On Fuzzy Bayesian Inference

Frühwirth-Schnatter, Sylvia

DOI:
10.57938/98506827-d124-4ed2-98d1-ef31d87ccf41

Published: 01/01/1990

Document Version:
Publisher's PDF, also known as Version of record

On Fuzzy Bayesian Inference

Sylvia Frühwirth-Schnatter

Institut für Statistik
Wirtschaftsuniversität Wien

Forschungsberichte

Bericht 17
1990

http://statmath.wu-wien.ac.at/

1 Introduction

The term "Fuzzy Bayesian Inference" was introduced by [7] meaning the generalization of Bayesian Statistics to fuzzy data. What is meant by fuzziness of data? Fuzziness of data occurs when it is impossible to assign a precise number to the observations in a sample. In Bayesian statistics fuzziness of data is not the only source of fuzziness. Fuzziness may enter through the prior when the prior's parameters are fuzzy, too. This fuzziness may occur in case it is not possible to formulate precise parameters. In this paper the generalization of Bayesian methods is discussed for both fuzzy data and fuzzy prior parameters.

Statistical inference for fuzzy data is a rather new issue. [3] developped a sophisticated mathematical framework to analyze fuzzy random variables. As a by-product they generalize some concepts from descriptive statistics to fuzzy data. A broader field of statistics for fuzzy data is discussed in a couple of papers by Viertl ([7], [8], [9], [10], [11]). However, these papers are not based on a consistent theory of statistical inference for fuzzy data.

Carrying on some ideas presented in the literature mentioned above, such a general mathematical framework of statistical inference for fuzzy data has been outlined in [6]. This theory is based on the principle of propagation of fuzziness ([5]) which is an application of the extension principle introduced in fuzzy set theory by Zadeh. The basic method is to regard the result of statistical inference as the fuzzy image of a fuzzy argument under a non-fuzzy mapping. The main ideas and results of this approach will be reviewed in a rather informal manner in Section 2.

The ideas of this approach are very general and may be widely applied in statistical research, ranging from simple descriptive statistics for fuzzy data ([5], [6]) to more complex statistical methods such as Kalman-filtering for fuzzy data ([4]). In the paper at hand we apply it to Bayesian statistics to obtain "Fuzzy Bayesian Inference". In the subsequent sections we will discuss a fuzzy valued likelihood function, Bayes' theorem for both fuzzy data and fuzzy priors, a fuzzy Bayes' estimator, fuzzy predictive densities and distributions, and fuzzy H.P.D.-Regions.

2 Propagation of Fuzziness in Statistics

2.1 Fuzzy Data and Fuzzy Samples

A fuzzy data point is a X-valued observation, $X \subseteq \mathbb{R}$, which cannot be described by a precise number. The imprecision of such a data point is modelled by a fuzzy number x^* with characterizing function $\varphi_{x^*} : X \mapsto [0,1]$ ([3] and [7]). A precise data point x_i may be viewed as a "fuzzy" observation with the characterizing function equal to the indicator function $\mathrm{I}_{\{x_i\}}(\cdot)$.

An equivalent description of a fuzzy number is given by a family $(C(x^*)_\alpha)_{\alpha\in(0,1]}$ of subsets of X called α-cuts. For fixed α the set $C(x^*)_\alpha$ contains all values in X with the characterizing function greater or equal to α:

$$C(x^*)_\alpha = \{x \in X : \varphi_{x^*}(x) \geq \alpha\}, \qquad \forall\alpha \in (0,1]. \tag{1}$$

Under certain conditions on the function $\varphi_{x_i^*}(\cdot)$ (see [6]) these α-cuts are closed, bounded, non-empty intervals $[\underline{C}(x_i^*)_\alpha, \overline{C}(x_i^*)_\alpha]$ for all $\alpha \in (0,1]$.

A fuzzy sample $\boldsymbol{x}^*$ contains n data points $x_1{}^*, \ldots, x_n{}^*$ where at least one of them really is a fuzzy number. Each of these data points is characterized by the function $\varphi_{x_i^*}(\cdot)$. Now we have to define the characterizing function $\varphi_{\boldsymbol{x}^*}(\cdot)$ of the whole sample $\boldsymbol{x}^*$ by combining the n characterizing functions $\varphi_{x_i^*}(\cdot)$ of the individual data points. Without further assumptions such a combining rule is not unique. In the paper at hand the following combining rule is considered ("minimum rule fuzzy sample", [6]):

$$\varphi_{\boldsymbol{x}^*}(\boldsymbol{x}) = \min_{i=1,\ldots,n} \varphi_{x_i^*}(x_i), \qquad \forall\, \boldsymbol{x} = (x_1, \ldots, x_n) \in X^n. \tag{2}$$

Again, an equivalent description of a fuzzy sample is given by the α-cut representation $(C(\boldsymbol{x}^*)_\alpha)_{\alpha\in(0,1]}$:

$$C(\boldsymbol{x}^*)_\alpha = \{\boldsymbol{x} \in X^n : \varphi_{\boldsymbol{x}^*}(\boldsymbol{x}) \geq \alpha\}, \qquad \forall\alpha \in (0,1]. \tag{3}$$

In [6] is has been shown that the shape of the α-cuts $C(\boldsymbol{x}^*)_\alpha$ determines the combining rule. For example, for the minimum rule fuzzy sample (2) each α-cut $C(\boldsymbol{x}^*)_\alpha$ is given by the cartesian product of the α-cuts $C(x_i^*)_\alpha$ of the fuzzy data points $x_i{}^*$.

2.2 Propagation of Fuzziness

In statistics (precise) data $\boldsymbol{x} = (x_1, \ldots, x_n)$ are assumed to be realizations of a stochastic quantity or a stochastic process the probability distribution of which is not exactly known. Statistical inference then aims at some "tool of inference" based on the data. Such a "tool of inference" is a certain point denoted by T in some space Y, e.g. a number in $Y = X$ in case of the sample mean of univariate data, a function in the space $Y = \mathcal{D}_n$ of all empirical distribution functions with at most n jumps, or one set out of the space $Y = \mathcal{A}$ of all subsets of a parameter space in case of confidence regions.

For each sample $\boldsymbol{x}$ in the sample space X^n a point T in Y is uniquely defined. Reversely, in general there will be many samples in X^n with the same image T – consider e.g. the sample mean. Thus statistical inference corresponds to a mapping say $S(\cdot)$ from the sample space X^n to Y which usually is not one-to-one. The set of samples with the same image T will be denoted by X_T:

$$X_T = \{\boldsymbol{x} \in X^n : S(\boldsymbol{x}) = T\}. \tag{4}$$

Now we proceed with statistical inference for fuzzy samples $\boldsymbol{x}^*$ as discussed in [6]. The basic idea of this approach is to regard the result of statistical inference denoted by T^* as the fuzzy image of the fuzzy sample $\boldsymbol{x}^*$ under the mapping $S(\cdot)$. Propagation of fuzziness is carried out by a standard method of fuzzy set theory, namely through the socalled extension principle ([2]), by mapping the imprecision of the data from the sample space to the image space:

$$\varphi_{T^*}(T) = \begin{cases} \sup\limits_{\boldsymbol{x} \in X_T} \varphi_{\boldsymbol{x}^*}(\boldsymbol{x}), & X_T \neq \emptyset, \\ 0 & \text{otherwise}, \end{cases} \qquad \forall T \in Y. \tag{5}$$

This principle determines the characterizing function $\varphi_{T^*}(\cdot)$ of the fuzzy image T^* directly from the characterizing function $\varphi_{\boldsymbol{x}^*}(\cdot)$ of the fuzzy sample $\boldsymbol{x}^*$. However, the extension principle has hardly any practical relevance as it does not provide a construction principle of the characterizing function of the image. Such a construction principle may be based on the α-cut representation.

As mentioned above fuzziness of data and samples may be described in terms of the α-cut representation $(C(\boldsymbol{x}^*)_\alpha, \alpha \in (0,1])$. The same holds for the fuzzy image:

$$C(T^*)_\alpha = \{T \in Y : \varphi_{T^*}(T) \geq \alpha\}, \qquad \forall \alpha \in (0,1]. \tag{6}$$

It is much easier to determine the α-cut representation $C(T^*)_\alpha$ of the result $T^* = S(\boldsymbol{x}^*)$ of statistical inference rather than to determine the characterizing function $\varphi_{T^*}(\cdot)$ directly by the extension principle (5). It may be easily proved (see [6]) that the α-cut representation $C(T^*)_\alpha$ of the image of a fuzzy sample under a mapping simply is the image of the α-cut representation of the sample:

$$C(T^*)_\alpha = \bigcup_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} S(\boldsymbol{x}), \qquad \forall \alpha \in (0,1]. \tag{7}$$

Therefore propagation of fuzziness is performed in practice by determining the image of $C(\boldsymbol{x}^*)_\alpha$ under the mapping $S(\cdot)$ for each α. In practice the image $C(T^*)_\alpha$ is determined for a finite number of α-levels $\alpha_1, \ldots, \alpha_M$.

Subsequently, we will study propagation of fuzziness on the α-cuts, only. However, it should be kept in mind that propagation of fuzziness through (7) and through the extension principle (5) lead to the same result: the α-cuts $C(T^*)_\alpha$ given by (7) are identical with the α-cuts obtained from the characterizing function $\varphi_{T^*}(\cdot)$ given by (5). Propagation through the α-cuts is preferred for computational convenience.

2.3 Statistical Inference for Fuzzy Data

If statistical inference for precise data leads to a "tool" T, statistical inference for fuzzy data leads to a "fuzzified tool" T^*. Let us explain this concept in more details.

If the result of statistical inference is a real number $T = S(\boldsymbol{x})$ (e.g. the sample mean or a Bayes' estimator) statistical inference for fuzzy data will result in a fuzzy number T^*. Let us consider propagation of fuzziness on the α-cuts. Assume that $S(\boldsymbol{x})$ is a real valued function continuous in $\boldsymbol{x}$. The α-cut representation $(C(T^*)_\alpha, \alpha \in (0,1])$ of the fuzzy image T^* is then the family of the following intervals (this result is obvious from the fact that $C(\boldsymbol{x}^*)_\alpha$ is a compact subset of the sample space):

$$C(T^*)_\alpha = \left[\min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} S(\boldsymbol{x}), \max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} S(\boldsymbol{x}) \right], \quad \forall \alpha \in (0,1]. \tag{8}$$

From a methodological point of view propagation of fuzziness is carried out on each α-level by minimization and maximization of the continuous function $S(\boldsymbol{x})$ on a compact subset of $\mathbb{R}^n$. For a minimum rule fuzzy sample these subsets are n-dimensional intervals leading to an optimization problem with linear constraints.

Now consider the case where statistical inference results in an interval $T = S(\boldsymbol{x}) = [a_1(\boldsymbol{x}), a_2(\boldsymbol{x})]$ where the boundaries depend on the sample $\boldsymbol{x}$ (e.g. a confidence region or a H.P.D. region for the unknown parameter of a distribution). For a fuzzy sample we obtain a fuzzy interval T^* in the parameter space with the following α-cut representation:

$$C(T^*)_\alpha = \left[\min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} a_1(\boldsymbol{x}), \max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} a_2(\boldsymbol{x}) \right], \quad \forall \alpha \in (0,1]. \tag{9}$$

Finally, consider the case where statistical inference for a precise sample $\boldsymbol{x}$ leads to a function $f(\omega|\boldsymbol{x})$, $f : \Omega \mapsto \mathbb{R}$, $\Omega \subseteq \mathbb{R}^q$ (e.g. the empirical distribution function of a univariate sample or the posterior distribution of a parameter in Bayesian statistics). For a fixed sample $\boldsymbol{x}$ the function $f(\omega|\boldsymbol{x})$ is a relation in $\Omega \times \mathbb{R}$ characterized by the following subset $T = S(\boldsymbol{x})$ in $\Omega \times \mathbb{R}$:

$$T = S(\boldsymbol{x}) = \bigcup_{\omega \in \Omega} \{\omega, f(\omega|\boldsymbol{x})\}.$$

For a fuzzy sample $\boldsymbol{x}^*$ statistical inference leads to fuzzy subset T^* in $\Omega \times \mathbb{R}$. Propagating fuzziness through (7) for each $\alpha \in (0,1]$ we find that the α-cut $C(T^*)_\alpha$ is given by the following relation in $\Omega \times \mathbb{R}$:

$$C(T^*)_\alpha = \bigcup_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} \bigcup_{\omega \in \Omega} \{\omega, f(\omega|\boldsymbol{x})\} = \bigcup_{\omega \in \Omega} \{\omega, C(f(\omega|\boldsymbol{x}^*))_\alpha\} \tag{10}$$

with

$$C(f(\omega|\boldsymbol{x}^*))_\alpha = \bigcup_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} f(\omega|\boldsymbol{x}). \tag{11}$$

Each ω is related with all functional values in the set $C(f(\omega|\boldsymbol{x}^*))_\alpha$. This set is a point for all $\alpha \in (0,1]$ iff the sample is precise. Therefore, if at least one data point is fuzzy at least one of the relations $C(T^*)_\alpha$ no more is a function in the sense of real analysis. The following property of these relations is obvious from the definiton of the α-cut:

$$C(T^*)_{\alpha_1} \subseteq C(T^*)_{\alpha_2}, \qquad \alpha_2 \leq \alpha_1. \tag{12}$$

A precise statement on the shape of the relation $C(T^*)_\alpha$ is possible, if $f(\omega|\boldsymbol{x})$ is continuous in $\boldsymbol{x}$ – not necessarily in ω. In [6] it has been shown that $C(f(\omega|\boldsymbol{x}^*))_\alpha$ is equal to an interval:

$$C(f(\omega|\boldsymbol{x}^*))_\alpha = [\underline{(f)}_\alpha(\omega), \overline{(f)}_\alpha(\omega)] = \left[\min_{\boldsymbol{x}\in C(\boldsymbol{x}^*)_\alpha} f(\omega|\boldsymbol{x}), \max_{\boldsymbol{x}\in C(\boldsymbol{x}^*)_\alpha} f(\omega|\boldsymbol{x})\right]. \tag{13}$$

Thus the contours of the relation $C(T^*)_\alpha$ are given by two functions in ω: the upper contour $\overline{(f)}_\alpha(\omega)$ and the lower contour $\underline{(f)}_\alpha(\omega)$.

