Kybernetika

Ali Reza Saeidi; Mohammad Ghasem Akbari; Mahdi Doostparast

Hypotheses testing with the two-parameter Pareto distribution on the basis of records in fuzzy environment

Kybernetika, Vol. 50 (2014), No. 5, 744–757

Persistent URL: `http://dml.cz/dmlcz/144104`

KYBERNETIKA — VOLUME *50* (2014), NUMBER 5, PAGES 744–757

HYPOTHESES TESTING WITH THE TWO-PARAMETER PARETO DISTRIBUTION ON THE BASIS OF RECORDS IN FUZZY ENVIRONMENT

Ali Reza Saeidi, Mohammad Ghasem Akbari and Mahdi Doostparast

In problems of testing statistical hypotheses, we may be confronted with fuzzy concepts. There are also situations in which the available data are *record statistics* such as weather and sports. In this paper, we consider the problem of testing fuzzy hypotheses on the basis of records. Pareto distribution is investigated in more details since it is used in applications including economic and life testing analysis. For illustrative proposes, a real data set on annual wage is analyzed using the results obtained.

Keywords: decision analysis, fuzzy hypotheses, pareto distribution, record data, testing hypotheses

Classification: 62F03, 62A86

1. INTRODUCTION

Statistical analysis, in the traditional form, is based on some concepts such as data, random variable, point estimation, hypotheses and parameter. There are many different situations in which the mentioned concepts are imprecise, and the theory of fuzzy sets is a tool for formulation and analysis of imprecise concepts. Therefore, the hypotheses testing with fuzzy data can be important. For more details, see Akbari and Rezaei [2] and Taheri and Behboodian [26].

The problem of statistical inference in fuzzy environments has been developed in different approaches. For example, Delgado et al. [12] considered the problem of fuzzy hypotheses testing with crisp data. Arnold [8] and Arnold [9] presented an approach to test fuzzily formulated hypotheses, in which he considered the fuzzy constraints on Types I and II errors. Holena [19] considered a fuzzy generalization of a sophisticated approach to exploratory data analysis. Holena [20] presented a different approach motivated by the observational logic and its success in automated knowledge discovery. Neyman–Pearson lemma for fuzzy hypotheses testing and Neyman–Pearson lemma for fuzzy hypotheses testing with vague data (and *crisp density function*) were given by Torabi et al. [27] and Taheri and Behboodian [26]. Filzmoser and Viertl [16] present an approach for statistical testing on the basis of fuzzy values by introducing the *fuzzy p−value*. Some

DOI: 10.14736/kyb-2014-5-0744

methods for statistical inference with fuzzy data were reviewed by Viertl [28]. Buckley [10] and Buckley [11] studied the problem of statistical inference in fuzzy environments. Taheri and Arefi [25] exhibited an approach for testing fuzzy hypotheses based on fuzzy test statistics. Parchami et al. [23] considered the problem of testing hypotheses, when the hypotheses are fuzzy based on crisp data. The bootstrap, using fuzzy data, is developed in different approaches. Montenegro et al. [22] have presented an asymptotic one-sample procedure. Körner's asymptotic development [21] concerned general fuzzy random variables. Gonzalez et al. [17] showed that the one-sample method for testing the mean of a fuzzy random variable can be extended to general ones. Akbari and Rezaei [3] described a bootstrap method for computing the variance designed directly for hypothesis testing with fuzzy data based on Yao-Wu signed distance [29]. Akbari et al. [4] studied statistical inference about the variance of fuzzy random variables based on L_2-metric. Akbari and Rezaei [5] investigated bootstrap testing fuzzy hypotheses based on fuzzy statistics. For a nonparametric approach for testing fuzzy hypotheses, see Akbari and Arefi [1].

In this paper, we consider the problem of fuzzy testing hypotheses when the parent population follows the two-parameter Pareto distribution. Therefore the rest of this paper is organized as follows. In Section 2, some preliminaries are given. In Section 3, a new definition for fuzzy probability density function (FPDF) is proposed. Also, some results on the basis of the new FPDF are also proposed in Section 3. A review on record-based Pareto analysis is presented in Section 4. For illustrative purposes, in Section 5, a real data set on annual wage, due to Dyer [15], is analyzed using the results obtained.

2. FUZZY HYPOTHESES

Testing statistical hypotheses is on of the most important topics in statistical inference. A statistical hypothesis is an assertion about the probability distribution of one or more random variable(s). It is usually assumed that the hypotheses for which we plan to test are well-defined. Sometimes, the statisticians use to make decision procedure in an unrealistic manner. For example, suppose that θ is the proportion of a population which have a disease. We take a random sample of the elements and study the sample for having some idea about θ. In crisp hypothesis testing, one may use the hypotheses of the form: $H_0 : \theta = 0.2$ versus $H_1 : \theta \neq 0.2$ or $H_0 : \theta \leq 0.2$ versus $H_0 : \theta > 0.2$. However, it is sometimes of interest to test more realistic hypotheses. For example, here, more realistic expressions about θ may be as: "small", "very small", "large", "approximately 0.2", "essentially larger" and so on. Therefore, more realistic formulations of the hypotheses are H_0: "θ is small" versus H_1: "θ is not small". We call such expressions *fuzzy hypotheses*. For more details, see Taheri and Behboodian [26].

