

JIRSS (2016)

Vol. 15, No. 2, pp 105-119

DOI: 10.18869/acadpub.jirss.15.2.105

Bayesian Fuzzy Hypothesis Testing with Imprecise Prior Distribution

Gholamreza Hesamian

Department of Statistics, Payame Noor University, Tehran, Iran

Abstract. This paper considers the classical (normal) Bayesian method for testing fuzzy hypotheses. For this purpose, using a notion of prior distribution with interval-valued or fuzzy-valued parameters, a concept of posterior probability of a fuzzy hypothesis is proposed and its main properties are also verified. The feasibility and effectiveness of the proposed methods are also clarified by some numerical examples.

Keywords. Fuzzy hypothesis, Bayes test, Interval parameter, Fuzzy parameter, Posterior probability.

MSC: 60A86, 03E72, 62C10.

1 Introduction

Hypothesis testing, also called significance testing, is the the most frequent statistical method widely used by scientists to guard against making claims unjustified by data. However, there are a variety of problems, difficulties, and inconsistencies associated with the most frequent methods of testing hypotheses that are overcome by using Bayesian analysis. However, the key ingredients to a Bayesian analysis are the likelihood function, which reflects information about the parameters contained in the data, and the prior distribution, which quantifies what is known about the parameters before observing data. The prior distribution and likelihood are then combined to the posterior distribution, which represents total knowledge about the parameters after the data have been

Gholamreza Hesamian (gh.hesamian@pnu.ac.ir)

observed.

Notably, a prior is a subjective assessment for the distribution of a parameter from an experienced expert. It is the probability distribution that would express expert's beliefs about a parameter(s) before some evidence is taken into account. However, when the experts agree about the family of prior distribution, the main concern is to assign some proper values for parameters which would satisfy all experts. For such cases, it is difficult to set exact values for parameters of the prior distribution due to uncertainty in estimations. In these cases, it is better that an expert estimates the value of a parameter θ based on the smallest possible value (or "pessimistic") and the highest possible value (or "optimistic") or some fuzzy sets instead of crisp ones. For instance, let X be a random variable denoting the expected time (per hour) to do a certain task. It is known that X is uniformly distributed in an interval $(0, \theta)$, $\theta > 0$. However, due to uncertainty related to the exact estimation, it is hard to say that our belief is that θ is distributed according to exponential distribution with an exact mean of 10 hours. For such a case, it is better to say that θ has an exponential distribution with a mean of about, for instance, 10 hours which may be modelled by an interval-valued $[8, 12]$ or a triangular fuzzy number of about, for instance, 10 hours.

On the other hand, the classical hypotheses testing may fail in many environmental practices and we may be forced to use fuzzy hypotheses testing as an alternative approach for demonstrating its utility in real applications Arnold (1996); Ortega *et al.* (2008); Parchami *et al.* (2011). Therefore, in order to achieve suitable Bayesian testing statistical methods dealing with imprecise information, there is a need to model the imprecise information and extend the classical methods of probability theory to imprecise environments.

The topic of probability theory with imprecise information has been studied by some authors (Baldwin *et al.* , 1996; Heilpern , 1993; Klement , 1982; Plasecki , 1985; Singpurwalla and Booker , 2004; Smets , 1982; Stein , 1985; Toth , 1992; Yager , 1984, see for example). Especially, introducing a loss function, Meeden (2008) suggested a fuzzy Bayes estimator for any given (crisp) prior density. He also investigated some conditions to transform the uncertainty expressed in a fuzzy decision into the (crisp) Bayesian paradigm. Moreover, there are several articles relevant to the Bayesian approach combining probability and/or fuzzy sets theory that can be found in literature Chib and Jacobi (2015); Chou and Yuan (1993); Gholizadeh *et al.* (2012); Jeet and Dhir (2012); Leu *et al.* (2015); Marttinen *et al.* (2009); Osoba *et al.* (2011, 2012); Viertl (2011); Xiao *et al.* (2013). In the sequel, there is a review of some studies relevant to testing statistical hypotheses based on Bayesian approaches. Casals (1993) considered the problem of Bayesian fuzzy hypotheses with fuzzy observations, as a special case of the fuzzy decision problem. Casals and Gil (1994) also

extended the Bayesian sequential problem of testing fuzzy hypotheses when the outcomes of which provide fuzzy information. Taheri and Behboodan Taheri and Behboodian (2001) considered the problem of Bayesian approach of testing hypotheses when the data (observations) are ordinary (crisp) and the hypotheses are fuzzy sets. Ortega *et al.* (2008) applied Taheri and Behboodian's method to evaluate whether the mother's serostatus and children's seroconversion to measles vaccine could be considered as high or low in a setting of fuzzy based-rules.

The purpose of this study is extending an approach to (Normal) Bayes test with unknown imprecise parameters given the crisp observations. For this purpose, after choosing a probability model for the (crisp) data, first a concept of prior distribution for the unknown fuzzy-valued or interval-valued parameters is introduced. Inspired by Zadeh's idea of probability of a fuzzy set, a posterior probability of the fuzzy hypotheses is introduced by averaging on the combination of the likelihood and prior distribution which quantifies the uncertainty in the values of the unknown imprecise parameters after the data are observed. The main properties of the proposed posterior probability are also put into investigation. For practical reasons, we will illustrate the proposed method by some numerical examples.