For ω fixed the family $(C(f(\omega|\boldsymbol{x}^*))_\alpha)_{\alpha\in(0,1]}$ is the α-cut representation of a fuzzy number $f(\omega|\boldsymbol{x}^*)$. Thus for fuzzy data $\boldsymbol{x}^*$ the function $f(\omega|\boldsymbol{x})$ is generalized to a "fuzzy valued function" $f(\omega|\boldsymbol{x}^*)$ assigning a fuzzy functional value to each ω. For the one-dimensional case $\Omega \subseteq \mathbb{R}$ the fuzzy valued function $f(\omega|\boldsymbol{x}^*)$ is a "mountain" in the $(\omega, f(\omega))$-plane with a height between 0 and 1. To represent this function the contours of this mountain are projected onto the $(\omega, f(\omega))$-plane for certain α-levels (α-level curves, [6]). These contours are identical with the contours of the relation $C(T^*)_\alpha$ given in (13). The functions given in (13) will be called α-contours of the fuzzy valued function $f(\omega|\boldsymbol{x}^*)$ throughout the rest of this paper.

3 Bayes' Theorem for Fuzzy Data and Fuzzy Prior Parameters

Consider a stochastic quantity $\tilde{x}$ which follows a distribution with probability density function (p.d.f.) belonging to a parametric family:

$$\tilde{x} \sim f(x|\boldsymbol{\theta}), \quad \boldsymbol{\theta} = (\theta_1, \ldots, \theta_m) \in \Theta \subseteq \mathbb{R}^m.$$

In Bayesian statistics one assumes that the parameter $\boldsymbol{\theta}$ of the p.d.f $f(x|\boldsymbol{\theta})$ is a stochastic quantity, too. A priori $\boldsymbol{\theta}$ is distributed according to some prior p.d.f. $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)$. The prior parameter $\boldsymbol{\eta}_0 \in E$ has to be known. Given a sample $\boldsymbol{x} = (x_1, \ldots, x_n)$ of n stochastically independent observations of $\tilde{x}$ the p.d.f. of $\boldsymbol{\theta}$ is updated by means of Bayes' theorem:

$$\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) \propto g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}), \tag{14}$$

$$g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0) \cdot L(\boldsymbol{\theta}|\boldsymbol{x}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0) \cdot \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}).$$

$\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ is called the posterior p.d.f. of $\boldsymbol{\theta}$, $g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ is called the non-normalized posterior p.d.f. In this section the non-normalized posterior $g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ obtained from Bayes' theorem is extended to fuzzy data and fuzzy prior parameters.

3.1 Bayes' Theorem for Fuzzy Data

[7] was the first who considered Bayes' theorem for fuzzy data by introducing a fuzzy valued posterior p.d.f. In this paper we will start our investigations with the non-normalized posterior provided by Bayes' theorem (14) rather than with the (normalized) posterior p.d.f. For fuzzy data the non-normalized posterior p.d.f. is a fuzzy valued function. Subsection 3.1.1 describes propagation of fuzziness of data through Bayes' theorem by expressing the fuzziness of the non-normalized posterior p.d.f. in terms of the fuzziness of the data.

3.1.1 Propagation of Fuzziness

Let $\tilde{x}$ be a stochastic quantity distributed according to $f(x|\boldsymbol{\theta})$, let $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)$ be the prior p.d.f. of $\boldsymbol{\theta}$ with known prior parameter $\boldsymbol{\eta}_0$ and let $\boldsymbol{x}^* = (x_1{}^*, \ldots, x_n{}^*)$ be a fuzzy sample with α-cut representation $(C(\boldsymbol{x}^*)_\alpha)_{\alpha \in (0,1]}$. The non-normalized posterior p.d.f. of $\boldsymbol{\theta}$ is then a fuzzy valued function $g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}^*)$ with the α-contours obtained from (13):

$$\underline{(g_n)}_\alpha(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0) \cdot \min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} L(\boldsymbol{\theta}|\boldsymbol{x}), \qquad \forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1, \tag{15}$$

$$\overline{(g_n)}_\alpha(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0) \cdot \max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} L(\boldsymbol{\theta}|\boldsymbol{x}), \qquad \forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1. \tag{16}$$

Lemma 1 If $\boldsymbol{x}^*$ is a minimum rule fuzzy sample, then the α-contours are given by:

$$\underline{(g_n)}_\alpha(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)\prod_{i=1}^{n}\underline{(f_i)}_\alpha(\boldsymbol{\theta}), \qquad \forall\boldsymbol{\theta}\in\Theta, \quad 0<\alpha\leq 1, \tag{17}$$

$$\overline{(g_n)}_\alpha(\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)\prod_{i=1}^{n}\overline{(f_i)}_\alpha(\boldsymbol{\theta}), \qquad \forall\boldsymbol{\theta}\in\Theta, \quad 0<\alpha\leq 1, \tag{18}$$

with

$$\underline{(f_i)}_\alpha(\boldsymbol{\theta}) = \min_{x\in C(x_i^*)_\alpha} f(x|\boldsymbol{\theta}), \quad \overline{(f_i)}_\alpha(\boldsymbol{\theta}) = \max_{x\in C(x_i^*)_\alpha} f(x|\boldsymbol{\theta}).$$

Proof. For a minimum rule fuzzy sample the α-cut $C(\boldsymbol{x}^*)_\alpha$ is equal to the cartesian product of the α-cuts of the components ([6]). From (15) we obtain:

$$\begin{aligned}\underline{(g_n)}_\alpha(\boldsymbol{\theta}) &= \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)\cdot\min_{\boldsymbol{x}\in C(\boldsymbol{x}^*)_\alpha} L(\boldsymbol{\theta}|\boldsymbol{x}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)\cdot\min_{\boldsymbol{x}\in C(\boldsymbol{x}^*)_\alpha}\prod_{i=1}^{n} f(x_i|\boldsymbol{\theta}) = \\ &= \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)\prod_{i=1}^{n}\min_{x_i\in C(x_i^*)_\alpha} f(x_i|\boldsymbol{\theta}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)\prod_{i=1}^{n}\underline{(f_i)}_\alpha(\boldsymbol{\theta}).\end{aligned}$$

This result proves (17). In the same way (18) is proved. □

The functions $\underline{(f_i)}_\alpha(\boldsymbol{\theta})$ and $\overline{(f_i)}_\alpha(\boldsymbol{\theta})$ are α-contours of a fuzzy valued function denoted by $f({x_i}^*|\boldsymbol{\theta})$.

3.1.2 Sequential Bayes' Analysis

Lemma 1 states that for a minimum rule fuzzy sample the fuzziness of data is propagated through Bayes' theorem indepently for each fuzzy data point ${x_i}^*$ by the fuzzy function $f({x_i}^*|\boldsymbol{\theta})$.

This important property is lost for other combining rules such as the product rule fuzzy samples (e.g. [10]): the α-contours of the non-normalized posterior given by (15) and (16) no longer factorize into n components each of them being a minimum or a maximum over x_i, only. This is due to the dependency of the components $x_1, \ldots, x_n$ in the α-cut $C(\boldsymbol{x}^*)_\alpha$ for combining rules other than the minimum rule. Using combining rules different from the minimum rule a new fuzzy data point ${x^*}_{n+1}$ requires the recalculation of the non-normalized posterior for *all* fuzzy data points ${x^*}_1, \ldots, {x^*}_n, {x^*}_{n+1}$. The sequential character of Bayes' theorem is lost although the data are stochastically independent.

For a minimum rule fuzzy sample, however, the sequential character of Bayes' theorem is preserved. Given the α-contours of the fuzzy valued non-normalized posterior $g_n(\boldsymbol{\theta}|\boldsymbol{x}^*, \boldsymbol{\eta}_0)$ the α-contours of the updated non-normalized fuzzy valued posterior $g_{n+1}(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}^*, {x_{n+1}}^*)$ simply are given by:

$$\underline{(g_{n+1})}_\alpha(\boldsymbol{\theta}) = \underline{(g_n)}_\alpha(\boldsymbol{\theta})\cdot\min_{x\in C(x_{n+1}^*)_\alpha} f(x|\boldsymbol{\theta}) = \underline{(g_n)}_\alpha(\boldsymbol{\theta})\cdot\underline{(f_{n+1})}_\alpha(\boldsymbol{\theta}), \tag{19}$$

$$\overline{(g_{n+1})}_\alpha(\boldsymbol{\theta}) = \overline{(g_n)}_\alpha(\boldsymbol{\theta})\cdot\max_{x\in C(x_{n+1}^*)_\alpha} f(x|\boldsymbol{\theta}) = \overline{(g_n)}_\alpha(\boldsymbol{\theta})\cdot\overline{(f_{n+1})}_\alpha(\boldsymbol{\theta}). \tag{20}$$

3.1.3 Fuzzy Valued Likelihood Function

For fixed $\boldsymbol{\theta}$ the lower α-contour of the non-normalized posterior given by (15) is obtained by searching for the sample $\boldsymbol{x}_{\min}$ in $C(\boldsymbol{x}^*)_\alpha$ which has the smallest contribution to the likelihood $L(\boldsymbol{\theta}|\boldsymbol{x})$, the upper α-contour is obtained by searching for the sample $\boldsymbol{x}_{\max}$ with the largest contribution to the likelihood. For a minimum rule fuzzy sample this n-dimensional optimization problem is reduced to a one-dimensional problem: the samples are easily found by searching for those data points $x_{i,\min}$ and $x_{i,\max}$ in $C(x_i^*)_\alpha$ that minimize and maximize the p.d.f. $f(x|\boldsymbol{\theta})$ for fixed $\boldsymbol{\theta}$. For fuzzy samples obtained from other combining rules the dimension of the optimization problem usually is not reducible to one dimension.

The functions given by $\underline{(L)}_\alpha(\boldsymbol{\theta}) := L(\boldsymbol{\theta}|\boldsymbol{x}_{\min})$ and $\overline{(L)}_\alpha(\boldsymbol{\theta}) := L(\boldsymbol{\theta}|\boldsymbol{x}_{\max})$ are the α-contours of a fuzzy valued function $L(\boldsymbol{\theta}|\boldsymbol{x}^*)$ which is called the "fuzzy valued likelihood function".

3.1.4 Example

Assume that the stochastic quantity $\tilde{x}$ follows an exponential distribution:

$$f(x|\lambda) = \lambda \exp^{-\lambda x} \tag{21}$$

with unknown parameter $\lambda \in I\!R^+$. A conjugate prior for λ is given by a gamma density with precise parameter $\boldsymbol{\eta}_0 = (\nu_0, \beta_0)$:

$$\pi(\lambda|\nu_0,\beta_0) = \frac{\beta_0^{\nu_0}}{\Gamma(\nu_0)}\lambda^{\nu_0-1}\exp(-\beta_0\lambda). \tag{22}$$

Assume that n fuzzy observations $x_i{}^*$ characterized by $\varphi_{x_i^*}(\cdot)$ or equivalently by their α-cut representation $C(x_i^*)_\alpha$ are given. Assume that these observations are combined by the minimum rule to a fuzzy sample $\boldsymbol{x}^*$.