Following Taheri and Behboodian [26], we consider some models, as fuzzy sets of real numbers, for modeling the extended versions of the simple, the one-sided, and the two-sided crisp hypotheses to the fuzzy ones.

Definition 2.1. Let $\theta_0 \in \Theta$ be a real and known constant where Θ is the parameter space.

- The hypothesis "θ is approximately equal to θ_0" is called *fuzzy simple hypothesis* and denoted by $H_0(\theta_0)$. Thus, the hypothesis "θ is not approximately equal to θ_0" is a *two-sided hypothesis*.

- The hypothesis "θ is essentially larger (smaller) than θ_0" is a fuzzy right (left) one-sided hypothesis.

Some fuzzy hypotheses are illustrated in Figure 1. The solid and dotted lines stand for the membership functions of the null and the alternative hypotheses, respectively. In Figure 1 (left), the fuzzy (null) simple hypothesis "θ is approximately equal to θ_0" against the alternative hypothesis "θ is not approximately equal to θ_0" are presented. We also provide membership functions for the composite fuzzy hypotheses. Notice that, the membership function of the hypotheses are determined based on the one's beliefs.

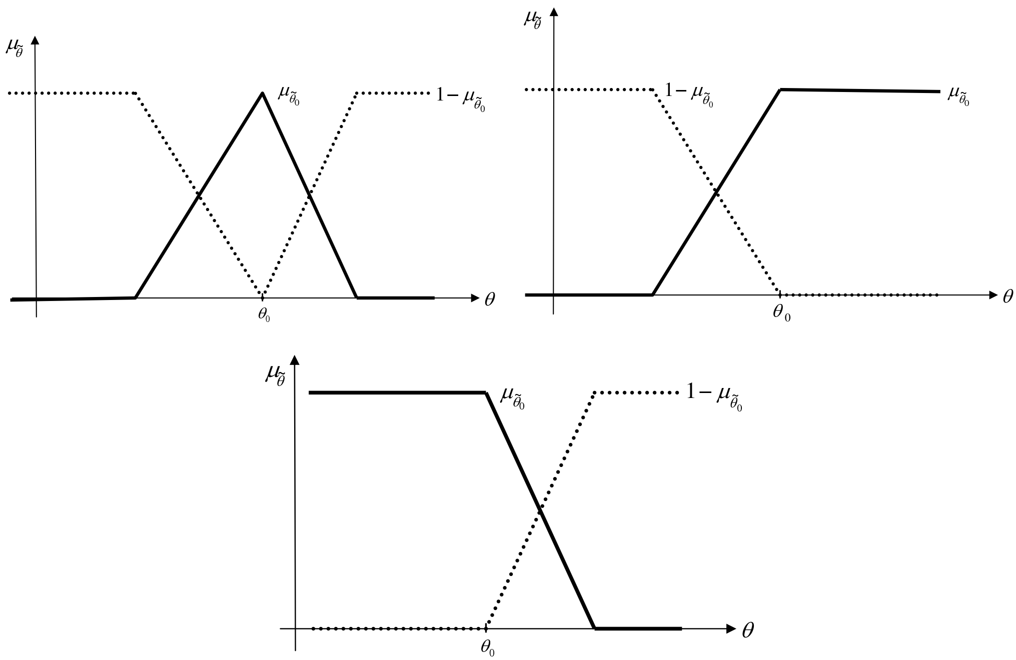

Fig. 1. Various forms of fuzzy hypotheses.

3. A NEW FUZZY DENSITY FUNCTION

Comparing crisp data is well-defined and clear. But comparison of fuzzy data is vague and therefore various methods have been proposed in literature. In the following definition, we provide a non-fuzzy function on the basis of a fuzzy parameter, say $\tilde{\theta}$, in which be a probability density function. To do this, let X be a given random variable and $S_X = \{x \in R : f(x|\theta) > 0,\ \theta \in \Theta\}$ denote the "support" of the random variable X where Θ is the parameter space and $f(x|\theta)$ is the corresponding (non-fuzzy) probability density function. For every $\theta \in \Theta$, we consider a weighted function, say $H(\theta)$, for the probability density function $f(x|\theta)$ and define the desired FPDF, denoted by $f(x|\tilde{\theta})$. As one can see, large values of α makes more concentrate for the proposed FPDF $f(x|\tilde{\theta})$.

In the proposed FPDF, there exists a parameter a which falls in the interval $[0, 1)$. It is easy to verify that $f(x|\tilde{\theta})$ tends to $f(x|\theta)$ as a goes to unity. Now, we give formally the definition.

Definition 3.1. The FPDF of the random variable X is defined by

$$f(x|\widetilde{\theta}) = \frac{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta) f(x|\theta)\,\mathrm{d}\theta \mathrm{d}\alpha}{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta)\,\mathrm{d}\theta \mathrm{d}\alpha}, \qquad a \in [0, 1), \tag{1}$$

where $H(\theta)$ is the membership function of the fuzzy hypothesis and $\widetilde{\theta}_\alpha$ is the corresponding α-cut.