This paper is organized as follows. In the next section, we briefly review some definitions of fuzzy numbers and some results from possibility theory. In addition, the classical normal Bayes test is also reviewed briefly in this section. In Section 2, based on a crisp random sample of observations, first we propose a new concept of prior distribution with imprecise parameters which are reported by closed intervals or fuzzy numbers. Then, a notion of posterior probability of the underlying fuzzy hypotheses are introduced and discussed. Some numerical examples are provided in Section 4 to clarify the discussions made in this paper and to show the applicability of the proposed method in the fuzzy environment. Finally, a brief conclusion is provided in Section 5.

2 Preliminaries

A fuzzy set of \mathbb{X} is a mapping $\tilde{A} : \mathbb{X} \rightarrow [0, 1]$, which assigns to each $x \in \mathbb{X}$ a degree of membership $0 \leq \tilde{A}(x) \leq 1$. For each $\alpha \in (0, 1]$, the subset $\{x \in \mathbb{X} \mid \tilde{A}(x) \geq \alpha\}$ is called the level set or α -cut of \tilde{A} and is denoted by $\tilde{A}[\alpha]$. The lower and upper bounds of $\tilde{A}[\alpha]$ are denoted by $\tilde{A}_\alpha^L = \inf\{x : x \in \tilde{A}[\alpha]\}$ and $\tilde{A}_\alpha^U = \sup\{x : x \in \tilde{A}[\alpha]\}$, respectively. The set $\tilde{A}[0] = \{x \in \mathbb{X} : \tilde{A}(x) \neq 0\}$ is called the support of \tilde{A} . A fuzzy set \tilde{A} of \mathbb{R} is called a fuzzy number if it is normal, i.e. there exists a $x_A^* \in \mathbb{R}$ with $\tilde{A}(x_A^*) = 1$ and, for every $\alpha \in (0, 1]$, the set $\tilde{A}[\alpha]$ is a non-empty compact interval in

\mathbb{R} . This interval will be denoted by $\widetilde{A}[\alpha] = [\widetilde{A}_\alpha^L, \widetilde{A}_\alpha^U]$, where $\widetilde{A}_\alpha^L = \inf\{x : x \in \widetilde{A}[\alpha]\}$ and $\widetilde{A}_\alpha^U = \sup\{x : x \in \widetilde{A}[\alpha]\}$. In what follows, a fuzzy number is called continuous if its membership function is continuous over \mathbb{R} . The set of all continuous fuzzy numbers will be denoted by $\mathcal{F}_c(\mathbb{R})$. The imprecision or vagueness can be treated by means of special fuzzy numbers, i.e. the LR-fuzzy numbers, since they can be characterized using three real numbers: the center, the left spread, and the right spread. A LR-fuzzy number $\widetilde{A} = (a^l, a^{c1}, a^{c2}, a^r)_T$, $a^l, a^{c1}, a^{c2}, a^r \geq 0$, is defined as

$$\widetilde{A}(x) = \begin{cases} L\left(\frac{a^{c1}-x}{a^l}\right) & a^{c1} - a^l \leq x \leq a^{c1}, \\ 1 & a^{c1} \leq x \leq a^{c2}, \\ R\left(\frac{x-a^{c2}}{a^r}\right) & a^{c2} < x \leq a^{c2} + a^r, \\ 0 & x \in \mathbb{R} - [a^{c1} - a^l, a^{c2} + a^r], \end{cases}$$

where a^l and a^r are the left-hand and the right-hand spreads, respectively. L and R are continuous and strictly decreasing functions with $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$. The α -cut of \widetilde{A} is given as

$$\widetilde{A}[\alpha] = [\widetilde{A}_\alpha^L, \widetilde{A}_\alpha^U] = [a^{c1} - L^{-1}(\alpha)a^l, a^{c2} + R^{-1}(\alpha)a^r], \quad \alpha \in [0, 1].$$

A special type of LR-fuzzy numbers is the so-called trapezoidal fuzzy numbers with the shape functions $L(x) = R(x) = \max\{0, 1 - |x|\}$, $x \in \mathbb{R}$. The membership function of a trapezoidal fuzzy number $\widetilde{A} = (a, a^l, a^{c1}, a^{c2}, a^r)_T$ is given by

$$\widetilde{A}(x) = \begin{cases} 0 & x < a^{c1} - a^l, \\ \frac{x-(a^{c1}-a^l)}{a^l} & a - a^{c1} \leq x \leq a^{c1}, \\ 1 & a^{c1} \leq x \leq a^{c2}, \\ \frac{a^{c2}+a^r-x}{a^r} & a^{c2} \leq x \leq a^{c2} + a^r, \\ 0 & x > a + a^r. \end{cases}$$

Remark 1. (Hesamian and Shams, 2016) Let $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ and \widetilde{A}_α be defined for each $\alpha \in [0, 1]$ as

$$\forall \alpha \in [0, 1], \quad \widetilde{A}_\alpha = \begin{cases} \widetilde{A}_{2\alpha}^U & \alpha \in [0, 0.5], \\ \widetilde{A}_{2(1-\alpha)}^L & \alpha \in (0.5, 1]. \end{cases} \quad (2.1)$$

Then, the α -cuts of a fuzzy number $\widetilde{A} \in \mathcal{F}_c(\mathbb{R})$ is equal to $\widetilde{A}[\alpha] = [\widetilde{A}_{1-\alpha/2}^L, \widetilde{A}_{\alpha/2}^U]$, $\alpha \in [0, 1]$.