To obtain the α-contours of the non-normalized posterior we have to minimize and maximize $f(x|\lambda)$ for fixed λ over $x \in C(x_i^*)_\alpha = [\underline{C}(x_i^*)_\alpha, \overline{C}(x_i^*)_\alpha]$ for $i = 1,\dots,n$. As $f(x|\lambda)$ is a decreasing function in x the arguments leading to the minimum and to the maximum, respectively, are given by:

$$x_{i,\min} = \overline{C}(x_i^*)_\alpha, \quad x_{i,\max} = \underline{C}(x_i^*)_\alpha.$$

Therefore we obtain the following α-contours of the fuzzy valued non-normalized posterior $g_n(\lambda|\boldsymbol{x}^*,\boldsymbol{\eta}_0)$:

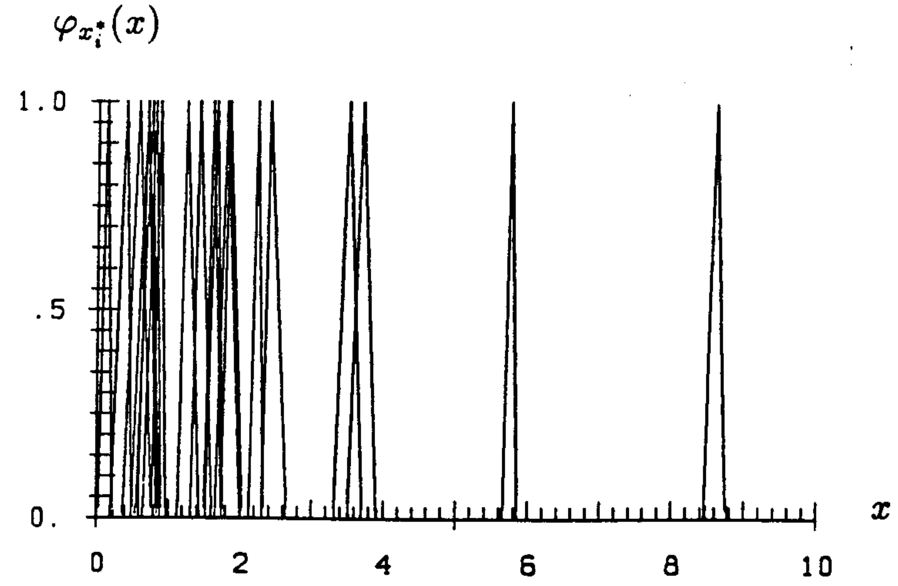
$$\underline{(g_n)}_\alpha(\lambda) = \frac{\beta_0^{\nu_0}}{\Gamma(\nu_0)}\lambda^{\nu_0+n-1}\exp\left(-(\beta_0 + \sum_{i=1}^n \overline{C}(x_i^*)_\alpha)\lambda\right),$$

$$\overline{(g_n)}_\alpha(\lambda) = \frac{\beta_0^{\nu_0}}{\Gamma(\nu_0)}\lambda^{\nu_0+n-1}\exp\left(-(\beta_0 + \sum_{i=1}^n \underline{C}(x_i^*)_\alpha)\lambda\right).$$

For fuzzy life time data the α-contours have a nice interpretation: the lower α-contour taking for each observation x_i the biggest possible value in $C(x_i^*)_\alpha$ corresponds to an optimistic and

Figure 1: Artificial fuzzy samples

a. Fuzzy sample 1



b. Fuzzy sample 2

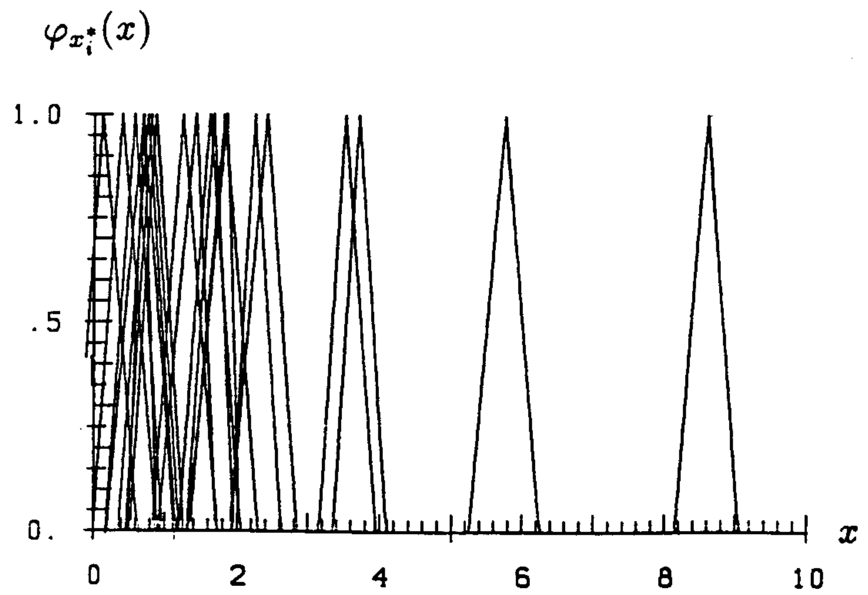
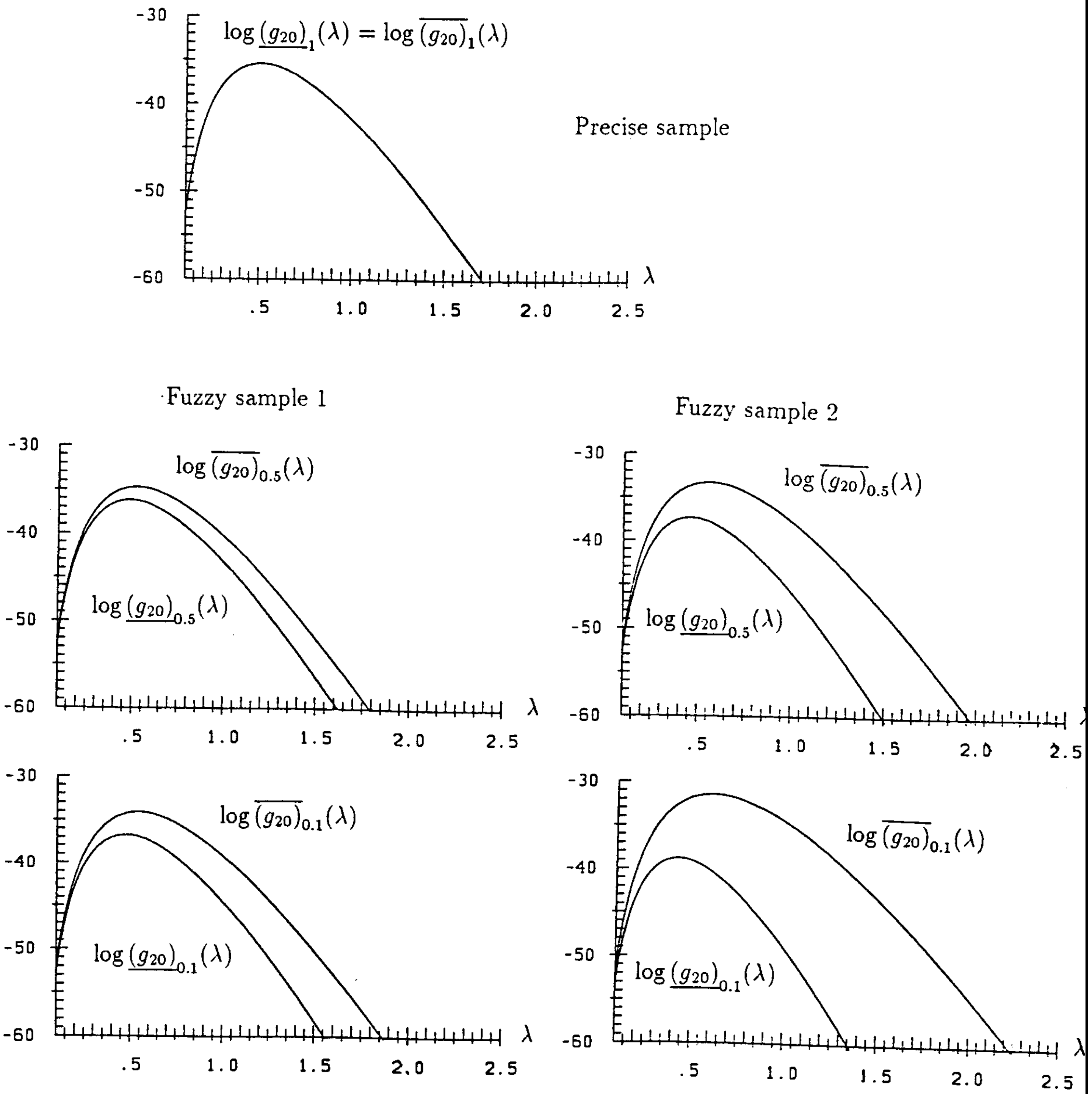


Figure 2: Fuzzy valued non-normalized posterior p.d.f.
(α-contours with $\alpha = 0.5$ and $\alpha = 0.1$)



daring analysis, whereas the upper α-contour taking for each observation x_i the smallest possible value in $C(x_i^*)_\alpha$ corresponds to the pessimistic and cautious one.

To give numerical results we have simulated a precise sample $m_1, \ldots, m_{20}$ of size $n = 20$ from an exponential distribution with $\lambda = 0.5$:

m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_{10}
0.40	1.41	1.80	0.87	0.81	2.40	2.23	1.66	0.69	0.58

m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_{16}	m_{17}	m_{18}	m_{19}	m_{20}
1.84	0.13	3.51	8.60	1.60	5.76	0.69	3.70	0.76	1.24

From this precise sample two fuzzy samples were derived by taking the simulated values m_i as the peak of a triangular fuzzy number $T^*(m_i, m_i, a_i, b_i)$ (see [6]). The parameter a_i and b_i determining the amount of fuzziness for the first fuzzy sample are smaller than the corresponding parameters for the second fuzzy sample:

Fuzzy sample 1

i	a_i	b_i	i	a_i	b_i
1	0.2	0.1	11	0.1	0.2
2	0.1	0.2	12	0.1	0.1
3	0.15	0.2	13	0.2	0.2
4	0.1	0.1	14	0.15	0.15
5	0.15	0.1	15	0.1	0.15
6	0.1	0.25	16	0.1	0.1
7	0.1	0.1	17	0.2	0.2
8	0.15	0.1	18	0.2	0.2
9	0.1	0.15	19	0.2	0.1
10	0.2	0.2	20	0.1	0.2

Fuzzy sample 2

i	a_i	b_i	i	a_i	b_i
1	0.2	0.5	11	0.5	0.2
2	0.5	0.3	12	0.4	0.5
3	0.45	0.5	13	0.35	0.45
4	0.35	0.4	14	0.45	0.45
5	0.35	0.4	15	0.4	0.45
6	0.5	0.45	16	0.5	0.5
7	0.3	0.4	17	0.2	0.2
8	0.35	0.3	18	0.35	0.4
9	0.5	0.45	19	0.4	0.4
10	0.4	0.4	20	0.4	0.5

These fuzzy samples are presented in Figure 1 by their characterizing functions.

Bayes' theorem was carried out with the prior's parameter equal to $\nu_0 = 2$ and $\beta_0 = 1$. In Figure 2 the non-normalized posterior for the precise sample is compared with the α-contours of the fuzzy valued non-normalized posterior of both fuzzy samples for $\alpha = 0.5$ and $\alpha = 0.1$. Especially for the second fuzzy sample the fuzziness of data has a considerable impact on the non-normalized posterior, e.g. on the posterior mode:

Posterior mode for Fuzzy Sample 1										
α	1.	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
upper α-contours	.504	.507	.511	.514	.518	.521	.525	.529	.532	.536
lower α-contours	.504	.500	.496	.493	.489	.486	.482	.479	.475	.472

Posterior mode for Fuzzy Sample 2										
α	1.	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
upper α-contours	.504	.514	.524	.535	.546	.558	.570	.582	.596	.610
lower α-contours	.504	.494	.485	.476	.467	.459	.451	.443	.436	.428

3.2 Bayes' Theorem for Non-precise Parameters of the Prior

The choice of the parameter $\boldsymbol{\eta}_0 \in E$ of the prior p.d.f. $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)$ may be subjected to uncertainty in practice. In the following we express this uncertainty in terms of a fuzzy set $\boldsymbol{\eta}_0{}^*$ with characterizing function $\varphi_{\boldsymbol{\eta}_0^*}(\boldsymbol{\eta}_0)$. A related approach is considered by robust Bayesian methods ([1]), where the uncertainty about $\boldsymbol{\eta}_0$ is expressed by a n-dimensional interval $M \subset E$. As an interval is a special case of a fuzzy set with characterizing function given by:

$$\varphi_{\boldsymbol{\eta}_0^*}(\boldsymbol{\eta}_0) = I_M(\boldsymbol{\eta}_0),$$

the approach discussed in this section may be viewed as a generalization of robust Bayesian methods.

3.2.1 Propagation of Fuzziness

Let $\tilde{x}$ be a stochastic quantity distributed according to $f(x|\boldsymbol{\theta})$, let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a precise sample of $\tilde{x}$, and let $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)$ be the prior p.d.f. of $\boldsymbol{\theta}$ with fuzzy prior parameter $\boldsymbol{\eta}_0{}^*$ characterized by the α-cut representation $(C(\boldsymbol{\eta}_0{}^*)_\alpha, \alpha \in (0,1])$. The non-normalized posterior p.d.f. of $\boldsymbol{\theta}$ is then a fuzzy valued function $g_n(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{\eta}^*{}_0)$ with the α-contours given by (13):

$$\underline{(g_n)}_\alpha(\boldsymbol{\theta}) = L(\boldsymbol{\theta}|\boldsymbol{x}) \cdot \min_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0), \qquad \forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1, \tag{23}$$

$$\overline{(g_n)}_\alpha(\boldsymbol{\theta}) = L(\boldsymbol{\theta}|\boldsymbol{x}) \cdot \max_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0), \qquad \forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1. \tag{24}$$

The fuzzy valued function in (23) and (24) with α-contours given by

$$\underline{(\pi)}_\alpha(\boldsymbol{\theta}) = \min_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0), \qquad \overline{(\pi)}_\alpha(\boldsymbol{\theta}) = \max_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0), \tag{25}$$

is called the fuzzy valued prior $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0{}^*)$.

