Sequential Test of Fuzzy Hypotheses

Mohammad Ghasem Akbari
Department of Statistics, Faculty of Sciences, University of Birjand, Southern Khorasan, Iran
E-mail: g_z_akbari@yahoo.com
Received May 22, 2011; revised June 10, 2011; accepted June 17, 2011

Abstract

In testing statistical hypotheses, as in other statistical problems, we may be confronted with fuzzy concepts. This paper deals with the problem of testing hypotheses, when the hypotheses are fuzzy and the data are crisp. We first give new definitions for notion of mass (density) probability function with fuzzy parameter, probability of type I and type II errors and then state and prove the sequential probability ratio test, on the basis of these new errors, for testing fuzzy hypotheses. Numerical examples are also provided to illustrate the approach.

Keywords: Canonical Fuzzy Number, Fuzzy Hypotheses, Type I and II Error Sizes, Sequential Probability Ratio Test

1. Introduction

Statistical analysis, in traditional form, is based on crispness of data, random variable, point estimation, hypotheses, parameter and so on. As there are many different situations in which the above mentioned concepts are imprecise. On the other hand, the theory of fuzzy sets is a well known tool for formulation and analysis of imprecise and subjective concepts. Therefore the sequential probability ratio test with fuzzy hypotheses can be important. The problem of statistical inference in fuzzy environments are developed in different approaches.


Thompson and Geyer [12] proposed the Fuzzy p-values in latent variable problems. Taheri and Arefi [13] exhibit an approach for testing fuzzy hypotheses based on fuzzy test statistics. Parchami et al. [14] consider the problem of testing hypotheses, when the hypotheses are fuzzy and the data are crisp. they first introduce the notion of fuzzy p-value, by applying the extension principle and then present an approach for testing fuzzy hypotheses by comparing a fuzzy p-value and a fuzzy significance level, based on a comparison of two fuzzy sets.

In present work, we first define a new approach for obtaining the probability (density) function, when the random variable is crisp and the parameter of interest is imprecise (fuzzy). Also, the type I and type II errors are introduced based on fuzzy hypotheses. Then, the sequential probability ratio test (SPRT) is defined and extended based on such hypotheses.

We organize the matter in the following way:

In section 2 we describe some basic concepts of fuzzy hypotheses, density (Mass) probability function with fuzzy parameter and necessary definitions. In section 3 we come up sequential probability ratio test based on fuzzy hypotheses. In section 4 the previous definitions and the sequential probability ratio test will be illustrated by examples.

2. Preliminaries

In this section we describe fuzzy hypotheses, density (Mass) probability function with fuzzy parameter and
necessary definitions.

Let \((\Omega,\mathcal{F},P)\) be a probability space, a random variable (RV) \(X\) is a measurable function from \((\Omega,\mathcal{F},P)\) to \((\mathcal{X},\mathcal{B},P_X)\), where \(P_X\) is the probability measure induced by \(X\) and is called the distribution of the RV \(X\), i.e.,

\[
P_X(A) = P(X \in A) = \int_{\mathcal{X}} dP \quad \forall A \in \mathcal{B}.
\]

If \(P_X\) is dominated by a \(\sigma\)-finite measure \(\nu\), i.e. \(P_X \ll \nu\) then by the Radon-Nikodym theorem (Billingsley, [15]), we have

\[
P_X(A) = \int_{X \in \mathcal{X}} f(x|\theta)d\nu(x) \quad \forall A \in \mathcal{B},
\]

where \(f(x|\theta)\) is the Radon-Nikodym derivative of \(P_X\) with respect to \(\nu\) and is called the probability density function of \(X\) with respect to \(\nu\). In a statistical context, the measure \(\nu\) is usually a “counting measure” or “Lebesgue measure”, hence \(P_X(A)\) is \(\sum_{x \in \mathcal{X}} P(X = x|\theta)\) or \(\int f(x|\theta)dx\), respectively.