Examples 2.1. For a given LR-fuzzy number $\tilde{A} = (a^l, a, a^r)_{LR}$, it is easily seen that

$$\tilde{A}_\alpha = \begin{cases} a + (a^r - a)R^{-1}(2\alpha) & \text{for } 0.0 \leq \alpha \leq 0.5, \\ a - (a - a^l)L^{-1}(2(1 - \alpha)) & \text{for } 0.5 \leq \alpha \leq 1.0. \end{cases}$$

If $\tilde{A} = (a^l, a, a^r)_T$ is a triangular fuzzy number, then

$$\tilde{A}_\alpha = \begin{cases} a^r - 2\alpha(a^r - a) & \text{for } 0.0 \leq \alpha \leq 0.5, \\ 2a - a^l - 2\alpha(a - a^l) & \text{for } 0.5 \leq \alpha \leq 1.0. \end{cases}$$

Remark 2. The following operators on fuzzy numbers \tilde{A} and \tilde{B} (Dubois and Prade, 1998) are used in this paper:

1. $\tilde{A} \subseteq \tilde{B}$ if and only if $\tilde{A}(x) \leq \tilde{B}(x)$, for all $x \in \mathbb{R}$.
2. \tilde{A}^c is the complement of fuzzy number \tilde{A} with the membership function $\tilde{A}^c(x) = 1 - \tilde{A}(x)$, for all $x \in \mathbb{R}$.
3. $\tilde{A} \cup \tilde{B}$ is the union of fuzzy numbers \tilde{A} and \tilde{B} with the membership function $(\tilde{A} \cup \tilde{B})(x) = \max\{\tilde{A}(x), \tilde{B}(x)\}$, for all $x \in \mathbb{R}$.
4. $\tilde{A} \cap \tilde{B}$ is the intersection of fuzzy numbers \tilde{A} and \tilde{B} with the membership function $(\tilde{A} \cap \tilde{B})(x) = \min\{\tilde{A}(x), \tilde{B}(x)\}$, for all $x \in \mathbb{R}$.

2.1 The classical Bayesian testing hypothesis: The normal case

Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be a random sample from a parametric density function f_θ , $\theta \in \Theta \subseteq \mathbb{R}$, where θ denotes a population parameter. The observed random sample is denoted by $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$. The general format of the null and alternative hypotheses is as $H_0 : \theta \in \Theta_0$ and $H_1 : \theta \in \Theta_0^c$, where Θ_0 is a subset of the parameter space and Θ_0^c is its complement. However, hypothesis testing problems may be formulated in a Bayesian framework. A Bayesian framework includes both the sampling distribution $f_\theta(x)$ and the prior distribution $\pi(\theta)$, with the prior distribution reflecting the experimenter's opinion about the parameter θ prior to sampling. The Bayesian paradigm prescribes that the sample information should be combined with the prior information using Bayes's Theorem to obtain the posterior distribution $\pi(\theta|\mathbf{x}) = \frac{f_\theta(\mathbf{x})\pi(\theta)}{m(\mathbf{x})}$ where $m(\mathbf{x}) = \int f_\theta(\mathbf{x})\pi(\theta)d\theta$. Therefore, all inferences about θ are investigated based on the posterior distribution. In the Bayesian testing hypothesis problem, the posterior distribution may be used to calculate the probabilities that H_0 and H_1 are true. One way a Bayesian hypothesis tester may choose to use the posterior distribution is to consider H_0 as true if $P(\theta \in \Theta_0|\mathbf{x}) \geq P(\theta \in \Theta_0^c|\mathbf{x})$ and to reject H_0

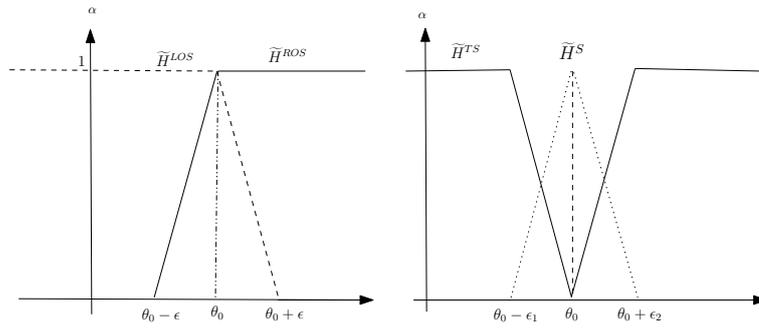


Figure 1: A typical type of fuzzy hypotheses.

otherwise. It is mentioned that the prior distribution for testing the two-sided hypothesis $H_0 : \theta = \theta_0$ v.s. $H_1 : \theta \neq \theta_0$ is given by

$$\pi(\theta) = \begin{cases} \pi_0 & \theta = \theta_0 \\ \pi_1 \pi'(\theta) & \theta \neq \theta_0, \end{cases}$$

where $\pi_0 + \pi_1 = 1$ (Berger , 1985; Bolstad , 2007, see for more).