Bayes' theorem for precise data and fuzzy prior parameter may be directly written in sequential form:

$$\underline{(g_n)}_\alpha(\boldsymbol{\theta}) = \underline{(g_{n-1})}_\alpha(\boldsymbol{\theta}) \cdot f(x_n|\boldsymbol{\theta}), \qquad \overline{(g_n)}_\alpha(\boldsymbol{\theta}) = \overline{(g_{n-1})}_\alpha(\boldsymbol{\theta}) \cdot f(x_n|\boldsymbol{\theta}), \tag{26}$$

with

$$\underline{(g_0)}_\alpha(\boldsymbol{\theta}) = \underline{(\pi)}_\alpha(\boldsymbol{\theta}), \qquad \overline{(g_0)}_\alpha(\boldsymbol{\theta}) = \overline{(\pi)}_\alpha(\boldsymbol{\theta}).$$

3.2.2 Example

Let us continue the example of Subsection 3.1.4. Again a conjugate prior is chosen with a precise number ν_0 of degrees of freedom; however, we are uncertain about β_0. We express this uncertainty in terms of a fuzzy number ${\beta_0}^*$ with α-cut representation $C(\beta_0^*)_\alpha = [\underline{C}(\beta_0^*)_\alpha, \overline{C}(\beta_0^*)_\alpha]$. Assume that a precise sample $\boldsymbol{x}$ is given.

Lemma 2 The fuzzy valued prior $\pi(\lambda|\nu_0, {\beta_0}^*)$ is a fuzzy valued function with α-contours given by:

$$\underline{(\pi)}_\alpha(\lambda) = \frac{\beta_{\min}(\lambda)^{\nu_0}}{\Gamma(\nu_0)}\lambda^{\nu_0-1}\exp(-\beta_{\min}(\lambda)\cdot\lambda),$$
$$\overline{(\pi)}_\alpha(\lambda) = \frac{\beta_{\max}(\lambda)^{\nu_0}}{\Gamma(\nu_0)}\lambda^{\nu_0-1}\exp(-\beta_{\max}(\lambda)\cdot\lambda),$$

with

$$\beta_{\min}(\lambda) = \begin{cases} \underline{C}(\beta_0^*)_\alpha, & \lambda \leq \dfrac{\nu_0}{b(\underline{C}(\beta_0^*)_\alpha, \overline{C}(\beta_0^*)_\alpha)}, \\ \overline{C}(\beta_0^*)_\alpha, & \lambda \geq \dfrac{\nu_0}{b(\underline{C}(\beta_0^*)_\alpha, \overline{C}(\beta_0^*)_\alpha)}, \end{cases}$$

$$b(\underline{C}(\beta_0^*)_\alpha, \overline{C}(\beta_0^*)_\alpha) = (\overline{C}(\beta_0^*)_\alpha - \underline{C}(\beta_0^*)_\alpha)\ln\frac{\overline{C}(\beta_0^*)_\alpha}{\underline{C}(\beta_0^*)_\alpha},$$

$$\beta_{\max}(\lambda) = \begin{cases} \overline{C}(\beta_0^*)_\alpha, & \lambda \leq \dfrac{\nu_0}{\overline{C}(\beta_0^*)_\alpha}, \\ \dfrac{\nu_0}{\lambda}, & \dfrac{\nu_0}{\overline{C}(\beta_0^*)_\alpha} \leq \lambda \leq \dfrac{\nu_0}{\underline{C}(\beta_0^*)_\alpha}, \\ \underline{C}(\beta_0^*)_\alpha, & \lambda \geq \dfrac{\nu_0}{\underline{C}(\beta_0^*)_\alpha}. \end{cases}$$

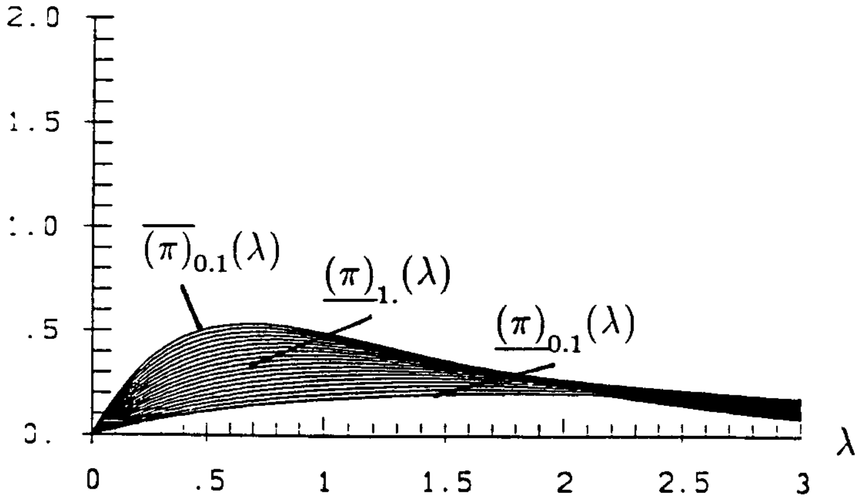
The fuzzy valued function $\pi(\lambda|\nu_0, {\beta^*}_0)$ is called "fuzzy valued gamma density".

Proof. Using (25) we have to determine the minimum and the maximum of the prior $\pi(\lambda|\nu_0,\beta)$ for β varying in $C(\beta_0^*)_\alpha$. Let us minimize and maximize the logarithm of this function. From

$$\frac{\partial \ln \pi(\lambda|\nu_0,\beta)}{\partial\beta} = \frac{\nu_0}{\beta} - \lambda \begin{cases} > 0 & \Leftrightarrow \quad \beta < \dfrac{\nu_0}{\lambda}, \\ = 0 & \Leftrightarrow \quad \beta = \dfrac{\nu_0}{\lambda}, \\ < 0 & \Leftrightarrow \quad \beta > \dfrac{\nu_0}{\lambda}, \end{cases} \tag{27}$$

it follows that for fixed $\lambda \in \mathbb{R}^+$ $\pi(\lambda|\nu_0,\beta)$ is an increasing function in β for $\beta < \dfrac{\nu_0}{\lambda}$ and a decreasing function in β for $\beta > \dfrac{\nu_0}{\lambda}$. Therefore $\pi(\lambda|\nu_0,\beta)$ takes the minimum in β on the

Figure 3: Fuzzy valued prior p.d.f.
(α-contours with $\alpha \geq 0.1$)



interval $C(\beta_0^*)_\alpha = [\underline{C}(\beta_0^*)_\alpha, \overline{C}(\beta_0^*)_\alpha]$ either at $\underline{C}(\beta_0^*)_\alpha$ or at $\overline{C}(\beta_0^*)_\alpha$ depending on λ. As it is easy to prove that for $\beta_1 < \beta_2$:

$$\pi(\lambda|\nu_0, \beta_1) \leq \pi(\lambda|\nu_0, \beta_2) \Leftrightarrow b(\beta_1, \beta_2) = (\beta_2 - \beta_1) \ln \frac{\beta_2}{\beta_1} \leq \frac{\nu_0}{\lambda},$$

$\beta_{\min}(\lambda)$ has obviously the form given above.

From (27) it follows that $\pi(\lambda|\nu_0, \beta)$ takes the maximum in β on the interval $C(\beta_0^*)_\alpha$ for $\beta_{\max}(\lambda) = \frac{\nu_0}{\lambda}$, if $\frac{\nu_0}{\lambda} \in C(\beta_0^*)_\alpha$. If $\frac{\nu_0}{\lambda} < \underline{C}(\beta_0^*)_\alpha$, $\pi(\lambda|\nu_0, \beta)$ is a decreasing function in β taking the maximum at $\underline{C}(\beta_0^*)_\alpha$; otherwise, if $\frac{\nu_0}{\lambda} > \overline{C}(\beta_0^*)_\alpha$, $\pi(\lambda|\nu_0, \beta)$ is an increasing function in β taking the maximum at $\overline{C}(\beta_0^*)_\alpha$. □

The α-contours of the prior together with (23) and (24) determine the α-contours of the fuzzy valued non-normalized posterior $g_n(\boldsymbol{\theta}|\boldsymbol{x}, \boldsymbol{\eta}_0{}^*)$:

$$\underline{(g_n)}_\alpha(\lambda) = \frac{\beta_{\min}(\lambda)^{\nu_0}}{\Gamma(\nu_0)} \lambda^{\nu_0+n-1} \exp\left(-(\beta_{\min}(\lambda) + \sum_{i=1}^n x_i)\lambda\right),$$
$$\overline{(g_n)}_\alpha(\lambda) = \frac{\beta_{\max}(\lambda)^{\nu_0}}{\Gamma(\nu_0)} \lambda^{\nu_0+n-1} \exp\left(-(\beta_{\max}(\lambda) + \sum_{i=1}^n x_i)\lambda\right).$$

The impact of the fuzziness of the prior parameter $\beta_0{}^*$ vanishes for increasing number of observations.

To give numerical results we combine the precise sample given in the example of Subsection 3.1.4 with a prior with $\nu_0 = 2$ and fuzzy parameter $\beta_0{}^*$ "about 1" by Bayes' theorem. The uncertainty about the prior parameter is expressed by a triangular fuzzy number $T^*(1, 1, 0.5, 0.5)$.

Figure 3 shows the α-contours of the fuzzy valued prior for $\alpha \geq 0.1$. Despite the considerable fuzziness of the prior the influence on the non-normalized posterior, e.g. on the posterior mode, is small:

Posterior mode for Precise Sample and Fuzzy Prior										
α	1.	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
upper level	.504	.504	.505	.506	.506	.507	.507	.508	.509	.509
lower level	.504	.503	.503	.502	.501	.501	.500	.500	.499	.498

3.3 Bayes' Theorem for Fuzzy Prior Parameters and Fuzzy Data

In this section we combine the results of the two previous sections in order to obtain a non-normalized posterior for both fuzzy prior parameter and fuzzy data. First we have to define a combined characterizing function $\varphi_{(\boldsymbol{\eta}_0, \boldsymbol{x})^*}(\boldsymbol{\eta}_0, \boldsymbol{x})$ for the fuzzy set $(\boldsymbol{\eta}_0, \boldsymbol{x})^*$. We use the minimum combining rule

$$\varphi_{(\boldsymbol{\eta}_0, \boldsymbol{x})^*}(\boldsymbol{\eta}_0, \boldsymbol{x}) = \min(\varphi_{\boldsymbol{\eta}_0^*}(\boldsymbol{\eta}_0), \varphi_{\boldsymbol{x}^*}(\boldsymbol{x})), \tag{28}$$

for which the α-cut representation of $(\boldsymbol{\eta}_0, \boldsymbol{x})^*$ is given by the cartesian product of the α-cut representations of $\boldsymbol{\eta}_0{}^*$ and $\boldsymbol{x}^*$:

$$C((\boldsymbol{\eta}_0, \boldsymbol{x})^*)_\alpha = C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha, \qquad \forall \alpha \in (0,1]. \tag{29}$$

The non-normalized posterior p.d.f. of $\boldsymbol{\theta}$ is then a fuzzy valued function $g_n(\boldsymbol{\theta}|\boldsymbol{x}^*, \boldsymbol{\eta}^*{}_0)$ with α-contours given by:

$$\underline{(g_n)}_\alpha(\boldsymbol{\theta}) = \left(\min_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0) \right) \cdot \left(\min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} L(\boldsymbol{\theta}|\boldsymbol{x}) \right), \tag{30}$$

$$\overline{(g_n)}_\alpha(\boldsymbol{\theta}) = \left(\max_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0) \right) \cdot \left(\max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} L(\boldsymbol{\theta}|\boldsymbol{x}) \right), \tag{31}$$

$$\forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1.$$

The choice of the combining rule (28) results in independent propagation of fuzziness of the prior parameter and the fuzziness of the data. Sequential analysis is possible for minimum rule combined fuzzy samples:

$$\underline{(g_n)}_\alpha(\boldsymbol{\theta}) = \underline{(g_{n-1})}_\alpha(\boldsymbol{\theta}) \cdot \min_{x \in C(x_n^*)_\alpha} f(x|\boldsymbol{\theta}), \tag{32}$$

$$\overline{(g_n)}_\alpha(\boldsymbol{\theta}) = \overline{(g_{n-1})}_\alpha(\boldsymbol{\theta}) \cdot \max_{x \in C(x_n^*)_\alpha} f(x|\boldsymbol{\theta}), \tag{33}$$

with

$$\underline{(g_0)}_\alpha(\boldsymbol{\theta}) = \underline{(\pi)}_\alpha(\boldsymbol{\theta}), \qquad \overline{(g_0)}_\alpha(\boldsymbol{\theta}) = \overline{(\pi)}_\alpha(\boldsymbol{\theta}).$$

It may be easily verified that the results of the Subsections 3.1 and 3.2 are special cases of the results of this subsection as a precise prior parameter $\boldsymbol{\eta}_0 = \boldsymbol{b}$ and a precise sample $\boldsymbol{x} = \boldsymbol{m}$ may be described by:

$$C(\boldsymbol{\eta}_0{}^*)_\alpha = \{\boldsymbol{b}\}, \qquad C(\boldsymbol{x}^*)_\alpha = \{\boldsymbol{m}\}, \quad \forall \alpha \in (0,1].$$

Therefore for the remaining part of the paper we will not distinguish between the source of fuzziness and speak of fuzzy prior parameter and fuzzy samples in general, even if the sample or the prior is precise.

4 Fuzzy Valued Posterior Probability Density Functions

4.1 General Remarks

The posterior p.d.f $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ is obtained from Bayes' theorem (14) by normalization:

$$\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) = \frac{g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})}{n(\boldsymbol{\eta}_0, \boldsymbol{x})}, \tag{34}$$

$$n(\boldsymbol{\eta}_0, \boldsymbol{x}) = \int g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})\, d\,\boldsymbol{\theta}. \tag{35}$$

In this subsection a fuzzy valued posterior p.d.f. is discussed. If the prior parameter is precise, our result is equivalent to the fuzzy valued posterior p.d.f. suggested in [11].

