(t^*_i) \cap (t; \infty) \neq \emptyset \}}{n} \quad \forall t \geq 0, \\
\hat{R}_{\delta,U} := & \frac{\# \{ t^*_i : C_\delta(t^*_i) \subseteq (t; \infty) \}}{n} \quad \forall t \geq 0.
\end{align*}$$

From this definition the above functions are step functions fulfilling

$$\hat{R}_{\delta,L}(t) \leq \hat{R}_{\delta,U}(t) \quad \forall t \geq 0. \quad (7)$$

Moreover $\hat{R}_{\delta,L}(0) = 1$ and $\hat{R}_{\delta,U}(0) = 1 \quad \forall \delta \in (0; 1]$ as well as

$$\lim_{t \to \infty} \hat{R}_{\delta,L}(t) = \lim_{t \to \infty} \hat{R}_{\delta,U}(t) = 0 \quad \forall \delta \in (0; 1]. \quad (9)$$

In figure 2 the FERF for the lifetime data in figure 1 is depicted. For $\delta_1 < \delta_2$ the following holds true:

$$\begin{align*}
\hat{R}_{\delta_1,U}(t) & \geq \hat{R}_{\delta_2,U}(t) \quad \forall t \geq 0 \quad (10) \\
\hat{R}_{\delta_1,L}(t) & \leq \hat{R}_{\delta_2,L}(t) \quad \forall t \geq 0 \quad (11)
\end{align*}$$

3 Bayesian reliability analysis

For parametric lifetime models $T \sim f(\cdot | \theta); \theta \in \Theta$ in Bayesian analysis also the parameter $\theta$ is described by a stochastic quantity $\tilde{\theta}$, whose probability distribution - before data are given - is called a-priori distribution. In case of continuous parameter space $\Theta$ the a-priori distribution is usually given by a probability density $\pi(\cdot)$ on $\Theta$, called a-priori density. In case of precise data $t_1, \ldots, t_n$ the updating of the information concerning the distribution of the parameter is the so-called Bayes' theorem, i.e.

$$\pi(\theta | t_1, \ldots, t_n) = \frac{\pi(\theta) \cdot l(\theta; t_1, \ldots, t_n)}{\int_\Theta \pi(\theta) \cdot l(\theta; t_1, \ldots, t_n) d\theta} \quad \forall \theta \in \Theta. \quad (12)$$

The conditional density $\pi(\cdot | t_1, \ldots, t_n)$ is called a-posteriori density of $\tilde{\theta}$. Based on the a-posteriori density Bayesian confidence regions, especially HPD-regions, as well as predictive distributions for lifetimes can be obtained. For fuzzy observed lifetimes $t^*_1, \ldots, t^*_n$ as described in section 1. Bayes theorem can be generalized in the following way:

For continuous stochastic models $X \sim f(\cdot | \theta); \theta \in \Theta$ with continuous pa-
rameter space $\Theta$ in general a-priori distributions as well as observations are fuzzy. Therefore it is necessary to generalize Bayes’ theorem to this situation.

### 3.1 Likelihood function for fuzzy data

In case of fuzzy data $t_1^*, \ldots, t_n^*$ the likelihood function $l(\theta; t_1, \ldots, t_n)$ has to be generalized to the situation of fuzzy numbers $t_1^*, \ldots, t_n^*$. The basis for that is the combined fuzzy sample element $t^*$ (see [3]). Then the generalized likelihood function $l^*(\theta; t^*)$ is represented by its $\delta$-level functions $\hat{l}_\delta(\cdot; t^*)$ and

![Diagram showing fuzzy sample and FERF for $\delta = 0.3$.]
For the \( \delta \)-cuts of the fuzzy value \( l^*(\theta; t^*) \) we have
\[
C_\delta(l^*(\theta; t^*)) = [l_\delta(\theta; t^*), l_\delta(\theta; t^*)].
\] (13)

Using this and the construction from [3] in order to keep the sequential property of the updating procedure in Bayes’ theorem, the generalization of Bayes’ theorem to the situation of fuzzy a-priori distribution and fuzzy data is possible.

Remark 1. The generalized likelihood function \( l^*(\cdot; t^*) \) is a fuzzy valued function, i.e. \( l^*: \Theta \rightarrow F_I([0, \infty)) \).

3.2 Bayes’ theorem for fuzzy a-priori distribution and fuzzy data

Using the averaging procedure of \( \delta \)-level curves of the a-priori density and combining it with the generalized likelihood function from section 3.1, the generalization of Bayes’ theorem is possible. The construction is based on \( \delta \)-level functions.

Based on a fuzzy a-priori density \( \pi^*(\cdot) \) on \( \Theta \) with \( \delta \)-level functions \( \pi^\delta(\cdot) \), and a fuzzy sample \( t_1^*, \ldots, t_n^* \) with combined fuzzy sample \( t^* \) whose vector-characterizing function is \( \zeta(\cdot, \ldots, \cdot) \), the characterizing function \( \psi_{l^*(\cdot; t^*)}(\cdot) \) of \( l^*(\theta; t^*) \) is obtained by the extension principle, i.e.
\[
\psi_{l^*(\theta; t^*)}(y) = \begin{cases} 
\sup \{ \zeta(t) : l(\theta; t) = y \} & \text{if } \exists t : l(\theta; t) = y \\
0 & \text{if } \nexists t : l(\theta; t) = y 
\end{cases} \quad \forall y \in \mathbb{R}. \tag{14}
\]