Notice that $f(x|\widetilde{\theta}) \geq 0$ and

$$\begin{aligned} \int_{x \in S_X} f(x|\widetilde{\theta})\,\mathrm{d}x &= \int_{x \in S_X} \frac{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta) f(x|\theta)\,\mathrm{d}\theta \mathrm{d}\alpha}{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta)\,\mathrm{d}\theta \mathrm{d}\alpha} \mathrm{d}x \\ &= \frac{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta) \int_{x \in S_X} f(x|\theta)\,\mathrm{d}x \mathrm{d}\theta \mathrm{d}\alpha}{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta)\,\mathrm{d}\theta \mathrm{d}\alpha} = 1. \end{aligned}$$

The FPDF $f(x|\widetilde{\theta})$ in (1) is sensitive with respect to the parameter a. If the spread of fuzzy parameter of interest is large, then we can improve the probability (density) function based on a. The parameter a is called *credit (control) index*.

Let $g(x) : \mathcal{R} \to \mathcal{R}$ be an arbitrary (crisp) function and the random variable X follows the FPDF (1). Then, the expectation of X is

$$\begin{aligned} E_{\widetilde{\theta}}(g(X)) &= \int_{x \in S_X} g(x) f(x|\widetilde{\theta})\,\mathrm{d}x \\ &= \frac{\int_{x \in S_X} \int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} g(x) H(\theta) f(x|\theta)\,\mathrm{d}\theta \mathrm{d}\alpha \mathrm{d}x}{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta)\,\mathrm{d}\theta \mathrm{d}\alpha} \\ &= \frac{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta) \int_{x \in S_X} g(x) f(x|\theta)\,\mathrm{d}x \mathrm{d}\theta \mathrm{d}\alpha}{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta)\,\mathrm{d}\theta \mathrm{d}\alpha} \\ &= \frac{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta) E_\theta(g(X))\,\mathrm{d}\theta \mathrm{d}\alpha}{\int_a^1 \int_{\theta \in \widetilde{\theta}_\alpha} H(\theta)\,\mathrm{d}\theta \mathrm{d}\alpha}. \end{aligned} \tag{2}$$

In the sequel, we use (2) for definitions of Fuzzy errors.

3.1. Fuzzy errors

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample coming from the FPDF $f(x|\widetilde{\theta})$, defined by (1).

Definition 3.2. Let $\psi(\mathbf{X})$ be a test function for testing $H_0(\theta_0)$ against the alternative $H_1(\theta_1)$. The probability of types I and II errors of $\psi(\mathbf{X})$ are, respectively, $\alpha_\psi = E_{\widetilde{\theta}_0}[\psi(\mathbf{X})]$, and $\beta_\psi = E_{\widetilde{\theta}_1}[1 - \psi(\mathbf{X})]$.

Definition 3.3. A test ψ is said to be a test of level α if $\alpha_\psi \leq \alpha$, where $\alpha \in [0,1]$ and α_ψ is called the size of ψ.

Definition 3.4. A test ψ of level α is said to be a best test of level α, if $\beta_\psi \leq \beta_{\psi^*}$ for all test ψ^* of the level α.

In the sequel, we use a version of the well-known Neyman–Pearson Lemma for testing fuzzy hypotheses based on the FPDF (1). The given procedure in this paper differs from the procedure of Taheri and Behboodian [26]. As we show, the new proposed procedure provides better results. First, we need the following lemma:

Lemma 3.5. Let $\mathbf{X}$ be a random sample arising from the FPDF (1), with the observed value $\mathbf{x} = (x_1, \ldots, x_n)$, and the parameter $\theta \in \Theta$ is unknown. Consider the problem of testing the simple fuzzy hypotheses $H_0 : \theta$ is approximately equal to θ_0 against the alternative $H_1 : \theta$ is approximately equal to θ_1. Then

(i) any test function with the following form, for some $k \geq 0$ and $0 \leq \delta(\mathbf{x}) \leq 1$, is the best test function of its size, say α_ψ,

$$\psi(\mathbf{x}) = \begin{cases} 1 & \frac{f(\mathbf{x}|\widetilde{\theta}_1)}{f(\mathbf{x}|\widetilde{\theta}_0)} > k, \\ \\ \delta(\mathbf{x}) & \frac{f(\mathbf{x}|\widetilde{\theta}_1)}{f(\mathbf{x}|\widetilde{\theta}_0)} = k, \\ \\ 0 & \frac{f(\mathbf{x}|\widetilde{\theta}_1)}{f(\mathbf{x}|\widetilde{\theta}_0)} < k. \end{cases} \tag{3}$$

If $k = \infty$, then the best test function is defined as

$$\psi(\mathbf{x}) = \begin{cases} 1 & f(\mathbf{x}|\widetilde{\theta}_0) = 0, \\ \\ 0 & f(\mathbf{x}|\widetilde{\theta}_0) > 0. \end{cases}$$

(ii) for $0 \leq \alpha \leq 1$, there exist a test function of form (i) with $\delta(\mathbf{x}) = \delta$ (a constant), for which $\alpha_\psi = \alpha$.