Let $\tilde{x}$ be a stochastic quantity distributed according to $f(x|\boldsymbol{\theta})$, let $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)$ be the prior p.d.f. of $\boldsymbol{\theta}$ with fuzzy prior parameter ${\boldsymbol{\eta}_0}^*$ characterized by the α-cut representation $(C({\boldsymbol{\eta}_0}^*)_\alpha, \alpha \in (0,1])$, and let $\boldsymbol{x}^* = ({x_1}^*, \ldots, {x_n}^*)$ be a fuzzy sample with α-cut representation $(C(\boldsymbol{x}^*)_\alpha, \alpha \in (0,1])$. The posterior p.d.f. of $\boldsymbol{\theta}$ is then a fuzzy valued function $\pi(\boldsymbol{\theta}|\boldsymbol{x}^*, {\boldsymbol{\eta}^*}_0)$ with the α-contours given by (13):

$$\underline{(\pi)}_\alpha(\boldsymbol{\theta}) = \min_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C({\boldsymbol{\eta}_0}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} \frac{g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})}{n(\boldsymbol{\eta}_0, \boldsymbol{x})}, \quad \forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1, \tag{36}$$

$$\overline{(\pi)}_\alpha(\boldsymbol{\theta}) = \max_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C({\boldsymbol{\eta}_0}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} \frac{g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})}{n(\boldsymbol{\eta}_0, \boldsymbol{x})}, \quad \forall \boldsymbol{\theta} \in \Theta, \quad 0 < \alpha \leq 1. \tag{37}$$

For precise data the posterior p.d.f. $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ and the non-normalized posterior p.d.f. $g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ are proportional. For fuzzy data and/or fuzzy priors a proportional relationship between the α-contours of $\pi(\boldsymbol{\theta}|\boldsymbol{x}^*, {\boldsymbol{\eta}_0}^*)$ and ${g_n}^*(\boldsymbol{\theta}|\boldsymbol{x}^*, {\boldsymbol{\eta}_0}^*)$ no longer holds. This loss of proportionality is due to the fact that the minimization and the maximization problem is not the same for the normalized (see (36) and (37)) and for the non-normalized posterior (see (30) and (31)) as the normalizing "constant" $n(\boldsymbol{\eta}_0, \boldsymbol{x})$ depends both on the data and the parameters of the prior.

The solution of the minimization and the maximization problem occuring in (36) and (37) may be untractable if the integration in (35) does not have an analytical solution. However, the solution is far more easily obtained for conjugate families.

4.2 Conjugate Fuzzy Bayesian Inference

For conjugate families the posterior p.d.f. of $\boldsymbol{\theta}$ belongs to the same parametric family as the prior p.d.f.:

$$\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_n), \qquad \boldsymbol{\eta}_n = S(\boldsymbol{\eta}_0, \boldsymbol{x}). \tag{38}$$

Therefore for conjugate families Bayes' theorem corresponds to a mapping $S(\cdot)$, $S : E \times I\!R^n \mapsto E$.

For conjugate families we suggest to propagate fuzziness in two steps:

1. Determination of the fuzzy posterior parameter $\boldsymbol{\eta}_n{}^*$ from the fuzzy sample $\boldsymbol{x}^*$ and the fuzzy prior parameter $\boldsymbol{\eta}_0{}^*$.
2. Fuzzification of the posterior p.d.f. $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) = \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_n)$ of $\boldsymbol{\theta}$ from the fuzzy posterior parameter $\boldsymbol{\eta}_n{}^*$.

(39) below describes the update of the fuzzy parameters of the posterior based on the fuzzy prior parameter and fuzzy data. Lemma 3 proves that this two-step-procedure is equivalent to the direct approach provided by (36) and (37).

4.2.1 Propagation of Fuzziness

Let $\tilde{x}$ be a stochastic quantity distributed according to $f(x|\boldsymbol{\theta})$, let $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0)$ be a conjugate prior p.d.f. of $\boldsymbol{\theta}$ with fuzzy prior parameter $\boldsymbol{\eta}_0{}^*$ characterized by the α-cut representation $(C(\boldsymbol{\eta}_0{}^*)_\alpha, \alpha \in (0,1])$, and let $\boldsymbol{x}^* = (x_1{}^*, \dots, x_n{}^*)$ be a fuzzy sample with α-cut representation $(C(\boldsymbol{x}^*)_\alpha, \alpha \in (0,1])$. The parameter of the posterior p.d.f. of $\boldsymbol{\theta}$ is then a fuzzy subset $\boldsymbol{\eta}_n{}^* \in E$ with the α-cuts given by (7):

$$C(\boldsymbol{\eta}_n^*)_\alpha = \bigcup_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \; \bigcup_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} S(\boldsymbol{\eta}_0, \boldsymbol{x}). \tag{39}$$

E.g., if the function $S(\cdot)$ is continuous in both arguments and the parameter $\eta \in E \subseteq \mathbb{R}$ is one-dimensional, then (39) simplifies to the following family of intervals:

$$C(\eta_n^*)_\alpha = \left[\min_{\eta_0 \in C(\eta_0{}^*)_\alpha} \min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} S(\eta_0, \boldsymbol{x}), \; \max_{\eta_0 \in C(\eta_0{}^*)_\alpha} \max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} S(\eta_0, \boldsymbol{x}) \right]. \tag{40}$$

Lemma 3 Let $\boldsymbol{\eta}_n{}^*$ be the fuzzy parameter of the posterior p.d.f. of $\boldsymbol{\theta}$ given by (39). The fuzzy valued function defined by the following α-contours:

$$\underline{(\pi)}_\alpha(\boldsymbol{\theta}) = \min_{\boldsymbol{\eta}_n \in C(\boldsymbol{\eta}_n{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_n), \qquad \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}, \forall \alpha \in (0,1], \tag{41}$$

$$\overline{(\pi)}_\alpha(\boldsymbol{\theta}) = \max_{\boldsymbol{\eta}_n \in C(\boldsymbol{\eta}_n{}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_n), \qquad \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}, \forall \alpha \in (0,1], \tag{42}$$

is identical with the fuzzy valued posterior p.d.f. given by (36) and (37).

Proof. Using Eq. (36) we obtain $\forall \boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $\forall \alpha \in (0,1]$:

$$\begin{aligned}
\underline{(\pi)}_\alpha(\boldsymbol{\theta}) &= \min_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} \frac{g_n(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})}{n(\boldsymbol{\eta}_0, \boldsymbol{x})} = \\
&= \min_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) = \\
&= \min_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} \pi(\boldsymbol{\theta}|S(\boldsymbol{\eta}_0, \boldsymbol{x})) = \\
&= \min_{\boldsymbol{\eta}_n \in C(\boldsymbol{\eta}_n^*)_\alpha} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_n).
\end{aligned}$$

This is exactly the same expression as in (41). In the same way the proof is carried out for the upper α-contours of the posterior. □

Figure 4: Fuzzy valued posterior p.d.f.
(α-contours with $\alpha \geq 0.1$)

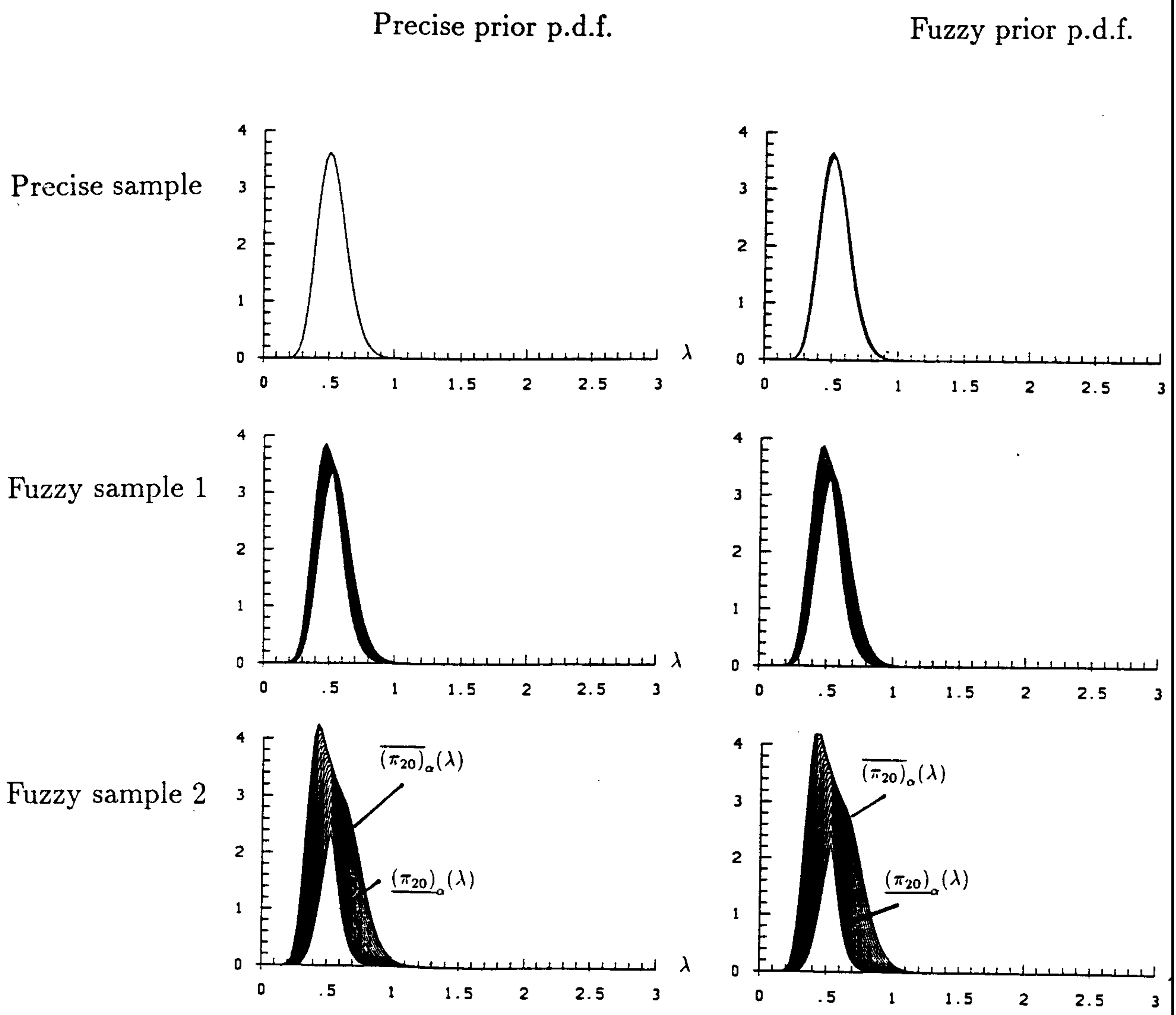
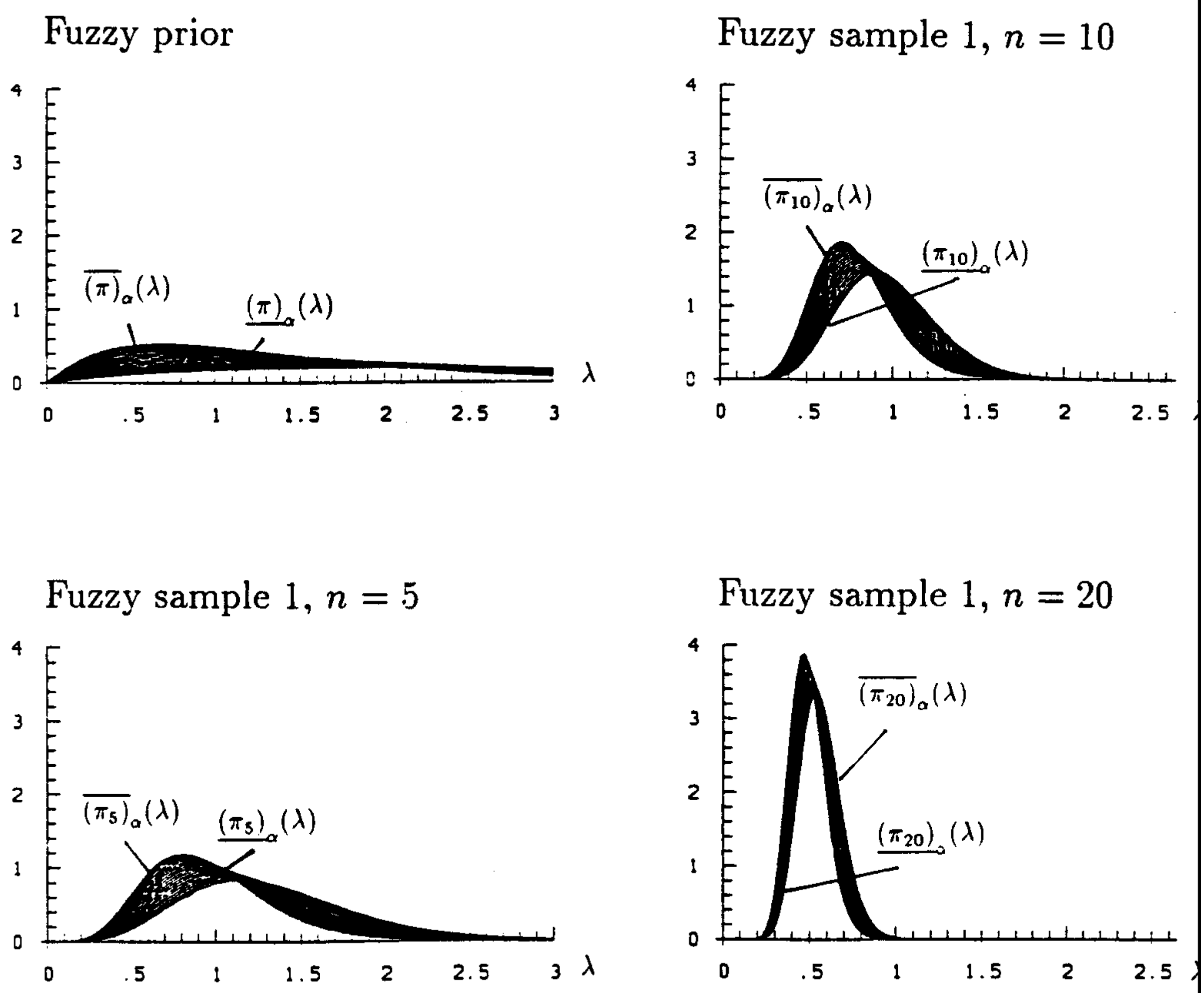