3 Bayesian testing for fuzzy hypothesis with imprecise prior distribution

This section proposes a procedure for Bayesian testing statistical fuzzy hypothesis in case where the unknown parameters of the prior density are given by closed intervals or fuzzy numbers.

3.1 Fuzzy hypothesis

Here, some models for extending one-sided or two-sided ordinary (crisp) hypotheses to fuzzy ones are introduced.

Definition 3.1. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be a random sample from a parametric density function f_θ , $\theta \in \Theta \subseteq \mathbb{R}$.

1. Any hypothesis of the form

$$H : \theta \text{ is approximately } \theta_0 \quad \equiv \quad \tilde{H} : \theta \text{ is } \tilde{H}^S,$$

is called a simple hypothesis.

2. Any hypothesis of the form

$$H : \theta \text{ is essentially smaller than } \theta_0 \quad \equiv \quad \tilde{H} : \theta \text{ is } \tilde{H}^{LOS}$$

is called a fuzzy left one-sided hypothesis.

3. Any hypothesis of the form

$$H : \theta \text{ is essentially greater than } \theta_0 \quad \equiv \quad \tilde{H} : \theta \text{ is } \tilde{H}^{ROS}$$

is called a fuzzy right one-sided hypothesis.

4. Any hypothesis of the form

$$H : \theta \text{ is not approximately } \theta_0 \quad \equiv \quad \tilde{H} : \theta \text{ is } \tilde{H}^{TS}$$

is called a two-sided hypothesis.

For the above hypotheses, we suppose that

$$\begin{aligned} \tilde{H}^S &= (\theta_0 - \epsilon_1, \theta_0, \theta_0 + \epsilon_2)_T, \\ \tilde{H}^{LOS}(\theta) &= \begin{cases} 1 & \theta \leq \theta_0, \\ \frac{\theta_0 + \epsilon - \theta}{\epsilon} & \theta_0 \leq \theta \leq \theta_0 + \epsilon, \\ 0 & \theta_0 + \epsilon < \theta, \end{cases} \\ \tilde{H}^{ROS}(\theta) &= \begin{cases} 0 & \theta \leq \theta_0 - \epsilon, \\ \frac{\theta_0 - \theta}{\epsilon} & \theta_0 - \epsilon \leq \theta \leq \theta_0, \\ 1 & \theta > \theta_0, \end{cases} \\ \tilde{H}^{TS} &= 1 \ominus \tilde{H}^S, \end{aligned}$$

where $0 < \epsilon, \epsilon_1 \leq \theta_0$ and \ominus denotes the subtraction operator between fuzzy numbers. A typical form of fuzzy hypotheses $\tilde{H}^S, \tilde{H}^{TS}, \tilde{H}^{LOS}$ and \tilde{H}^{ROS} are shown in Fig. 1. In this paper, we are going to test the following hypotheses:

$$(1) \begin{cases} H_0 : \theta \text{ is } \tilde{H}^S \\ H_1 : \theta \text{ is } \tilde{H}^{TS} \end{cases} \quad (2) \begin{cases} H_0 : \theta \text{ is } \tilde{H}^{LOS} \\ H_1 : \theta \text{ is } \tilde{H}^{ROS} \end{cases} \quad (3) \begin{cases} H_0 : \theta \text{ is } \tilde{H}^{ROS} \\ H_1 : \theta \text{ is } \tilde{H}^{LOS} \end{cases}$$

3.2 Posterior probability of a fuzzy hypothesis

Suppose that the prior distribution $\pi(\theta)$ has a vector of imprecise parameters. In this regards, two sources of uncertainty are considered in this paper: 1) a vector of interval parameters $\bar{\gamma} = (\bar{\gamma}_1, \bar{\gamma}_2, \dots, \bar{\gamma}_k) \in (C(\Theta))^k$, where $C(\Theta)$ is the

class of nonempty compact intervals on real numbers \mathbb{R} and 2) a vector of fuzzy parameters $\tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_k) \in (\mathcal{F}_c(\mathbb{R}))^k$. Now, let X be a random sample from a parametric density function f_θ , $\theta \in \Theta \subseteq \mathbb{R}$. It is worth noting that Zadeh (1965) suggested to evaluate the probability of a fuzzy event \tilde{A} by $\tilde{P}(\tilde{A}) = \int_0^1 \int_{\tilde{A}[\alpha]} f_\theta(x) dx d\alpha$ (for more discussions on Zadeh's criterion see Singpurwalla and Booker (2004)). Here, inspired by Zadeh's idea, we introduce a posterior probability of the null and alternative fuzzy hypotheses as follows.