2.1. Canonical Fuzzy Numbers

Let \(\Theta = \{\theta | f(x|\theta) > 0\}\) be the “support” or “sample space” of \(\Theta\), then a fuzzy subset \(\tilde{\theta}\) of \(S_X\) is defined by its membership function \(\mu_{\tilde{\theta}}: \Theta \rightarrow [0,1]\). We denote by \(\tilde{\theta}_\alpha = \{\theta: \mu_{\tilde{\theta}}(x) \geq \alpha\}\) the \(\alpha\)-cut set of \(\tilde{\theta}\) and \(\tilde{\theta}_1\) is the closure of the set \(\{\theta: \mu_{\tilde{\theta}}(x) > 0\}\), and

1) \(\tilde{\theta}\) is called a normal fuzzy set if there exists \(\theta \in \Theta\) such that \(\mu_{\tilde{\theta}}(x) = 1\);

2) \(\tilde{\theta}\) is called a convex fuzzy set if

\[
\mu_{\tilde{\theta}}(\lambda x + (1-\lambda) y) \geq \min\{\mu_{\tilde{\theta}}(x), \mu_{\tilde{\theta}}(y)\} \quad \text{for all} \quad \lambda \in [0,1];
\]

3) \(\tilde{\theta}\) is called a fuzzy number if \(\tilde{\theta}\) is a normal convex fuzzy set and its \(\alpha\)-cut sets, are bounded \(\forall \alpha > 0\);

4) \(\tilde{\theta}\) is called a closed fuzzy number if \(\hat{\theta}\) is a fuzzy number and its membership function \(\mu_{\tilde{\theta}}\) is upper semicontinuous;

5) \(\tilde{\theta}\) is called a bounded fuzzy number if \(\tilde{\theta}\) is a fuzzy number and the support of its membership function \(\mu_{\tilde{\theta}}\) is compact.

If \(\tilde{\theta}\) is a closed and bounded fuzzy number with \(\theta'_\alpha = \inf\{\theta: \theta \in \tilde{\theta}_\alpha\}\) and \(\theta''_\alpha = \sup\{\theta: \theta \in \tilde{\theta}_\alpha\}\) and its membership function be strictly increasing on the interval \(\left[\theta'_\alpha, \theta''_\alpha\right]\) and strictly decreasing on the interval \(\left[\theta''_\alpha, \theta'_\alpha\right]\), then \(\tilde{\theta}\) is called a canonical fuzzy number (Klir and Yuan, [16]).

The fuzzy canonical numbers (such as triangular or trapezoidal fuzzy numbers) are very realistic in fuzzy set theory, so we use this numbers for our goal.

2.2. Fuzzy Hypotheses

We define some models, as fuzzy sets of real numbers, for modeling the extended versions of the simple, the one-sided, and the two-sided ordinary (crisp) hypotheses to the fuzzy ones.

Testing statistical hypothesis is a main branch of statistical inference. Typically, a statistical hypothesis is an assertion about the probability distribution of one or more random variable(s). Traditionally, all statisticians assume the hypothesis for which we wish provide a test are well-defined. This limitation, sometimes, force the statistician to make decision procedure in an unrealistic manner. This is because in realistic problems, we may come across non-precise (fuzzy) hypothesis. For example, suppose that \(\theta\) is the proportion of a population which have a disease. We take a random sample of elements and study the sample for having some idea about \(\theta\). In crisp hypothesis testing, one uses 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 \(H_1: \theta > 0.2\), and so on. However, we would sometimes like to test more realistic hypotheses. In this example, more realistic expressions about \(\theta\) would be considered as: “small”, “very small”, “large”, “approximately 0.2”, “essentially larger” and so on. Therefore, more realistic formulation of the hypotheses might be \(H_0: \theta\) is small, versus \(H_1: \theta\) is not small. We call such expressions as fuzzy hypotheses.

We define 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 (Akbari and Rezaei, [17]).