Figure 5: Fuzzy valued prior p.d.f. and fuzzy valued posterior p.d.f.
(α-contours with $\alpha \geq 0.1$)



4.2.2 Example

We consider again the case of an exponentially distributed random variable with unknown parameter λ. The updating mapping $S(\cdot)$ in (38) for the parameter $\boldsymbol{\eta}_n = (\nu_n, \beta_n)$ of the posterior of λ is given by:

$$\boldsymbol{\eta}_n = S(\boldsymbol{\eta}_0, \boldsymbol{x}) = \begin{pmatrix} \nu_0 + n \\ \beta_0 + \sum_{i=1}^{n} x_i \end{pmatrix}.$$

Now assume that the sample and the parameter ${\beta_0}^*$ of the prior are fuzzy. The first parameter ν_n remains precise whereas the second parameter ${\beta_n}^*$ is a fuzzy number with α-cuts given by (40):

$$\begin{aligned} C(\beta_n^*)_\alpha &= \left[\min_{\beta_0 \in C(\beta_0^*)_\alpha} \min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} (\beta_0 + \sum_{i=1}^{n} x_i), \max_{\beta_0 \in C(\beta_0^*)_\alpha} \max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} (\beta_0 + \sum_{i=1}^{n} x_i) \right] = \\ &= \left[\underline{C}(\beta_0^*)_\alpha + \min_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} (\sum_{i=1}^{n} x_i), \overline{C}(\beta_0^*)_\alpha + \max_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} (\sum_{i=1}^{n} x_i) \right]. \end{aligned} \tag{43}$$

For a minimum rule fuzzy samples (43) simplifies to:

$$C(\beta_n^*)_\alpha = \left[\underline{C}(\beta_0^*)_\alpha + \sum_{i=1}^{n} \underline{C}(x_i^*)_\alpha, \overline{C}(\beta_0^*)_\alpha + \sum_{i=1}^{n} \overline{C}(x_i^*)_\alpha \right].$$

The fuzzy valued posterior of λ is a fuzzy valued gamma density (see Lemma 2). It is derived from ν_n and the fuzzy parameter ${\beta_n}^*$ exactly in the same way as the fuzzy prior of λ was derived from ν_0 and the fuzzy prior parameter ${\beta_0}^*$ in Example 3.2.2.

To illustrate these results we consider again the precise and the two fuzzy samples introduced in Example 3.1.4 and combine them with the precise prior of λ used in Example 3.1.4 and the fuzzy prior of λ given in Example 3.2.2. Figure 4 presents the α-contours, $\alpha \geq 0.1$ of the fuzzy valued posterior density of λ for each of these six cases.

In this example the effect of the fuzziness of the prior on the fuzziness of the functional values of the posterior p.d.f. is small compared to the impact of the fuzziness of the data.

In order to illustrate the difference between randomness and fuzziness we compare the fuzzy valued posterior of λ with the fuzzy valued prior of λ for the first fuzzy sample by increasing n (n = 5, 10, 20). Figure 5 shows how increasing information – although fuzzy – reduces random variation in the parameter λ of the exponential distribution: the posterior gets sharper and sharper although the fuzziness of the functional values of the posterior p.d.f. remains.

5 Further Methods of Fuzzy Bayesian Inference

In this section we discuss the generalization of further methods of Bayesian inference for fuzzy data and fuzzy prior parameters, namely fuzzy posterior Bayes' estimates (Subsection 5.1), fuzzy predictive analysis (Subsection 5.2) and fuzzy H.P.D.-regions (Subsection 5.3).

5.1 Fuzzy Posterior Bayes' Estimate

The posterior Bayes' estimate $\hat{\boldsymbol{\theta}}_n$ is defined as the posterior expectation

$$\hat{\boldsymbol{\theta}}_n = \int_{\Theta} \boldsymbol{\theta} \pi(\boldsymbol{\theta} | \boldsymbol{\eta}_0, \boldsymbol{x}) \, d\boldsymbol{\theta}. \tag{44}$$

For conjugate families $\hat{\boldsymbol{\theta}}_n$ is a function of the parameter $\boldsymbol{\eta}_n$ of the posterior p.d.f.: $\hat{\boldsymbol{\theta}}_n = B(\boldsymbol{\eta}_n), \quad B : E \mapsto \Theta$.

5.1.1 Propagation of Fuzziness for Conjugate Families

Let $\boldsymbol{\eta}_n{}^*$ be the fuzzy posterior parameter given by Subsection 4.2.1. The posterior Bayes' estimate is then a fuzzy set $\hat{\boldsymbol{\theta}}_n{}^*$ in Θ with the α-cut representation given by (7):

$$C((\hat{\boldsymbol{\theta}}_n)^*)_\alpha = \bigcup_{\boldsymbol{\eta} \in C(\boldsymbol{\eta}_n^*)_\alpha} B(\boldsymbol{\eta}). \tag{45}$$

5.1.2 Example

For an exponential distribution the posterior Bayes' estimate of the parameter λ is given by:

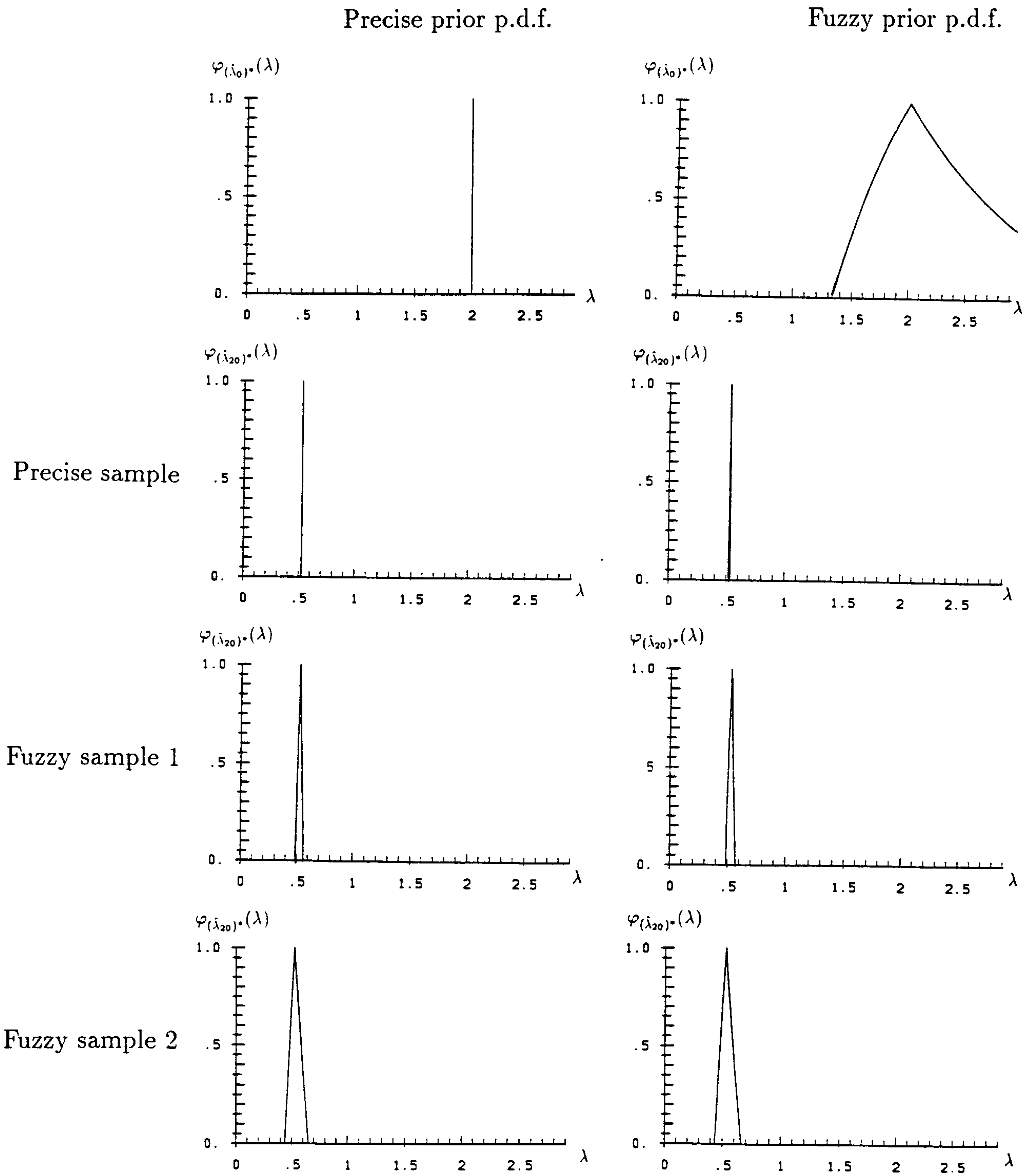
$$\hat{\lambda}_n = \frac{\nu_n}{\beta_n}.$$

Therefore

$$C((\hat{\lambda}_n)^*)_\alpha = \left[\min_{\beta \in C(\beta_n^*)_\alpha} \frac{\nu_n}{\beta}, \max_{\beta \in C(\beta_n^*)_\alpha} \frac{\nu_n}{\beta} \right] = \left[\frac{\nu_n}{\overline{C}(\beta_n^*)_\alpha}, \frac{\nu_n}{\underline{C}(\beta_n^*)_\alpha} \right].$$

Figure 6 compares the posterior Bayes' estimate $(\hat{\lambda}_n)^*$ for all posterior densities shown in Figure 4 with the prior expectation $(\hat{\lambda}_0)^*$. It is interesting to study the impact of the information and the fuzziness contained both in the prior and in the sample on the expectation. From the (fuzzy) prior we obtain the statement "Based on the prior the expected value of λ is (about) 2". For the fuzzy prior the imprecision associated with the term "about" is expressed by the fuzzy number shown in Figure 6. Given the (fuzzy) sample we obtain the statement "Based both on the prior and the data the expected value of λ is (about) 0.53". This considerable shift in the expectation of λ results from the stochastic information contained in the sample and takes place for precise and more or less fuzzy samples. For the fuzzy prior the imprecision associated with the term "about" is fairly reduced.

Figure 6: Fuzzy posterior Bayes' estimate



5.2 Fuzzy Predictive Analysis

Predictive Analysis is inference on the stochastic quantity $\tilde{x}$ in terms of the predictive density and the predictive distribution conditional on prior information and data. The predictive density $f(x|\boldsymbol{\eta}_0, \boldsymbol{x})$ is given by:

$$f(x|\boldsymbol{\eta}_0, \boldsymbol{x}) = \int_{\Theta} f(x|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})\, d\,\boldsymbol{\theta}. \tag{46}$$

The predictive distribution $F(x|\boldsymbol{\eta}_0, \boldsymbol{x})$ is obtained by the following integral:

$$F(x|\boldsymbol{\eta}_0, \boldsymbol{x}) = \int_{-\infty}^{x} f(z|\boldsymbol{\eta}_0, \boldsymbol{x})\, d\, z. \tag{47}$$

For a fuzzy sample and fuzzy prior parameters the predictive density and the predictive distribution are fuzzy valued functions which are defined below. If the prior parameter is precise, our result reduces to the fuzzy valued predictive density given in [10].