Proof. **(i)** Let $\alpha_\psi = \alpha$ and $\psi^*(\mathbf{x})$ be another test function of level α. We prove that $\beta_\psi \leq \beta_{\psi^*}$. To do this, suppose that

$$\begin{aligned} A_1 &= \{\mathbf{x} : f(\mathbf{x}|\widetilde{\theta}_1) - kf(\mathbf{x}|\widetilde{\theta}_0) > 0\}, \\ A_2 &= \{\mathbf{x} : f(\mathbf{x}|\widetilde{\theta}_1) - kf(\mathbf{x}|\widetilde{\theta}_0) = 0\}, \\ A_3 &= \{\mathbf{x} : f(\mathbf{x}|\widetilde{\theta}_1) - kf(\mathbf{x}|\widetilde{\theta}_0) < 0\}. \end{aligned}$$

Now we have

$$\begin{aligned} \int [\psi(\mathbf{x}) - \psi^*(\mathbf{x})](f(\mathbf{x}|\widetilde{\theta}_1) - kf(\mathbf{x}|\widetilde{\theta}_0))\,\mathrm{d}\mathbf{x} &= \int_{A_1} + \int_{A_2} + \int_{A_3} \\ &\geq 0 + 0 + 0 = 0. \end{aligned}$$

Therefore,

$$\begin{aligned}\int [\psi(\mathbf{x}) - \psi^*(\mathbf{x})](f(\mathbf{x}|\widetilde{\theta}_1))\,\mathrm{d}\mathbf{x} &\geq k\int [\psi(\mathbf{x}) - \psi^*(\mathbf{x})]f(\mathbf{x}|\widetilde{\theta}_0))\,\mathrm{d}\mathbf{x}\\ &\geq \alpha - \alpha = 0.\end{aligned}$$

Thus, the left side of the above inequality is positive i. e.,

$$E_{\widetilde{\theta}_1}(\psi(\mathbf{X})) \geq E_{\widetilde{\theta}_1}(\psi^*(\mathbf{X}))$$

or $\beta_\psi \leq \beta_{\psi^*}$. For the case $k = \infty$, we first prove that the size of test function is zero. Let

$$\begin{aligned}A_4 &= \{\mathbf{x} : f(\mathbf{x}|\widetilde{\theta}_0) = 0\},\\ A_5 &= \{\mathbf{x} : f(\mathbf{x}|\widetilde{\theta}_0) > 0\}.\end{aligned}$$

Now, we have

$$\begin{aligned}\alpha_\psi = E_{\widetilde{\theta}_0}(\psi(\mathbf{X})) &= \int \psi(\mathbf{x})(f(\mathbf{x}|\widetilde{\theta}_0))\,\mathrm{d}\mathbf{x}\\ &= \int_{A_4} \psi(\mathbf{x})(f(\mathbf{x}|\widetilde{\theta}_0))\,\mathrm{d}\mathbf{x} + \int_{A_5} \psi(\mathbf{x})(f(\mathbf{x}|\widetilde{\theta}_0))\,\mathrm{d}\mathbf{x}\\ &= 0 + 0 = 0.\end{aligned}$$

It is trivial that, if ψ^* is another test of size zero, it must be zero on A_5. It remains to show $\beta_\psi \leq \beta_{\psi^*}$. For this purpose, we write

$$\begin{aligned}\int [\psi(\mathbf{x}) - \psi^*(\mathbf{x})]f(\mathbf{x}|\widetilde{\theta}_1)\,\mathrm{d}\mathbf{x} &= \int_{A_4} [\psi(\mathbf{x}) - \psi^*(\mathbf{x})]f(\mathbf{x}|\widetilde{\theta}_1)\,\mathrm{d}\mathbf{x}\\ &\qquad + \int_{A_5} [\psi(\mathbf{x}) - \psi^*(\mathbf{x})]f(\mathbf{x}|\widetilde{\theta}_1)\,\mathrm{d}\mathbf{x}\\ &= \int_{A_4} [\psi(\mathbf{x}) - \psi^*(\mathbf{x})]f(\mathbf{x}|\widetilde{\theta}_1)\,\mathrm{d}\mathbf{x} + 0\\ &\geq 0.\end{aligned}$$

Therefore, we have $E_{\widetilde{\theta}_1}(\psi(\mathbf{X})) \geq E_{\widetilde{\theta}_1}(\psi^*(\mathbf{X}))$ or $\beta_\psi \leq \beta_{\psi^*}$ and the desired result follows.

(ii) We restrict ourselves to the case $0 < \alpha \leq 1$, since the best test of size zero is given by (i). The size of a test of the form (3) is α, we have $E_{\widetilde{\theta}_0}(\psi(\mathbf{X})) = \alpha$ or

$$P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} > k\right) + \delta P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} = k\right) = \alpha.$$

Hence,

$$P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} \leq k\right) - \delta P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} = k\right) = 1 - \alpha.$$

It is clear that $P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} \leq k\right)$ is nondecreasing and right continuous in k since it is a cumulative distribution function. Now, if there exists a k_0 such that $P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} \leq k_0\right) = 1-\alpha$, then we take $\delta(\mathbf{x}) = 0$ and $k = k_0$. Otherwise, there exists a k_0 such that

$$P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} < k_0\right) \leq 1-\alpha < P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} \leq k_0\right).$$

Now, we take $k = k_0$ and

$$\delta(\mathbf{x}) = \frac{P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} \leq k_0\right) - (1-\alpha)}{P\left(\frac{f(\mathbf{X}|\widetilde{\theta}_1)}{f(\mathbf{X}|\widetilde{\theta}_0)} = k_0\right)},$$

which satisfies $E_{\widetilde{\theta}_0}(\psi(\mathbf{X})) = \alpha$, and $\delta(\mathbf{x}) \in [0,1]$. Hence, the Proof is complete. The proof for the discrete case of FPDF is similar. □