**Definition 2.1** Let \(\theta_0\) be a real number and known.

1) Any hypothesis of the form \((H : \theta.3cmis.3mapproximately.3cm\theta_0)\) is called to be a fuzzy simple hypothesis.

2) Any hypothesis of the form \((H : \theta.3cmis.15cmnot.3mapproximately.3cm\theta_0)\) is called to be a fuzzy two-sided hypothesis.

3) Any hypothesis of the form \((H : \theta.3cmis.15cmessentially.5cmlarger.15cmthan.15cm.13cm\theta_0)\) is called to be a fuzzy right one-sided hypothesis.

4) Any hypothesis of the form \((H : \theta.3cmis.15cmessentially.15cmsgsmaller.15cmthan.15cm.3cm\theta_0)\) is called to be a fuzzy left one-sided hypothesis.

We denote the above definitions by

\[
(a).3cm\{H _0 : \theta.3cmis.3mapproximately \theta_0 + 0.06\}cm\(\theta_0 + 0.14\}
\]

\[
(b).3cm\{H_0 : \theta.3cmis.3mnessentially \theta_0 + 0.06\}cm\theta_0 + 0.14\}
\]
(c) $3cm \{H_0 : \theta.3cm\text{is}3cm\text{essentially}3cm\text{smaller}\ 3cm\text{than}3cm\theta_0\text{1.6cm}\}$

(d) $3cm \{H_0 : \theta.3cm\text{is}3cm\text{approximately}3cm\theta_0\text{65cm}\}$

2.3. Density (Mass) Probability Function

Let $X$ be a RV and let $S_x = \{x \in R : f(x|\theta) > 0\}$ be the “support” or “sample” space of $X$ and

$$f(x|\theta) = \int_{s_X} H(\theta)f(x|\theta)d\theta d\alpha,$$

where $H(\theta)$ is the membership function of canonical fuzzy hypothesis and $\theta_0$ is its $\alpha$-cuts.

We call the new density $f(x|\theta)$ as the fuzzy probability density (mass) function (FPDF) of $X$ (Akbari and Rezaei [18]). We note that, $f(x|\theta) \geq 0$ and

$$\int_{s_X} f(x|\theta) dx = \int_{s_X} \int_{\theta_0}^\infty H(\theta)f(x|\theta)d\theta d\alpha$$

$$= \int_0^\infty \int_{\theta_0}^\infty H(\theta)f(x|\theta) d\theta d\alpha$$

$$= 1$$

(substitute the summation by integral in discrete cases).

Let $g(x): R \rightarrow R$ be arbitrary function in $x$. Then we define

$$E_\theta(g(X)) = \int_{s_X} g(x)f(x|\theta) dx$$

$$= \int_{s_X} \int_{\theta_0}^\infty g(x) H(\theta)f(x|\theta)d\theta d\alpha dx$$

$$= \int_0^\infty \int_{\theta_0}^\infty H(\theta)g(x)f(x|\theta) d\theta d\alpha$$

$$= \int_0^\infty \int_{\theta_0}^\infty H(\theta)E_\theta(g(X)) d\theta d\alpha$$

$$= \int_0^\infty \int_{\theta_0}^\infty H(\theta) \frac{E_\theta(g(X))}{\theta_0} d\theta d\alpha$$

Let $X = (X_1, X_2, \ldots, X_n)$ be a random sample, with observed value $x = (x_1, x_2, \ldots, x_n)$, where $X_i$ has the FPDF $f(x|\theta)$ with unknown $\theta \in \Theta$. For testing

$$H_0 : \theta \text{ is } H_0(\theta) \text{ or } \theta \sim \theta_0$$

$$H_1 : \theta \text{ is } H_1(\theta) \text{ or } \theta \sim \theta_1$$

we state the following definitions:

**Definition 2.2** Let $\psi(X)$ be a test function. The probability of type I error of $\psi(X)$ is

$$\alpha_\psi = E_{\theta_0}[\psi(X)],$$

and the probability of type II error of $\psi(X)$ is

$$\beta_\psi = 1 - E_{\theta}[\psi(X)] = E_{H_0}[1 - \psi(X)].$$

**Definition 2.3** A test is said to be a test of level $\alpha$ if $\alpha_\psi \leq \alpha$, where $\alpha \in [0,1]$.

we call $\alpha_\psi$ the size of $\psi$.