5.2.1 Propagation of Fuzziness

Let $\tilde{x}$ be a stochastic quantity distributed according to $f(x|\boldsymbol{\theta})$, let $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0{}^*)$ be the prior p.d.f. of $\boldsymbol{\theta}$ with fuzzy prior parameter $\boldsymbol{\eta}_0{}^*$ and let $\boldsymbol{x}^* = (x_1{}^*, \ldots, x_n{}^*)$ be a fuzzy sample of $\tilde{x}$. The predictive density and the predictive distribution of $\tilde{x}$ are then fuzzy valued functions $f(x|\boldsymbol{\eta}_0{}^*, \boldsymbol{x}^*)$ and $F(x|\boldsymbol{\eta}_0{}^*, \boldsymbol{x}^*)$ with α-contours given by:

$$\underline{(f)}_\alpha(x) = \min_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} f(x|\boldsymbol{\eta}_0, \boldsymbol{x}), \quad \forall x \in I\!R, \quad 0 < \alpha \leq 1, \tag{48}$$

$$\overline{(f)}_\alpha(x) = \max_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} f(x|\boldsymbol{\eta}_0, \boldsymbol{x}), \quad \forall x \in I\!R, \quad 0 < \alpha \leq 1, \tag{49}$$

$$\underline{(F)}_\alpha(x) = \min_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} F(x|\boldsymbol{\eta}_0, \boldsymbol{x}) \quad \forall x \in I\!R, \quad 0 < \alpha \leq 1, \tag{50}$$

$$\overline{(F)}_\alpha(x) = \max_{(\boldsymbol{\eta}_0, \boldsymbol{x}) \in C(\boldsymbol{\eta}_0{}^*)_\alpha \times C(\boldsymbol{x}^*)_\alpha} F(x|\boldsymbol{\eta}_0, \boldsymbol{x}) \quad \forall x \in I\!R, \quad 0 < \alpha \leq 1. \tag{51}$$

5.2.2 Example

Consider again an exponentially distributed random variable $\tilde{x}$. The predictive p.d.f. and the predictive distribution of x based on prior parameters $\boldsymbol{\eta}_0$ and based on a sample $\boldsymbol{x}$ is given by:

$$f(x|\boldsymbol{\eta}_0, \boldsymbol{x}) = f(x|\boldsymbol{\eta}_n) = \nu_n \cdot \frac{\beta_n^{\nu_n}}{(\beta_n + x)^{\nu_n + 1}},$$

$$F(x|\boldsymbol{\eta}_0, \boldsymbol{x}) = F(x|\boldsymbol{\eta}_n) = 1 - \left(\frac{\beta_n}{\beta_n + x}\right)^{\nu_n}.$$

Now let $\beta_0{}^*$ and $\boldsymbol{x}^*$ be fuzzy. This fuzziness causes $\beta_n{}^*$ to be fuzzy, whereas ν_n remains precise. In order to obtain the α-contours $\underline{(f)}_\alpha(x)$, $\overline{(f)}_\alpha(x)$, $\underline{(F)}_\alpha(x)$, and $\overline{(F)}_\alpha(x)$ of the predictive density and the predictive distribution we have to minimize and maximize $f(x|\boldsymbol{\eta}_n)$and

Figure 7: Fuzzy valued predicitve densities
(α-contours with $\alpha \geq 0.1$)

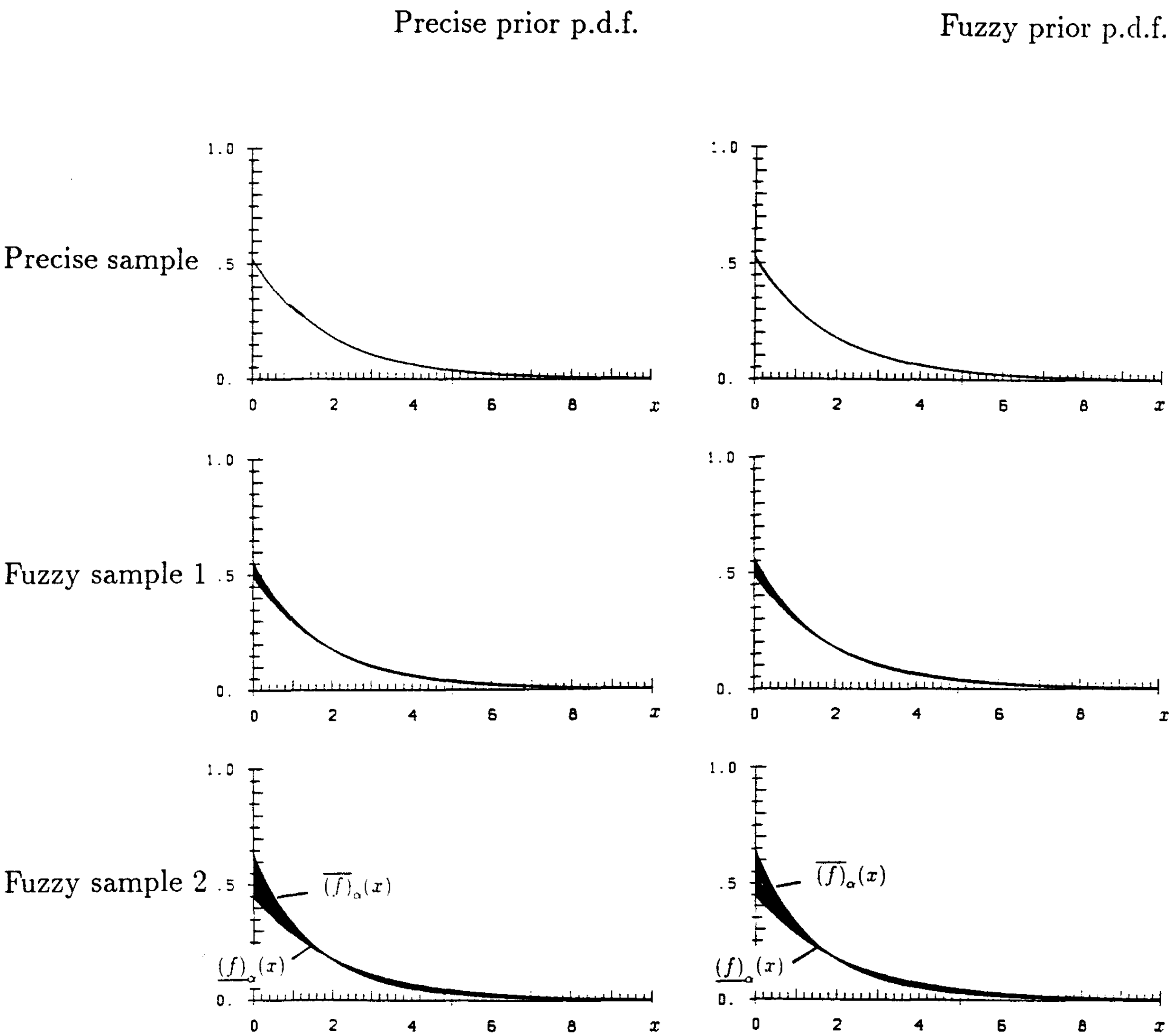
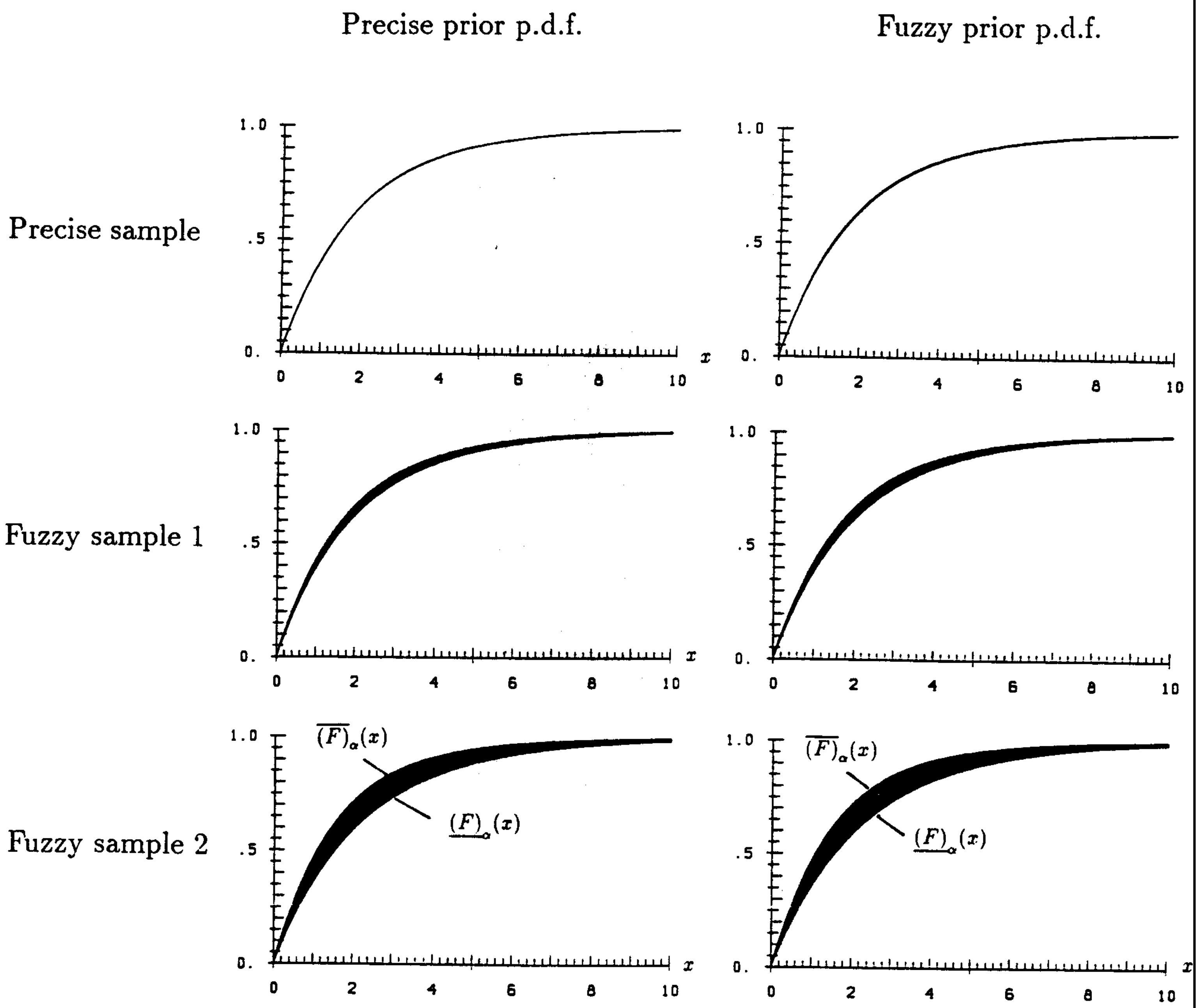


Figure 8: Fuzzy valued predicitve distributions
(α-contours with $\alpha \geq 0.1$)



$F(x|\boldsymbol{\eta}_n)$ on $C(\beta_n^*)_\alpha$ $\forall x \in I\!R$, $\forall \alpha \in (0,1]$. Let us minimize and maximize the logarithm of these functions. From

$$\frac{\partial \ln f(x|\nu_n,\beta)}{\partial \beta} = \frac{\nu_n}{\beta} - \frac{\nu_n+1}{\beta+x} \begin{cases} >0 & \Leftrightarrow & \beta < x\nu_n, \\ =0 & \Leftrightarrow & \beta = x\nu_n, \\ <0 & \Leftrightarrow & \beta > x\nu_n, \end{cases} \tag{52}$$

it follows that for fixed $x \in I\!R^+$ $f(x|\nu_n,\beta)$ is an increasing function in β for $\beta \le x\nu_n$ and a decreasing function in β for $\beta \ge x\nu_n$. Therefore $f(x|\nu_n,\beta)$ takes the minimum in β on the interval $C(\beta_n^*)_\alpha = [\underline{C}(\beta_n^*)_\alpha, \overline{C}(\beta_n^*)_\alpha]$ either at $\underline{C}(\beta_n^*)_\alpha$ or at $\overline{C}(\beta_n^*)_\alpha$ depending on x. As it is easy to prove that for $\beta_1 < \beta_2$:

$$f(x|\nu_n,\beta_1) \le f(x|\nu_n,\beta_2) \Leftrightarrow x \ge \frac{\beta_2\beta_1^{\frac{\nu_n}{\nu_n+1}} - \beta_1\beta_2^{\frac{\nu_n}{\nu_n+1}}}{\beta_2^{\frac{\nu_n}{\nu_n+1}} - \beta_1^{\frac{\nu_n}{\nu_n+1}}} = b(\beta_1,\beta_2),$$

we obtain:

$$\underline{(f)}_\alpha(x) = \nu_n \cdot \frac{\beta_{\min}(x)^{\nu_n}}{(\beta_{\min}(x)+x)^{\nu_n+1}},$$

with

$$\beta_{\min}(x) = \begin{cases} \underline{C}(\beta_n^*)_\alpha, & x \ge b(\underline{C}(\beta_n^*)_\alpha, \overline{C}(\beta_n^*)_\alpha), \\ \overline{C}(\beta_n^*)_\alpha, & x \le b(\underline{C}(\beta_n^*)_\alpha, \overline{C}(\beta_n^*)_\alpha). \end{cases}$$