Definition 3.2. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be a random sample from a parametric density function f_θ , $\theta \in \Theta \subseteq \mathbb{R}$, and $\{\pi_{\tilde{\gamma}} : \tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_k) \in (C(\mathbb{R}))^p\}$ be a family of a prior distribution with interval-valued parameters. Then, the posterior probability of a fuzzy hypothesis \tilde{H} under the induced posterior density function $\pi_{\tilde{\gamma}}$ is defined as

$$\tilde{P}(\tilde{H}|\mathbf{x}) = \int_0^1 \int_{\tilde{H}[\beta]} g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) d\theta d\beta, \quad (3.1)$$

where

1. $\tilde{\gamma}_\beta = (\tilde{\gamma}_{1\beta}, \tilde{\gamma}_{2\beta}, \dots, \tilde{\gamma}_{k\beta})_T$ and $\tilde{\gamma}_{j\beta} = (1-\beta)\gamma_j^L + \beta\gamma_j^U$ is the convex combination of the lower and upper bounds of the closed interval $\tilde{\gamma}_j = [\gamma_j^L, \gamma_j^U]$,

- 2.

$$g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) = \frac{\pi_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) + \pi_{\tilde{\gamma}_{1-\beta}}(\theta|\mathbf{x})}{2},$$

in which $\pi_{\tilde{\gamma}_\beta}(\theta|\mathbf{x})$ denotes the posterior density function of \mathbf{X} with respect to the prior distribution $\{\pi_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) : \tilde{\gamma}_\beta = (\tilde{\gamma}_{1\beta}, \tilde{\gamma}_{2\beta}, \dots, \tilde{\gamma}_{k\beta})_T \in \mathbb{R}^k\}$ for any $\beta \in [0, 1]$.

Remark 3. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be a random sample from a parametric density function f_θ , $\theta \in \Theta \subseteq \mathbb{R}$, and $\{\pi_{\tilde{\gamma}} : \tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_k) \in (\mathcal{F}_c(\mathbb{R}))^p\}$ be a family of prior distribution with fuzzy-valued parameters. Then, similar to Definition 3.2, the probability of a fuzzy hypothesis \tilde{H} under the posterior density function $\pi_{\tilde{\gamma}}(\theta|\mathbf{x})$ is given as

$$\tilde{P}(\tilde{H}|\mathbf{x}) = \int_0^1 \int_{\tilde{H}[\beta]} g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) d\theta d\beta, \quad (3.2)$$

where

1. $\tilde{\gamma}_\beta = (\tilde{\gamma}_{1\beta}, \tilde{\gamma}_{2\beta}, \dots, \tilde{\gamma}_{k\beta})_T$ in which $\tilde{\gamma}_{j\beta}$ is introduced in Eq. (3.1).

2.

$$g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) = \frac{\pi_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) + \pi_{\tilde{\gamma}_{1-\beta}}(\theta|\mathbf{x})}{2},$$

in which $\pi_{\tilde{\gamma}_\beta}(\theta|\mathbf{x})$ denotes the posterior density function of \mathbf{X} with respect to the prior distribution $\{\pi_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) : \tilde{\gamma}_\beta = (\tilde{\gamma}_{1\beta}, \tilde{\gamma}_{2\beta}, \dots, \tilde{\gamma}_{k\beta})^T \in \mathbb{R}^k\}$, for any $\beta \in [0, 1]$.

From the operations on fuzzy numbers introduced in Remark 2, we have the following results.

Lemma 3.1. *Let \tilde{A} and \tilde{B} be two fuzzy numbers. Then,*

- 1) $\tilde{P}(\tilde{A}|\mathbf{x}) \in [0, 1]$,
- 2) if $\tilde{A} \subseteq \tilde{B}$, then $\tilde{P}(\tilde{A}|\mathbf{x}) \leq \tilde{P}(\tilde{B}|\mathbf{x})$,
- 3) $\tilde{P}(\tilde{A} \cup \tilde{B}|\mathbf{x}) = \tilde{P}(\tilde{A}|\mathbf{x}) + \tilde{P}(\tilde{B}|\mathbf{x}) - \tilde{P}(\tilde{A} \cap \tilde{B}|\mathbf{x})$,
- 4) if $\tilde{A}^c = 1 \ominus \tilde{A}$, then $\tilde{P}(\tilde{A}^c|\mathbf{x}) = 1 - \tilde{P}(\tilde{A}|\mathbf{x})$.

Proof. Let $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ be a random sample from a parametric density function f_θ , $\theta \in \Theta \subseteq \mathbb{R}$, and $\{\pi_{\tilde{\gamma}} : \tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_k) \in (\mathcal{F}_c(\mathbb{R}))^p\}$ be a family of prior distribution with fuzzy-valued parameters. Therefore, the assertions 1)-3) are immediately followed. To prove the assertion 4), note that

$$\tilde{P}(\tilde{A}^c|\mathbf{x}) = \int_0^1 \int_{(\tilde{A}[1-\beta])^c} g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) d\theta d\beta = 1 - \int_0^1 \int_{\tilde{A}[1-\beta]} g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) d\theta d\beta.$$

Let $y = 1 - \beta$. Since $g_{\tilde{\gamma}_\beta}(\theta|\mathbf{x}) = g_{\tilde{\gamma}_{1-\beta}}(\theta|\mathbf{x})$ for all $\beta \in [0, 1]$, it is easily concluded that

$$\tilde{P}(\tilde{A}^c|\mathbf{x}) = 1 - \int_0^1 \int_{\tilde{A}[y]} g_{\tilde{\gamma}_y}(\theta|\mathbf{x}) d\theta dy = 1 - \tilde{P}(\tilde{A}|\mathbf{x}),$$

which completes the proof. The proof for the case that the family of prior distributions is considered as $\{\pi_{\tilde{\gamma}} : \tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_k) \in (C(\mathbb{R}))^p\}$ with interval-valued parameters is similar. \square

Remark 4. It is pointed out that if the vector of interval-valued parameter $\tilde{\gamma}$ or fuzzy-valued $\tilde{\gamma}$ reduces to the crisp one, i.e. a vector of the real-valued parameter γ , then it is easy to observe that the posterior probability of the fuzzy hypothesis \tilde{H} is reduced as

$$\tilde{P}(\tilde{H}|\mathbf{x}) = \int_0^1 \int_{\tilde{H}[\beta]} \pi_\gamma(\theta|\mathbf{x}) d\theta d\beta = \int_{\mathbb{R}} \tilde{H}(\theta) \pi_\gamma(\theta|\mathbf{x}) d\theta,$$

which is the same result as Taheri and Behbodian's method (Taheri and Behbodian, 2001).