3. Sequential Probability Ratio Test

Consider testing a null fuzzy hypothesis against an alternative fuzzy hypothesis. In other words, suppose a sample can be drawn from one of two FPDFs and it is desired to test that the sample came from one distribution against the possibility that it came from the other. If $X_1, X_2, \ldots$ denotes the random variables, we want to test $H_0 : X_i \sim f(\theta_0)$ versus $H_1 : X_i \sim f(\theta_1)$. The simple likelihood-ratio test was of the following form:

$$\text{reject } H_0 \text{ if } \frac{L_0}{L_1} \leq k \text{ for some constant } k > 0.$$
A_n = \{ x : k_0 < \lambda(x_1, x_2, \ldots, x_j) \leq k_1, \}
\quad j = 1, 2, \ldots, n-1, \lambda(x_1, x_2, \ldots, x_n) \geq k_i \}

When we considered the simple likelihood-ratio test for fixed sample size \( n \), we determined \( k \) so that the test would have preassigned size \( \alpha \). We want to determine \( k_0 \) and \( k_1 \) so that the sequential probability ratio test will have preassigned \( \alpha \) and \( \beta \) for its respective sizes of type I and type II errors. Note that
\[
\alpha = P(\text{reject } H_0 | H_0 \text{ is true}) = \sum_{n=1}^{\infty} L_0(n)
\]
and
\[
\beta = P(\text{accept } H_0 | H_0 \text{ is false}) = \sum_{n=1}^{\infty} L_1(n),
\]
where, as before, \( \int C_n L_0(n) \) is a shortened notation for \( \int \cdots \int \left[ \prod_{i=1}^{n} f(x_i | \theta_0) \right] dx_i \).

For fixed \( \alpha \) and \( \beta \), the above equations are two equations in the two unknown \( k_0 \) and \( k_1 \). A solution of these two equations would give the sequential probability ratio test having the desired preassigned error sizes \( \alpha \) and \( \beta \). As might be anticipated, the actual determination of \( k_0 \) and \( k_1 \) from above equations can be a major computational project.

We note that the sample size of a sequential probability ratio test is a random variable. The procedure says to continue sampling until \( \lambda_0 \) first falls outside the interval \((k_0, k_1)\). The actual sample size then depend on which \( X_i \)'s observed; it is a function of the random variables \( X_1, X_2, \ldots \) and consequently is itself a RV. Denote it by \( N \). Ideally, we would like to know the distribution of \( N \) or at least the expectation of \( N \). One way of assessing the performance of the sequential probability ratio test would be to evaluate the expected sample size that is required under each hypothesis. The following lemma, given without proof (Lehmann, [20]), state that the sequential probability ratio test with crisp hypotheses is an optimal test if performance is measured using expected sample size. We can similarly prove this lemma with fuzzy hypotheses based on introduced FDPF.

**Lemma 3.1** The sequential probability ratio test with error sizes \( \alpha \) and \( \beta \) minimizes both
\[
E(N | H_0 \text{ is true}) \quad \text{and} \quad E(N | H_1 \text{ is true})
\]
among all tests which satisfy the following:
\[
P(H_0 \text{ is rejected } | H_0 \text{ is true}) \leq \alpha,
\]
\[
P(H_0 \text{ is rejected } | H_0 \text{ is true}) \leq \beta,
\]
and the expected sample size is finite.

We noted above that the determination of \( k_0 \) and \( k_1 \) that defines that particular sequential probability ratio test which has error sizes \( \alpha \) and \( \beta \) is in general computationally quite difficult. The following lemma (with simple proof) gives an approximation to \( k_0 \) and \( k_1 \).