Lemma 3.6. If H_0 and H_1 are composite crisp hypotheses, i. e. their membership functions are exactly the indicator functions of sets Θ_0 and Θ_1 i. e.,

$$H_0(\theta) = \begin{cases} 1 & \theta \in \Theta_0 \\ 0 & \theta \in \Theta_1, \end{cases} \qquad H_1(\theta) = \begin{cases} 0 & \theta \in \Theta_0 \\ 1 & \theta \in \Theta_1. \end{cases}$$

Now using Lemma 3.5 we reject H_0 if $f(\mathbf{x}|\widetilde{\theta}_1)/f(\mathbf{x}|\widetilde{\theta}_0) > k$ or equivalently

$$\int_{\theta\in\Theta_1} f(\mathbf{x}|\theta)\,\mathrm{d}\theta > k \int_{\theta\in\Theta_0} f(\mathbf{x}|\theta)\,\mathrm{d}\theta.$$

This is the same as rejection of H_0 in a Bayes test with a $0-1$ loss function and an improper prior density function for θ, such as $\pi(\theta) = \text{Constant}$.

4. FUZZY RECORD-BASED PARETO ANALYSIS

In this section, we apply the results obtained in the preceding sections for fuzzy hypotheses testing on the basis of records arising from the Pareto distribution.

4.1. Form of data

Let $X_1, X_2, \ldots$ be a sequence of continuous random variables. X_i is defined to be a lower record value if its value is smaller than all preceding values $X_1, \ldots, X_{i-1}$. By definition, X_1 is a lower record value. An analogous definition can be provided for upper record values. As pointed out by Gulati and Padgett [18], often, in industrial testing, meteorological data, and some other situations, measurements may be made sequentially and only values smaller (or larger) than all previous ones are recorded. Such data may be represented by $(\mathbf{r}, \mathbf{\Delta}) := (r_1, \Delta_1, \ldots, r_m, \Delta_m)$, where r_i is the ith record value, meaning new minimum (or maximum), and Δ_i is the number of trials following the observation of r_i that are needed to obtain a new record value r_{i+1}. There are two sampling schemes for generating such a record-breaking data. Under the inverse sampling scheme, units are taken sequentially and sampling is terminated when the mth minimum is observed. In this case, the total number of units sampled is a random number, and Δ_m is defined to be 1 for convenience, while under the random sampling scheme, a random sample $X_1, \dots, X_n$ is examined sequentially and successive minimum values are recorded. In this setting, we have $N(n)$, the number of records obtained, to be random and, given a value of m for it, we have in this case $\sum_{i=1}^{m} \Delta_i = n$. Interested readers may refer to Arnold et al. [7].

4.2. The family of Pareto distributions

A random variable X is said to have a two-parameter Pareto distribution, denoted by $\mathrm{Par}(\gamma, \beta)$, if its cumulative distribution function (cdf) is

$$F(x; \gamma, \beta) = 1 - \left(\frac{\beta}{x}\right)^{\gamma}, \qquad x \geq \beta > 0,\ \gamma > 0, \tag{4}$$

with the corresponding probability density function (pdf) as

$$f(x; \gamma, \beta) = \frac{\gamma\beta^{\gamma}}{x^{\gamma+1}}, \qquad x \geq \beta > 0,\ \gamma > 0. \tag{5}$$

The family of Pareto distributions has been used commonly to model naturally occurring phenomena in which the distribution of random variables of interest have long tails. For more details, see Arnold [6].

4.3. Point estimation

Let $X_1, X_2, \dots$ be i.i.d. random variables, each drawn from a population with the cdf and the pdf $F(\cdot)$ and $f(\cdot)$, respectively. Then, the likelihood function associated with the record data $(\mathbf{r}, \boldsymbol{\Delta})$ reads

$$L(\mathbf{r}, \boldsymbol{\Delta}) = \prod_{i=1}^{m} f(r_i)[1 - F(r_i)]^{(\Delta_i - 1)} I_{(-\infty, r_{i-1})}(r_i), \tag{6}$$

where $r_0 \equiv \infty$, $\Delta_m \equiv 1$ for inverse sampling and $\Delta_m = n - \sum_{i=1}^{m-1} \Delta_i$ for random sampling schemes, and $I_A(x)$ is the indicator function of the set A; for details, one may refer to Samaniego and Whitaker [24]. Substituting (4) and (5) into (6), we have

$$\begin{aligned} L(\gamma, \beta; \mathbf{r}, \boldsymbol{\Delta}) &= \textstyle\prod_{i=1}^{m} \frac{\gamma}{r_i} \left(\frac{\beta}{r_i}\right)^{\gamma\Delta_i} \\ &= \frac{\gamma^m}{\prod_{i=1}^{m} r_i} \left(\frac{\beta^{T_m}}{\prod_{i=1}^{m} {r_i}^{\Delta_i}}\right)^{\gamma} \\ &= \gamma^m \beta^{\gamma \sum_{i=1}^{m} \Delta_i} \textstyle\prod_{i=1}^{m} r_i^{-(\gamma\Delta_i + 1)}, \quad r_1 > \dots > r_m > \beta > 0,\ \gamma > 0. \end{aligned} \tag{7}$$