4 Numerical examples

In this section, we apply two numerical examples to demonstrate the proposed method.

Examples 4.1. Consider determining the true IQ for a student called Martin. Martin is taking an IQ test and his score x is normally distributed with the mean θ and variance $\sigma^2 = 100$. Before the IQ test is given, we strongly believe Martin's IQ has normal distribution with interval-mean $\bar{\mu} = [95, 105]$ and variance $\bar{\tau}^2 = [220, 230]$. If θ represents Martin's true IQ, suppose we are interested in testing the hypotheses H_0 : the Martin's IQ is essentially less than 100 against the alternative hypothesis H_1 : the Martin's IQ is essentially larger than 100. Based on the fuzzy hypothesis modeling introduced in this paper, we deal with a fuzzy hypothesis, which may be represented as

$$\begin{cases} H_0 : \theta \text{ is } \tilde{H}^{LOS} \\ H_1 : \theta \text{ is } \tilde{H}^{ROS}, \end{cases}$$

where

$$\tilde{H}^{LOS}(\theta) = \begin{cases} 1 & \theta \leq 100, \\ \frac{105-\theta}{5} & 100 \leq \theta \leq 105, \\ 0 & 105 < \theta, \end{cases}$$

$$\tilde{H}^{ROS}(\theta) = \begin{cases} 0 & \theta \leq 95, \\ \frac{100-\theta}{5} & 95 \leq \theta \leq 100, \\ 1 & \theta > 100. \end{cases}$$

Assuming that $\bar{y} = (\bar{\mu}, \bar{\sigma}^2)^T$ and regarding to Martin's observed IQ test score that is $x = 120$, we know that the posterior density for θ is normal with mean $\frac{100\bar{\mu}_\beta + 120\bar{\tau}_\beta^2}{100 + \bar{\tau}_\beta^2}$ and variance $\frac{100\bar{\tau}_\beta^2}{100 + \bar{\tau}_\beta^2}$, where $\bar{\mu}_\beta = 95 + 5\beta$ and $\bar{\tau}_\beta^2 = 220 + 5\beta$ at each level of $\beta \in [0, 1]$. According to Eq. (3.2), the posterior probability of \tilde{H}_0 is then equal to

$$\begin{aligned} \tilde{P}(\tilde{H}_0|x = 120) &= \frac{1}{2} \int_0^1 \int_{-\infty}^{105-5\beta} (n(\bar{\mu}_\beta, \bar{\tau}_\beta^2, \theta) + n(\bar{\mu}_{1-\beta}, \bar{\tau}_{1-\beta}^2, \theta)) d\theta d\beta \\ &= 0.0761, \end{aligned}$$

where $n(\mu, \tau^2)$ denotes the density function of normal distribution with mean μ and variance τ^2 . Similarly, the posterior probability of \tilde{H}_1 is given by

$$\begin{aligned} \tilde{P}(\tilde{H}_1|x = 120) &= \frac{1}{2} \int_0^1 \int_{95+5\beta}^{\infty} (n(\bar{\mu}_\beta, \bar{\tau}_\beta^2, \theta) + n(\bar{\mu}_{1-\beta}, \bar{\tau}_{1-\beta}^2, \theta)) d\theta d\beta \\ &= 0.9650. \end{aligned}$$

In this example, since $\tilde{P}(\tilde{H}_1|x = 120) \gg \tilde{P}(\tilde{H}_0|x = 120)$, the data strongly supports the alternative hypothesis that Martin’s IQ is essentially greater than the standard intelligence 100.

Now, assume that our prior beliefs are normal with the mean of about 100 and variance of about 225 represented by triangular fuzzy numbers $\tilde{\mu} = (95, 100, 105)_T$ and $\tilde{\sigma}^2 = (220, 225, 230)_T$, respectively. In this case, the posterior probability of \tilde{H}_0 and \tilde{H}_1 are evaluated as

$$\begin{aligned} \tilde{P}(\tilde{H}_0|x = 120) &= \frac{1}{2} \int_0^1 \int_{-\infty}^{105-5\beta} (n(\tilde{\mu}_\beta, \tilde{\tau}^2_\beta, \theta) + n(\tilde{\mu}_{1-\beta}, \tilde{\tau}^2_{1-\beta}, \theta))d\theta d\beta \\ &= 0.0630, \\ \tilde{P}(\tilde{H}_1|x = 120) &= \frac{1}{2} \int_0^1 \int_{95+5\beta}^{\infty} (n(\tilde{\mu}_\beta, \tilde{\tau}^2_\beta, \theta) + n(\tilde{\mu}_{1-\beta}, \tilde{\tau}^2_{1-\beta}, \theta))d\theta d\beta \\ &= 0.9701, \end{aligned}$$

where $\tilde{\mu}_\beta = 105 - 10\beta$ and $\tilde{\tau}^2_\beta = 230 - 10\beta$. So, with respect to the observed IQ value of 125, the null hypothesis H_1 is also strongly accepted in this case.