**Lemma 3.2** Let \( k_0 \) and \( k_1 \) be defined so that the sequential probability ratio test corresponding to \( k_0 \) and \( k_1 \) has error sizes \( \alpha \) and \( \beta \); then \( k_0 \) and \( k_1 \) can be approximated by, say \( k'_0 \) and \( k'_1 \), where
\[
k'_0 = \frac{\alpha}{1 - \beta} \quad \text{and} \quad k'_1 = \frac{1 - \alpha}{\beta}.
\]

**Lemma 3.3** Let \( \alpha' \) and \( \beta' \) be the error sizes of the sequential probability ratio test defined by \( k'_0 \) and \( k'_1 \) given in before lemma. Then \( \alpha' + \beta' \leq \alpha + \beta \).

The procedure used in performing a sequential probability ratio test is to continue sampling as long as \( k_0 < \lambda_m < k_1 \) and stop sampling as soon as \( \lambda_m \leq k_0 \) or \( < \lambda_m \geq k_1 \). If \( \lambda_m = \sum_{i=1}^{N} z_i \leq \sum_{i=1}^{n} \hat{\theta} \), an equivalent test
\[
\sum_{i=1}^{Z_i} z_i = \sum_{i=1}^{n} \hat{\theta}
\]
given by the following: continue sampling as long as \( \sum_{i=1}^{n} z_i < \sum_{i=1}^{n} \hat{\theta} \) and stop sampling as soon as \( \sum_{i=1}^{n} z_i \geq \sum_{i=1}^{n} \hat{\theta} \). As before, let \( N \) be a RV denoting the sample size of the sequential probability ratio test, and let \( z_i = \ln \left[ \frac{f(x_i | \hat{\theta})}{f(x_i | \theta)} \right] \).

If the sequential probability ratio test leads to rejection of \( H_0 \), then the RV \( \sum_{i=1}^{n} Z_i \geq \ln k_0 \), but \( \sum_{i=1}^{n} Z_i \) is close to \( \ln k_0 \) since \( \sum_{i=1}^{n} Z_i \) first became less than or equal to \( \ln k_0 \) at the \( N \) th observation; hence
\[
E\left[ \sum_{i=1}^{n} Z_i \right] \approx E\left[ \sum_{i=1}^{n} Z_i \right] \approx \ln k_0 \quad \text{Similarly} \quad E\left[ \sum_{i=1}^{n} Z_i \right] \approx \ln k_0 \quad \text{hence} \quad E\left[ \sum_{i=1}^{n} Z_i \right] \approx \xi \ln k_0 + (1 - \xi) \ln k_i \quad \text{where} \quad \xi = P(H_0 \text{ is rejected})
\]

Using Wald's equation (Casella and Berger, [19])
\[
E(N) = \frac{E\left[ \sum_{i=1}^{n} z_i \right]}{E(Z_i)} \approx \frac{\xi \ln k_0 + (1 - \xi) \ln k_i}{E(Z_i)}
\]
we obtain

\[
E \left( N | H_0 \text{ is true} \right) \approx \frac{\alpha \ln k_0 + (1 - \alpha) \ln k_1}{E \left( Z | N | H_0 \text{ is true} \right)}
\]

\[
\alpha \ln \left( \frac{\alpha}{1 - \beta} \right) + (1 - \alpha) \ln \left( \frac{1 - \alpha}{\beta} \right)
\]

and

\[
E \left( N | H_0 \text{ is flase} \right) \approx \frac{(1 - \beta) \ln k_0 + \beta \ln k_1}{E \left( Z | H_0 \text{ is flase} \right)}
\]

\[
(1 - \beta) \ln \left( \frac{\alpha}{1 - \beta} \right) + \beta \ln \left( \frac{1 - \alpha}{\beta} \right)
\]

4. Numerical Examples

In this section, we illustrate the proposed approach for some distributions and use the ability of package “Maple 6” [21] for this examples.