From Equation (6), it is apparent that a minimal sufficient statistic for the parameter (γ, β) is $(\sum_{i=1}^{m} \Delta_i \log R_i, T_m, R_m)$ where $T_m = \sum_{i=1}^{m} \Delta_i$. Hence, the log-likelihood function is reduced to

$$l(\gamma, \beta; \mathbf{r}, \mathbf{\Delta}) = \log L(\gamma, \beta; \mathbf{r}, \mathbf{\Delta}) = m \ \log \gamma + \gamma \sum_{i=1}^{m} \Delta_i \log(\frac{\beta}{r_i}) - \sum_{i=1}^{m} \log \ r_i, \tag{8}$$

where (and throughout this article) "log" denotes the natural logarithm. Then, the maximum likelihood estimates ($MLEs$) of the unknown parameters γ and β under both sampling schemes are

$$\widehat{\gamma}_{ML} = \frac{m}{\log \prod_{i=1}^{m} (R_i/R_m)^{\Delta_i}}, \tag{9}$$

and

$$\widehat{\beta}_{ML} = R_m, \tag{10}$$

respectively. Under the inverse random sampling scheme, it can be shown that the expression of denominator of $\widehat{\gamma}_{ML}$ in (9) is distributed as Gamma with shape and scale parameters $m-1$ and γ^{-1}, respectively; see Doostparast et al. [13]. Under the inverse sampling scheme, it can be shown that

$$f_{R_m}(x) = \frac{f(x)[-\log F(x)]^{m-1}}{\Gamma(m)}, \quad 0 < x < \infty, \tag{11}$$

and

$$\sum_{i=1}^{m} \ \Delta_i \ \log(R_i/R_m) \sim \Gamma(m-1, \gamma^{-1}). \tag{12}$$

5. ANNUAL WAGE DATA SET

In this section, we illustrate the proposed approach in the preceding sections for records from the Pareto distribution. Dyer [15] reported an annual wage data set (in multiplies of 100 US dollars) of a random sample of 30 production-line workers in a large industrial firm, as presented in Table 1. He argued that the Pareto distribution provided an adequate fit for these data. The records extracted from the data set are presented in Table 2. The $MLEs$ of γ and β were obtained by Doostparast and Balakrishnan [14] as $\widehat{\gamma}_{ML} = 4.197$ and $\widehat{\beta}_{ML} = 101$. Now, consider the problem of testing the fuzzy simple hypothesis $H_0 : \gamma \simeq 5$ against the alternative $H_1 : \gamma \simeq 7$ with the following membership functions

$$H_0(\gamma) = \begin{cases} \gamma - 4 & 4 < \gamma < 5 \\ \frac{(7-\gamma)}{2} & 5 < \gamma < 7 \end{cases}$$

and

$$H_1(\gamma) = \begin{cases} \gamma - 6 & 6 < \gamma < 7 \\ \frac{(9-\gamma)}{2} & 7 < \gamma < 9. \end{cases}$$

112	154	119	108	112	156	123	103	115	107
125	119	128	132	107	151	103	104	116	140
108	105	158	104	119	111	101	157	112	115

Tab. 1. Annual wage data (in multiplies of 100 U.S. dollars).

i	1	2	3	4
R_i	112	108	103	101
Δ_i	3	4	19	4

Tab. 2. Record data coming from annual wage data.

k	$\alpha_\psi(a=0)$	$\alpha_\psi(a=0.1)$	$\alpha_\psi(a=0.3)$	$\beta_\psi(a=0)$	$\alpha_{\psi_{T-B}}$	$\beta_{\psi_{T-B}}$	α_{exact}	β_{exact}
0	*	*	*	*	*	*	0.000	1.000
0.10	0.017	0.016	0.016	0.962	0.017	0.961	0.014	0.966
0.15	0.046	0.046	0.045	0.902	0.048	0.890	0.041	0.910
0.20	0.090	0.090	0.088	0.821	0.094	0.816	0.080	0.933
0.25	0.147	0.146	0.143	0.727	0.152	0.721	0.131	0.744
0.30	0.211	0.210	0.207	0.630	0.218	0.623	0.191	0.650
0.35	0.279	0.278	0.274	0.535	0.288	0.528	0.256	0.557
0.40	0.350	0.349	0.345	0.447	0.359	0.441	0.323	0.469
0.45	0.419	0.418	0.414	0.369	0.429	0.363	0.391	0.390
0.50	0.486	0.484	0.480	0.301	0.429	0.363	0.456	0.321
0.55	0.549	0.547	0.543	0.242	0.557	0.237	0.519	0.261
0.60	0.606	0.605	0.601	0.194	0.615	0.190	0.577	0.210
0.65	0.659	0.657	0.654	0.154	0.667	0.150	0.630	0.168
0.70	0.706	0.705	0.701	0.121	0.713	0.118	0.679	0.133
0.75	0.748	0.747	0.743	0.095	0.754	0.093	0.723	0.105
0.80	0.785	0.784	0.781	0.074	0.790	0.072	0.762	0.082
0.85	0.817	0.816	0.814	0.057	0.822	0.056	0.796	0.064
0.90	0.845	0.844	0.842	0.044	0.849	0.043	0.826	0.050
0.95	0.869	0.868	0.867	0.034	0.872	0.033	0.853	0.038
1.00	0.889	0.889	0.888	0.026	0.893	0.025	0.875	0.030

Tab. 3. The probabilities of errors of types I and II.