Examples 4.2. Bolstad (2007) In the study of water quality in New Zealand streams, a high level of Campylobacter was defined as a level of greater than 100 per 100 ml of stream water. A random sample of size $n = 116$ is taken from streams having a high environmental impact from birds. Out of these, $y = 11$ had a high Campylobacter level. Let θ be the true probability that a sample of water from this type of stream has a high Campylobacter level. We try to test the following two-sided fuzzy hypothesis:

$$\begin{cases} H_0 : \theta \text{ is } \tilde{H}^S = (0.05, 0.10, 0.15)_T \\ H_1 : \theta \text{ is not } \tilde{H}^{TS} = 1 \ominus \tilde{H}^S. \end{cases}$$

Let the prior distribution be suggested as

$$\pi(\theta) = \begin{cases} 0.5 & \theta = 0.1 \\ 0.5\pi'(\theta) & \theta \neq 0.1, \end{cases}$$

where $\pi'(\theta)$ is the beta distribution with fuzzy parameters $\tilde{a} = (0.7, 1, 1.4)_T$ and $\tilde{b} = (8, 10, 13)_T$. For calculating the probability of \tilde{H}_0 , first note that, for any $\beta \in [0, 1]$, the posterior probability is evaluated as

$$\pi_\beta(p|y = 11) = \frac{1}{m_\beta(\mathbf{x})} (0.5(0.1)^{11}(0.9)^{105} + (0.5) \frac{1}{\beta(\tilde{a}_\beta, \tilde{b}_\beta)} p^{10+\tilde{a}_\beta} (1-p)^{104+\tilde{b}_\beta}),$$

where

$$m_{\beta}(y = 11) = 0.5(0.5)^{11}(0.5)^{105} + 0.5 \frac{\beta(11 + \tilde{a}_{\beta}, 105 + \tilde{b}_{\beta})}{\beta(\tilde{a}_{\beta}, \tilde{b}_{\beta})},$$

in which $\beta(a, b)$ denotes the beta function and

$$\tilde{a}_{\beta} = \begin{cases} 1.4 - 0.8\beta & 0 \leq \beta \leq 0.5, \\ 0.7 + 0.6(1 - \beta) & 0.5 \leq \beta \leq 1, \end{cases}$$

$$\tilde{b}_{\beta} = \begin{cases} 13 - 6\beta & 0 \leq \beta \leq 0.5, \\ 8 + 4(1 - \beta) & 0.5 \leq \beta \leq 1. \end{cases}$$

Therefore, the posterior probability of \tilde{H}_0 is evaluated as

$$\begin{aligned} \tilde{P}(\tilde{H}_0|y = 11) &= \frac{1}{2} \int_0^1 \int_{0.15-0.05\beta}^{0.05+0.05\beta} (\pi_{\beta}(p|y = 11) + \pi_{1-\beta}(p|y = 11)) dp d\beta \\ &= 0.141. \end{aligned}$$

In addition, since $\tilde{H}_1 = 1 \ominus \tilde{H}_0$, we get $\tilde{P}(\tilde{H}_1|y = 11) = 1 - \tilde{P}(\tilde{H}_0|y = 11) = 0.859$. So, it is observed that $\tilde{P}(\tilde{H}_1|y = 11) \gg \tilde{P}(\tilde{H}_0|y = 11)$ and thus this result has provided decisive evidence that \tilde{H}_1 should be true.

5 Conclusion

There are many situations in applied statistics where we deal with imprecise information. In this paper, the subject of Bayesian test based on fuzzy hypothesis is introduced and discussed. Unlike the classical Bayesian method that usually applies a two stage hierarchical model for priors, the proposed method assumes that the unknown parameters of the prior distribution in the first stage are closed intervals or fuzzy numbers instead of putting a second prior distribution on the prior. Averaging over the interval-valued or fuzzy-valued parameters, we then construct a notion of posterior probability of the fuzzy hypotheses. Numerical results show that we invoked successfully the induced prior distribution with imprecise parameters to overcome such uncertainties which is often arisen in many real applications of Bayesian hypothesis testing. Notably, the proposed method is general which can be applied for other kinds of Bayesian testing methods of fuzzy hypothesis such as Bayes factor. Extending the proposed method with imprecise observations and hypotheses is also a potential subject for future studies.

Acknowledgements

I would like to thank the editor and anonymous referee for his/her constructive comments and suggestions, which helped me improve the manuscript.