From (52) it follows that $f(x|\nu_n,\beta)$ takes the maximum in β on the interval $C(\beta_n^*)_\alpha$ for $\beta_{\max}(x) = x\nu_n$, if $x\nu_n \in C(\beta_n^*)_\alpha$. If $x\nu_n < \underline{C}(\beta_n^*)_\alpha$, $f(x|\nu_n,\beta)$ is a decreasing function in β taking the maximum at $\underline{C}(\beta_n^*)_\alpha$; otherwise, if $x\nu_n > \overline{C}(\beta_n^*)_\alpha$, $f(x|\nu_n,\beta)$ is an increasing function in β taking the maximum at $\overline{C}(\beta_n^*)_\alpha$. Therefore

$$\overline{(f)}_\alpha(x) = \nu_n \cdot \frac{\beta_{\max}(x)^{\nu_n}}{(\beta_{\max}(x)+x)^{\nu_n+1}},$$

with

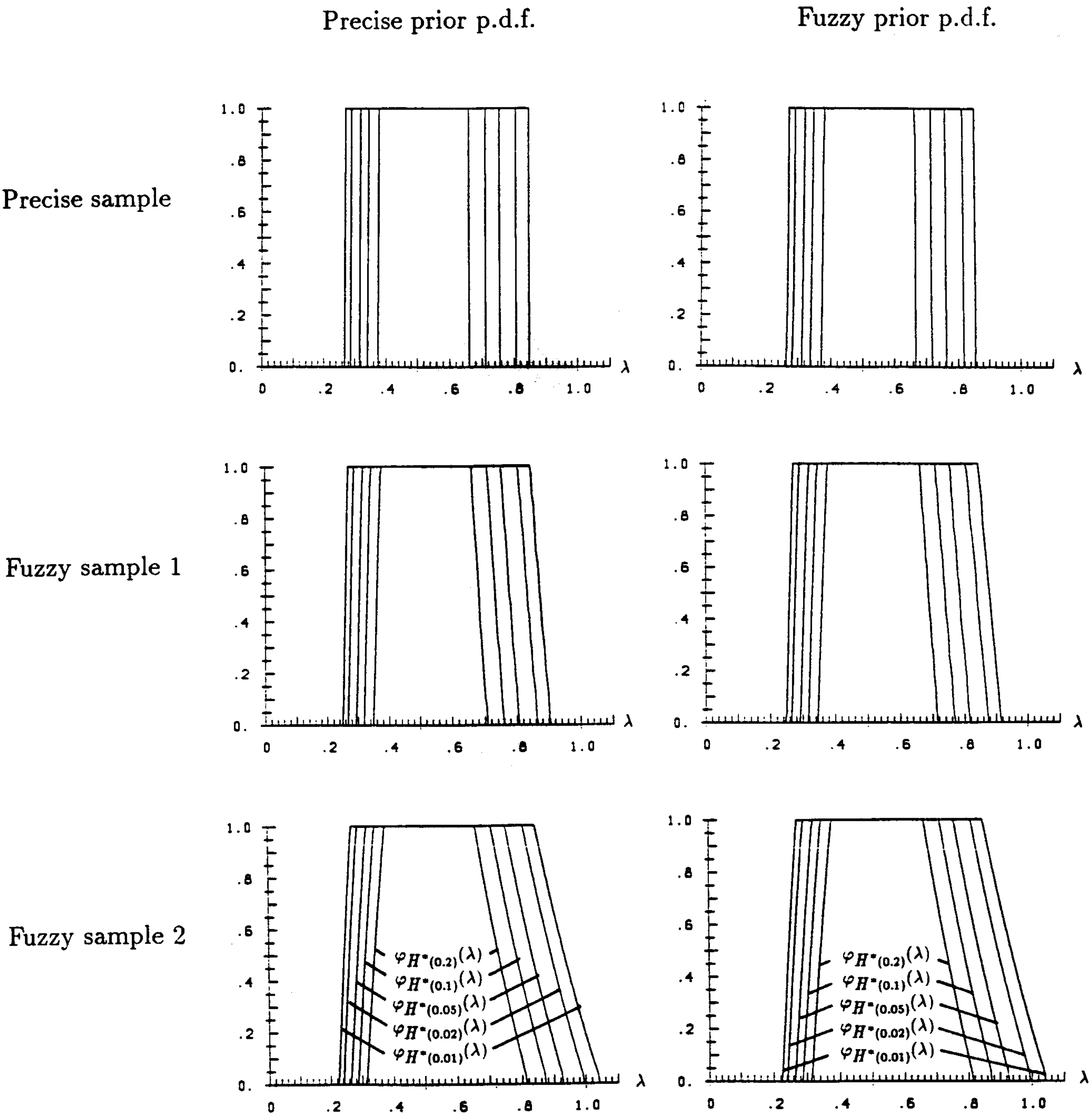
$$\beta_{\max}(x) = \begin{cases} \underline{C}(\beta_n^*)_\alpha, & x \le \frac{\underline{C}(\beta_n^*)_\alpha}{\nu_n}, \\ x\nu_n, & \frac{\underline{C}(\beta_n^*)_\alpha}{\nu_n} \le x \le \frac{\overline{C}(\beta_n^*)_\alpha}{\nu_n}, \\ \overline{C}(\beta_n^*)_\alpha, & x \ge \frac{\overline{C}(\beta_n^*)_\alpha}{\nu_n}. \end{cases}$$

From

$$\frac{\partial \ln(1-F(x|\nu_n,\beta))}{\partial \beta} = \nu_n\left(\frac{1}{\beta} - \frac{1}{\beta+x}\right) > 0$$

it follows that for fixed $x \in I\!R^+$ $F(x|\nu_n,\beta)$ is a decreasing function in β. Therefore $F(x|\nu_n,\beta)$ takes the minimum in β on the interval $C(\beta_n^*)_\alpha$ at $\overline{C}(\beta_n^*)_\alpha$ and the maximum at $\underline{C}(\beta_n^*)_\alpha$.

Figure 9: Fuzzy H.P.D. regions



Finally we obtain the following α-level curves of the predictive distribution:

$$\underline{(F)}_\alpha(x) = 1 - \left(\frac{\overline{C}(\beta_n^*)_\alpha}{\overline{C}(\beta_n^*)_\alpha + x}\right)^{\nu_n},$$

$$\overline{(F)}_\alpha(x) = 1 - \left(\frac{\underline{C}(\beta_n^*)_\alpha}{\underline{C}(\beta_n^*)_\alpha + x}\right)^{\nu_n}.$$

Figure 7 shows the fuzzy valued predictive densities for all cases discussed in the example of Subsection 4.2.2 ($\alpha \geq 0.1$) and Figure 8 shows the corresponding predictive distributions.

5.3 Fuzzy H.P.D.-Regions

A H.P.D.-region $H(\gamma)$ is a subset of the parameter space $\boldsymbol{\Theta}$ derived from the posterior p.d.f $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})$ in such a way that

1. $P(\boldsymbol{\theta} \in H(\gamma)) = \int_{H(\gamma)} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x})\, d\boldsymbol{\theta} \geq 1 - \gamma$
2. $\pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}) \leq d_\gamma,\ \forall \boldsymbol{\theta} \notin H(\gamma), \quad d_\gamma = \inf_{\boldsymbol{\theta} \in H(\gamma)} \pi(\boldsymbol{\theta}|\boldsymbol{\eta}_0, \boldsymbol{x}).$

For an unimodal posterior p.d.f. $H(\gamma)$ is the smallest region containing the unknown parameter with a probability greater or equal $1 - \gamma$.

Let γ be a fixed value. $H(\gamma)$ is then a subset of $\boldsymbol{\Theta}$ depending on the prior parameter $\boldsymbol{\eta}_0$ and on the sample $\boldsymbol{x}$. For fuzzy prior parameter ${\boldsymbol{\eta}_0}^*$ and for a fuzzy sample $\boldsymbol{x}^*$ we obtain a fuzzy H.P.D.-region which is a fuzzy set in $\boldsymbol{\Theta}$. A different approach of fuzzifying H.P.D.-regions was suggested in [11].

5.3.1 Propagation of Fuzziness

Let $H(\gamma, \boldsymbol{\eta}_0, \boldsymbol{x})$ be a H.P.D.-region depending on the sample $\boldsymbol{x}$ and on the prior parameter $\boldsymbol{\eta}_0$. Let $\boldsymbol{x}^*$ be a fuzzy sample and let ${\boldsymbol{\eta}_0}^*$ be a fuzzy prior parameter. Let $(\boldsymbol{\eta}_0, \boldsymbol{x})$ be combined by the minimum rule. $H(\gamma, \boldsymbol{\eta}_0, \boldsymbol{x})$ is then a fuzzy set $H(\gamma, {\boldsymbol{\eta}_0}^*, \boldsymbol{x}^*)$ in $\boldsymbol{\Theta}$ with an α-cut representation given by (7):

$$C(H(\gamma, {\boldsymbol{\eta}_0}^*, \boldsymbol{x}^*))_\alpha = H(\gamma, C({\boldsymbol{\eta}_0}^*)_\alpha, C(\boldsymbol{x}^*)_\alpha) = \bigcup_{\boldsymbol{\eta}_0 \in C({\boldsymbol{\eta}_0}^*)_\alpha} \bigcup_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} H(\gamma, \boldsymbol{\eta}_0, \boldsymbol{x}). \quad (53)$$

5.3.2 Example

For an exponential distribution the H.P.D.-region of the parameter λ is given by the interval

$$H(\gamma, \boldsymbol{\eta}_0, \boldsymbol{x}) = H(\gamma, \boldsymbol{\eta}_n) = H(\gamma, \nu_n, \beta_n) = \left[\frac{R_1(\gamma, \nu_n)}{\beta_n}, \frac{R_2(\gamma, \nu_n)}{\beta_n}\right],$$

where $R_1(\gamma, \nu)$ and $R_2(\gamma, \nu)$ is the H.P.D.-region of the standardized posterior

$$\frac{1}{\Gamma(\nu)} \lambda^{\nu-1} e^{-\lambda}.$$

The standard H.P.D.-regions have to be determined numerically for each ν and each γ. If the parameter $\beta_n{}^*$ of the posterior is fuzzy, we use (53) to construct the fuzzy interval $H(\gamma, \nu_n, \beta_n{}^*)$:

$$C(H^*(\gamma, \nu_n, \beta_n{}^*))_\alpha = \bigcup_{\boldsymbol{\eta}_0 \in C(\boldsymbol{\eta}_0{}^*)_\alpha} \bigcup_{\boldsymbol{x} \in C(\boldsymbol{x}^*)_\alpha} H(\gamma, \boldsymbol{\eta}_0, \boldsymbol{x}) =$$

$$= \bigcup_{\boldsymbol{\eta}_n \in C(\boldsymbol{\eta}_n{}^*)_\alpha} H(\gamma, \boldsymbol{\eta}_n) = \bigcup_{\beta_n \in C(\beta_n^*)_\alpha} H(\gamma, \nu_n, \beta_n) = \left[\frac{R_1(\gamma, \nu_n)}{\overline{C}(\beta_n^*)_\alpha}, \frac{R_2(\gamma, \nu_n)}{\underline{C}(\beta_n^*)_\alpha}\right].$$

Figure 9 shows these H.P.D.-regions for posterior p.d.f. of the example in Subsection 4.2.2 for different values of γ ($\gamma = 0.2, 0.1, 0.05, 0.02, 0.01$). The following table summarizes $H(\gamma, \nu_n, \beta_n{}^*)$ for $\alpha = 0.025$.

prior	sample	$\gamma = 0.01$	$\gamma = 0.02$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.2$
precise	precise	[.270,.844]	[.288,.806]	[.317,.752]	[.343,.708]	[.375,.659]
fuzzy	precise	[.267,.854]	[.285,.816]	[.313,.761]	[.339,.717]	[.371,.667]
precise	fuzzy 1	[.252,.903]	[.269, .863]	[.296,.805]	[.320,.758]	[.350,.705]
fuzzy	fuzzy 1	[.249,.915]	[.266, .874]	[.292,.815]	[.316,.768]	[.346,.714]
precise	fuzzy 2	[.227,1.040]	[.242, .993]	[.266,.927]	[.288,.873]	[.315,.812]
fuzzy	fuzzy 2	[.225,1.055]	[.240, 1.008]	[.264,.940]	[.285,.885]	[.312,.824]

The intervals given in the first line correspond to $H(\gamma, \boldsymbol{\eta}_n)$ for the non-fuzzy case. These intervals grow larger and larger the more fuzziness enters. The fuzziness of the prior parameter has little influence on the H.P.D.-regions compared to the impact of the fuzziness of the data.

To conclude we see that in this example the fuzziness of the data is dominated by the stochastic uncertainty: the imprecision of the boundaries of the H.P.D.regions is small compared to the length of the intervals representing stochastic uncertainty.

Acknowledgements

This research was supported by the Austrian Science Foundation, Project Nr. 6360.

References

[1] J.O. Berger: The Robust Bayesian Viewpoint. In: J. Kadane (ed.): *Robustness in Bayesian Statistics*, pp. 63-144. Amsterdam: North-Holland, 1984.

[2] D. Dubois, H. Prade: *Fuzzy Sets and Systems.* Theory and Applications. New York/London: Academic Press, 1980.

[3] R. Kruse, K.D. Meyer: *Statistics with Vague Data.* Theory and Decision Library. Series B, Mathematical and Statistical Methods. Dordrecht/Boston: D. Reidel Publishing Company, 1987.

[4] S. Schnatter: Linear Dynamic Systems and Fuzzy Data. In: R. Trappl (Ed.): *Cybernetics and Systems '90*, Singapore/New Jersey/London: World Scientific Publishing, 1990.

[5] S. Schnatter: On the Propagation of Fuzziness of Data. To appear in *Environmetrics*, **2**, 1991.

[6] S. Schnatter: *On Statistical Inference for Fuzzy Data with Applications to Descriptive Statistics.* Preprint. Submitted for publication to *Fuzzy Sets and Systems.*

[7] R. Viertl: Is it necessary to develop a Fuzzy Bayesian Inference? In: R. Viertl (Ed.): *Probability and Bayesian Statistics*, pp. 471 - 475. New York: Plenum Publishing Company, 1987.

[8] R. Viertl: Modelling of Fuzzy Measurements in Reliability Estimation. In: V. Colombari (Ed.): *Reliability Data Collection and Use in Risk and Availability Assessment.* Berlin: Springer-Verlag, 1989.

[9] R.Viertl: Estimation of the Reliability Function Using Fuzzy Life Time Data. In: P.K. Bose, S.P. Mukherjee, K.G. Ramamurthy (Eds.): *Quality for Progress and Development.* New Delhi: Wiley Eastern, 1989.

[10] R. Viertl: Statistical Inference for Fuzzy Data in Environmetrics. *Environmetrics*, **1**, 1990, 37-42.

[11] R. Viertl, H. Hule: On Bayes' Theorem for Fuzzy Data. *Statistical Papers*, **32**, 2, 1991, 115-122.