By Lemma 3.6 and the FPDF (1), the most powerful (MP) test of H_0 versus H_1, in the sense of Definition 3.2, is

$$\psi_1^\star(\mathbf{r}, \mathbf{\Delta}) = \begin{cases} 1 & T \leq k \\ 0 & T > k \end{cases} \tag{13}$$

where $T = \sum_{i=1}^{m} \Delta_i \, \log(R_i/R_m)$. Since T has the gamma distribution with shape and scale parameters $m-1=3$ and γ^{-1}, respectively, (Doostparast and Balakrishnan [14]), the power function of the test $\psi_1^\star$ in (13) is

$$\beta_{\psi_1^\star}(\gamma) = P_\alpha(T < k) = 1 - \left[e^{-k\gamma} \left(\frac{k^2\gamma^2}{2} + k\gamma + 1 \right) \right],$$

which is decreasing in γ. For some selected values of k and $a = 0, 0.1, 0.3$, the probabilities of types I and II are obtained and displayed in Table 3.

Similarly, consider $H_0 : \beta \simeq 90$ versus $H_1 : \beta \simeq 100$ with membership functions

$$H_0(\beta) = \begin{cases} \frac{(\beta-85)}{5} & \text{for } 85 < \beta < 90 \\ \frac{(108-\beta)}{18} & \text{for } 90 < \beta < 108 \end{cases}$$

and

$$H_1(\beta) = \begin{cases} \frac{(\beta-95)}{5} & \text{for } 95 < \beta < 100 \\ \frac{(105-\beta)}{5} & \text{for } 100 < \beta < 105. \end{cases}$$

By Lemma 3.6 and the FPDF (1), we reject H_0, if $R_m > k$. Hence, the best test function of H_0 against H_1 is derived as

$$\psi_2^\star(\mathbf{r}, \mathbf{\Delta}) = \begin{cases} 1 & R_m > k, \\ 0 & R_m \leq k. \end{cases} \tag{14}$$

From Doostparast and Balakrishnan [14], the pdf of R_4 is

$$f_{R_4}(x) = \frac{\gamma \beta^\gamma [-\log(1 - (\frac{\beta}{x})^\gamma)]^3}{\Gamma(4) x^{\gamma+1}}. \tag{15}$$

So, the power function of the test (14) is obtained from (15) as

$$\beta_{\psi_2^\star}(\beta) = \int_0^{\frac{\beta}{k}} \frac{\gamma \left[-\log(1-u^\gamma)\right]^3 u^{\gamma-1}}{6} \, \mathrm{d}u. \tag{16}$$

For $\gamma = 4$ in (16), and the identity $-\log(1-a) = \sum_{k=0}^{\infty} a^{k+1}/(k+1)$, for $|a| < 1$, we conclude that

$$\beta_\psi(\beta) = \int_0^{\frac{\delta}{k}} \frac{2}{3} \left[u^3 + \frac{u^6}{2} + \frac{u^9}{3} + \frac{u^{12}}{4} + \frac{u^{15}}{5} + \cdots \right]^3 u^3 \, \mathrm{d}u.$$

For some selected values of k and $a = 0$, the probabilities of types I and II errors are given in Tables 4 and 5, respectively. From Tables 3−5, one can see that $\alpha_{\psi_{\text{new}}}$ and $\beta_{\psi_{\text{new}}}$ are closer to $\alpha_{\psi_{\text{exact}}}$ and $\beta_{\psi_{\text{exact}}}$, respectively, with respect to the method of Taheri and Behboodian [26], denoted by $\alpha_{\psi_{T-B}}$ and $\beta_{\psi_{T-B}}$.

k	$\alpha_{\psi_{new}}(a=0)$	$\alpha_{\psi_{T-B}}$	$\alpha_{\psi_{\text{exact}}}$
95	0.603	0.613	0.592
97	0.426	0.335	0.306
99	0.182	0.189	0.173
100	0.139	0.141	0.134
102	0.092	0.117	0.081
105	0.000	0.000	0.000

Tab. 4. The probability of type I error for the annual wage data set.

k	$\beta_{\psi_{\text{new}}}(a=0)$	$\beta_{\psi_{T-B}}$	$\beta_{\psi_{\text{exact}}}$
105	0.268	0.253	0.287
107	0.583	0.571	0.619
109	0.763	0.759	0.778
110	0.811	0.803	0.827
112	0.873	0.866	0.891
115	0.933	0.928	0.941

Tab. 5. The probability of type II error for the annual wage data set.

6. CONCLUSIONS

In the real world, some parameters of interest may be imprecise. A method to formulate these parameters is provided by fuzzy environments. In this paper, we considered the problem of testing hypotheses with fuzzy parameters on the basis of record data coming from the two-parameter Pareto distribution. For introducing a test function, we defined a new FPDF based on the parameters and then the Neyman–Pearson Lemma was applied based on such a density. Since, the proposed approach in this paper is different from Taheri and Behboodian's approach, we compare these two methods with the type I and type II errors of crisp (precise) case for a real data set. Based on the results as shown in Table 3, the type I error (type II error) of the new method is smaller (larger) than method of Taheri and Behboodian [26] and is near to crisp one. Notice that one sample does not tell us so much and more investigations must be carried out.

ACKNOWLEDGEMENT

The authors are grateful to the Editor-in-Chief, the Associate Editor, and the anonymous referees for their useful suggestions and comments on an earlier version of this paper, which resulted in this improved version of the manuscript.