References

- Arnold, B. F. (1996), An approach to fuzzy hypothesis testing, *Metrika*. **44**, 119-126.
- Angers, J. F. and Delampady, M. (2008), Fuzzy sets in nonparametric Bayes regression, *Collections*. **3**, 89-104.
- Baldwin, J. F., Lawry, J. and Martin, T. P. (1996), A mass assignment theory of the probability of fuzzy events, *Fuzzy Sets and Systems*. **83**, 353-367.
- Berger, J. O. (1985), *Statistical decision theory and Bayesian analysis*. 2nd ed., Springer-Verlag, New York.
- Bolstad, W. M. (2007), *Introduction to Bayesian statistics*. 2nd ed. Hoboken, NJ; Wiley.
- Casals, M. R. (1993), Bayesian testing of fuzzy parametric hypotheses from fuzzy information, *RAIRO, Operation Reaserch*. **27**, 189-199.
- Casals, M. R. and Gil, P. (1994), Bayesian sequential test for fuzzy parametric hypotheses from fuzzy information, *Information Sciences*, **80**, 283-298.
- Chib, S. and Jacobi, L. (2015), Bayesian fuzzy regression discontinuity analysis and returns to compulsory schooling, *Journal of Applied Econometrics*. DOI: 10.1002/jae.2481.
- Chou, K. and Yuan, J. (1993), Fuzzy Bayesian approach to reliability of existing structures, *Journal of Structural Engineering*, **119**, 3276-3290.
- Dubois, D. and Prade, H. (1998), *Fuzzy sets and systems: theory and applications*, Academic, New York.
- Gholizadeh, R., Mastani Shirazi, A. R. and Sadeghpour Gildeh, B. (2012), Fuzzy Bayesian system reliability assessment based on prior two-parameters exponential distribution under different loss functions, *Software Testing, Verification and Reliability*. **22**, 203-217.
- Heilpern, S. (1993), Fuzzy subsets of the space of probability measures and expected value of fuzzy variable, *Fuzzy Sets and Systems*. **54**, 301-309.

- Klement, E. P. (1982), Some remarks on a paper of R.R. Yager, *Information Sciences*, **27**, 211-220.
- Jeet, K. and Dhir, R. (2012). Bayesian and fuzzy approach to assess and predict the maintainability of software: a comparative study, *International Scholarly Research Network*, 1-10. DOI:10.5402/2012/202980.
- Hesamian, G. and Shams, M. (2016), Parametric testing statistical hypotheses for fuzzy random variables, *Soft Computing*. **20**, 1537–1548.
- Leu, S., Son, P. and Nhung, P. (2015), Hybrid Bayesian Fuzzy-game model for improving the negotiation effectiveness of construction material procurement, *Journal of Computing in Civil Engineering*. **29**, 401-409.
- Marttinen, P., Tang, J., De Baets, B., Dawyndt, P. and Corander, J. (2009), Bayesian clustering of fuzzy feature vectors using a quasi-likelihood approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **31**, 74-85.
- Meeden, G. (2008), Fuzzy Set Representation of a Prior Distribution, in Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh; Bertrand Clarke and Subhashis Ghosal, eds. *Institute of Mathematical Statistics, Collections*. **3**, 82-88.
- Ortega, N. R. S., Massad, E. and Struchiner, C.J. (2008), A Bayesian approach to fuzzy hypotheses testing for the estimation of optimal age for vaccination against measles, *Mathematics and Computers in Simulation*. **79**, 1-13.
- Osoba, O., Mitaim, S. and Kosko, B. (2011), Bayesian inference with adaptive fuzzy priors and likelihoods, *IEEE Transaction on Systems, Man, and Cybernetics-Part B*. **41**, 1183-1197.
- Osoba, O., Mitaim, S. and Kosko, B. (2012), Triply fuzzy function approximation for hierarchical Bayesian inference, *Fuzzy Optimization and Decision Making*, **11**, 241-268.
- Parchami, A., Ivani, R. and Mashinchic, M. (2011), An application of testing fuzzy hypotheses: Soil study on the bio-availability of cadmium, *Scientia Iranica*. **18**, 470-478.
- Plasecki, K. (1985), Probability of fuzzy events defined as denumerable additivity measure, *Fuzzy Sets and Systems*, **17**, 271-287.
- Singpurwalla, N. D. and Booker, J. M. (2004), Membership functions and probability measures of fuzzy sets (with discussion), *Journal of American Statistical Association*. **99**, 867-889.

Smets, P. (1982), Probability of a fuzzy event: An axiomatic approach, *Fuzzy Sets and Systems*. **7**, 153-164.

Stein, W. E. (1985), Fuzzy probability vectors, *Fuzzy Sets and Systems*. **15**, 263-267.

Taheri, S. M. and Behboodan, J. (2001), A Bayesian approach to fuzzy hypotheses testing, *Fuzzy Sets and Systems*. **123**, 39-48.

Toth, H. (1992), Probabilities and fuzzy events: An operational approach, *Fuzzy Sets and Systems*. **48**, 113-127.

Viertl, R. (2011), *Statistical methods for fuzzy data*, John Wiley & Sons, Chichester, Ltd.

Xiao, N. C., Li, Y. F., Wang, Z., Peng, W. and Huang, H. Z. (2013), Bayesian reliability estimation for deteriorating systems with limited samples using the maximum entropy approach, *Entropy*. **15**, 5492-5509.

Yager, R. R. (1984), A representation of the probability of a fuzzy subset, *Fuzzy Sets and Systems*. **13**, 273-283.

Zadeh, L. A. (1965), Fuzzy sets, *Information and Control*. **8**, 338-356.