The \( \delta \)-level curves of the fuzzy a-posteriori density \( \pi^*(\cdot|t_1^*, \ldots, t_n^*) = \pi^*(\cdot|t^*) \) are defined in the following way:
\[
\pi^\delta(\theta|t^*) := \frac{\pi^\delta(\theta) \cdot I_\delta(\theta; t^*)}{\int_\Theta \frac{1}{2} [\pi^\delta(\theta) \cdot I_\delta(\theta; t^*) + \pi^\delta(\theta) \cdot I_\delta(\theta; t^*)] d\theta} \tag{15}
\]
and
\[
\pi^\delta(\theta|t^*) := \frac{\pi^\delta(\theta) \cdot I_\delta(\theta; t^*)}{\int_\Theta \frac{1}{2} [\pi^\delta(\theta) \cdot I_\delta(\theta; t^*) + \pi^\delta(\theta) \cdot I_\delta(\theta; t^*)] d\theta} \tag{16}
\]
for all \( \delta \in (0; 1] \).
3.3 Generalized fuzzy HPD-regions

From the a-posteriori density a generalization of confidence regions, especially highest a-posteriori density regions (HPD-regions) can be constructed. Let \( \pi^*(\cdot|t^*_1, \ldots, t^*_n) \) be the fuzzy a-posteriori density of \( \theta \), and \( \Theta \subseteq \mathbb{R}^k, \delta \in (0; 1], \alpha \in (0; 1), \alpha \ll 1 \) and \( 1 - \alpha \) the coverage probability. Moreover defining \( D_\delta \) to be the set of classical probability densities \( g \) on \( \Theta \) for which \( \pi^*_\delta(\theta) \leq g(\theta) \leq \pi_\delta(\theta) \) \( \forall \theta \in \Theta \), we define the generating system of subsets of \( \Theta \) from which the generalized HPD-region, denoted as HPD*-region is obtained.

For \( g \in D_\delta \) let \( \delta \text{HPD}_{1-\alpha}(g) \) be the standard HPD-region for \( \theta \) with coverage probability \( 1 - \alpha \). Then the family of generating subsets of \( \Theta \), denoted by \( (A_\delta; \delta \in (0; 1]) \), is defined by

\[
A_\delta := \bigcup_{g \in D_\delta} \delta \text{HPD}_{1-\alpha}(g) \quad \forall \delta \in (0; 1]. \tag{17}
\]

The membership function \( \varphi(\cdot) \) of the HPD*-region is given by the so-called construction lemma, i.e.

\[
\varphi(\theta) := \sup\{ \delta \cdot 1_{A_\delta}(\theta) : \delta \in [0; 1]\} \quad \forall \theta \in \Theta. \tag{18}
\]

Remark 2. In case of classical a-posteriori density \( \pi(\cdot|t_1, \ldots, t_n) \), the membership function \( \varphi(\cdot) \) coincides with the indicator function \( 1_{\text{HPD}_{1-\alpha}}(\cdot) \) of the classical HPD-region. This is seen by \( D_\delta = \{ \pi(\cdot) \} \forall \delta \in (0; 1] \) and therefore \( \delta \text{HPD}_{1-\alpha} = \text{HPD}_{1-\alpha} \forall \delta \in (0; 1] \) which yields

\[
\bigcup_{g \in D_\delta} \delta \text{HPD}_{1-\alpha}(g) = \text{HPD}_{1-\alpha} \quad \forall \delta \in (0; 1]. \tag{19}
\]

Therefore \( A_\delta = \text{HPD}_{1-\alpha} \forall \delta \in (0; 1], \) and \( \varphi(\cdot) = 1_{\text{HPD}_{1-\alpha}}(\cdot). \)

3.4 Fuzzy predictive densities

Another application of fuzzy a-posteriori densities is the construction of generalized predictive densities \( p(\cdot|t_1^*, \ldots, t_n^*) \) for lifetimes. In the classical case the predictive density is defined as the marginal density of the joint density of \( (\theta, T) \), i.e.

\[
p(t|t_1, \ldots, t_n) = \int_{\Theta} f(t|\theta)\pi(\theta|t_1, \ldots, t_n)d\theta \quad \forall t \geq 0. \tag{20}
\]

In case of fuzzy a-posteriori densities \( \pi^*(\cdot|t_1^*, \ldots, t_n^*) \) the above integral has to be generalized. This can be done in different ways (compare [4]). The most suitable generalization seems to be the following: Again we look at \( D_\delta \) from
above and define for every $\delta \in (0; 1]$ the closed interval $[a_\delta; b_\delta]$ by

\[
b_\delta := \sup \left\{ \int_\Theta f(x|\theta)g(\theta)d\theta : g \in D_\delta \right\} \\
a_\delta := \inf \left\{ \int_\Theta f(x|\theta)g(\theta)d\theta : g \in D_\delta \right\}.
\] (21) (22)

The characterizing function $\psi_t(\cdot)$ of the value $p^*(t|t^*_1,\ldots,t^*_n) \forall t \geq 0$ of the generalized fuzzy predictive density $p^*(\cdot|t^*_1,\ldots,t^*_n)$ is defined by the construction lemma:

\[
\psi_t(y) := \sup_{\delta \in [0; 1]} \delta \cdot 1_{[a_\delta; b_\delta]}(y) \quad \forall y \in \mathbb{R}
\] (23)

Remark 3. For precise a-posteriori density the result coincides with the result from standard Bayesian inference.

References