(Received June 15, 2013)

REFERENCES

[1] M. G. Akbari and M. Arefi: Statistical nonparametric test based on the intuitionistic fuzzy data. J. Intell. Fuzzy Systems *25* (2013), 525–534.

[2] M. G. Akbari and A. Rezaei: An uniformly minimum variance unbiased point estimator using fuzzy observations. Austrian J. Statist. *36* (2007), 307–317.

[3] M. G. Akbari and A. Rezaei: Bootstrap statistical inference for the variance based on fuzzy data. Austrian J. Statist. *38* (2009), 121–130.

[4] M. G. Akbari, A. Rezaei, and Y. Waghei: Statistical inference about the variance of fuzzy random variables. Sankhya *71-B* (2009), 1–15.

[5] M. G. Akbari and A. Rezaei: Bootstrap testing fuzzy hypotheses and observations on fuzzy statistic. Expert Systems Appl. *37* (2010), 8, 5782–5787.

[6] B. C. Arnold: Pareto Distributions. International Co-operative Publishing House, Fairland 1983.

[7] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja: Records. John Wiley and Sons, New York 1998.

[8] B. F. Arnold: An approach to fuzzy hypothesis testing. Metrika *44* (1996), 119–126.

[9] B. F. Arnold: Testing fuzzy hypotheses with crisp data. Fuzzy Sets and Systems *94* (1998), 323–333.

[10] J. J. Buckley: Fuzzy Probabilities: New Approach and Applications. Springer-Verlag, Berlin, Heidelberg 2005.

[11] J. J. Buckley: Fuzzy Probability and Statistics. Springer-Verlag, Berlin, Heidelberg 2006.

[12] M. Delgado, J. L. Verdegay, and M. A. Vila: Testing fuzzy hypotheses, a Bayesian approach. In: Approximate Reasoning in Expert Systems (M. M. Gupta, ed.) 1985, pp. 307–316.

[13] M. Doostparast, M. G. Akbari, and N. Balakrishnan: Bayesian analysis for the two-parameter Pareto distribution based on record values and times. J. Stat. Comput. Simul. *81* (2011), 11, 1393–1403.

[14] M. Doostparast and N. Balakrishnan: Pareto record-based analysis. Statistics *47* (2013), 5, 1075–1089.

[15] D. Dyer: Structural probability bounds for the strong Pareto laws. Canad. J. Statist. *9* (1981), 71–77.

[16] P. Filzmoser and R. Viertl: Testing hypotheses with fuzzy data: the fuzzy $p-$value. Metrika *59* (2004), 21–29.

[17] G. Gonzalez-Rodriguez, M. Montenegro, A. Colubi, and M. A. Gil: Bootstrap techniques and fuzzy random variables: Synergy in hypothesis testing with fuzzy data. Fuzzy Sets and Systems *157* (2006), 2608–2613.

[18] S. Gulati and W. J. Padgett: Smooth nonparametric estimation of the distribution and density functions from record-breaking data. Commun. Commun. Statist. – Theory and Methods *23* (1994), 1259–1274.

[19] M. Holena: Fuzzy hypotheses for GUHA implications. Fuzzy Sets and Systems *98* (1998), 101–125.

[20] M. Holena: Fuzzy hypotheses testing in the framework of fuzzy logic. Fuzzy Sets and Systems *145* (2004), 229–252.

[21] R. Körner: An asymptotic $\alpha-$cut for the expectation of random fuzzy variables. J. Statist. Plann. Inferences *83* (2000), 331–346.

[22] M. Montenegro, A. Colubi, M. R. Casals, and M. A. Gil: Asymptotic and bootstrap techniques for testing the expected value of a fuzzy random variable. Metrika *59* (2004), 31–49.

[23] A. Parchami, S. M. Taheri, and M. Mashinchi: Fuzzy p-value in testing fuzzy hypotheses with crisp data. Statist. Papers *51* (2010), 1, 209–226.

[24] F. J. Samaniego and L. R. Whitaker: On estimating population characteristics from record-breaking observations II: NonParametric results. Naval Res. Logist. *35* (1988), 221–236.

[25] S. M. Taheri and M. Arefi: Testing fuzzy hypotheses based on fuzzy statistics. Soft Computing *13* (2009), 617–625.

[26] S. M. Taheri and J. Behboodian: Neyman–Pearson lemma for fuzzy hypotheses testing. Metrika *49* (1999), 3–17.

[27] H. Torabi, J. Behboodian, and S. M. Taheri: Neyman-Pearson Lemma for fuzzy hypotheses testing with vague data. Metrika *64* (2006), 289–304.

[28] R. Viertl: Univariate statistical analysis with fuzzy data. Comput. Statist. Data Anal. *51* (2006), 133–147.

[29] J. S. Yao and K. Wu: Ranking fuzzy numbers based on decomposition principle and signed distance. Fuzzy Sets and Systems *11* (2000), 275–288.

Ali Reza Saeidi, Department of Statistics, School of Mathematical Sciences, University of Birjand, Southern Khorasan. Iran.
e-mail: alirezasaeidi.s@gmail.com

Mohammad Ghasem Akbari, Department of Statistics, School of Mathematical Sciences, University of Birjand, Southern Khorasan. Iran.
e-mail: g_z_akbari@yahoo.com

Mahdi Doostparast, Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad. Iran.
e-mail: doustparast@um.ac.ir