Example 4.1 (Taheri and Behboodian, [6]) Let \( X \) be a continues r.v. with PDF

\[
f(x|\theta) = 2\theta x + 2(1 - \theta)(1 - x) \quad 0 < x < 1 \quad 0 < \theta < 1.
\]

we want to test

\[
\begin{align*}
H_0 : \theta & \text{ is approximately } \frac{1}{3} \\
H_1 : \theta & \text{ is approximately } \frac{1}{2}
\end{align*}
\]

where the membership functions \( H_0 \) and \( H_1 \) are defined in the following way:

\[
H_0(\theta) = \begin{cases} 
3\theta & 0 \leq \theta < \frac{1}{3} \\
2 - 3\theta & \frac{1}{3} \leq \theta < \frac{2}{3} \\
0 & \text{otherwise}
\end{cases}
\]

\[
H_1(\theta) = \begin{cases} 
2\theta & 0 \leq \theta < \frac{1}{2} \\
2 - 2\theta & \frac{1}{2} \leq \theta < 1 \\
0 & \text{otherwise}
\end{cases}
\]

We can interpret \( \hat{\theta}_0 \) and \( \hat{\theta}_1 \) as the value of “near to \( \frac{1}{3} \)” and “near to \( \frac{1}{2} \)”.

Let \( \alpha = 0.05 \) and \( \beta = 0.01 \). We obtain

\[
\ln(k_0') = -2.986, \quad \ln(k_1') = 4.554.
\]

Hence,

\[
E \left( N | H_0 \text{ is true} \right) = 38.556, \quad \text{and we must take } n = 39,
\]

whereas \( E \left( N | H_1 \text{ is true} \right) = 25.432 \), thus we take \( n = 25 \).

Example 4.2 Let \( X = (X_1, X_2, \ldots, X_n) \) be a random sample where \( X_i \sim (\theta, 1) \) population, i.e.,

\[
f(x|\theta) = \frac{1}{\sqrt{2\pi}} \exp \left\{ \frac{1}{2}(x - \theta)^2 \right\} \quad x, \theta \in R
\]

and \( H_i(\theta) \) s are our fuzzy hypotheseos with membership functions given by:

\[
H_0(\theta) = \begin{cases} 
\theta - 11 & 11 \leq \theta < 12 \\
13 - \theta & 12 \leq \theta < 13 \\
0 & \text{otherwise}
\end{cases}
\]

\[
H_1(\theta) = \begin{cases} 
\theta - 9 & 9 \leq \theta < 10 \\
11 - \theta & 10 \leq \theta < 11 \\
0 & \text{otherwise}
\end{cases}
\]

We can interpret \( H_0 \) and \( H_1 \) as the value of “near to 12” and “near to 10”.

Let \( \alpha = \beta = 0.05 \). Hence, \( E \left( N | H_0 \text{ is true} \right) = 85.155 \), and we must take \( n = 85 \), whereas \( E \left( N | H_1 \text{ is true} \right) = 87.63 \), thus we take \( n = 88 \).

Example 4.3 Let \( X \) be a random sample where \( X_i \sim E(\theta, 1) \) population, i.e.,

\[
f(x|\theta) = \exp((-x - \theta)) \quad x > \theta, \theta > 0,
\]

and \( H_i(\theta) \) s are our triangular fuzzy parameters with membership functions

\[
H_i(\theta) = \begin{cases} 
\theta - a_i + a & a_i - a \leq \theta \leq a_i \\
\frac{a_i + b - \theta}{b} & a_i \leq \theta \leq a_i + b,
\end{cases}
\]

for \( a, b \geq 0 \).

We can interpret the canonical parameters as having values that are “near to \( a_i \)”.

Let \( a_0 = 8, \quad a_i = 4, \quad \alpha = 0.05 \) and \( \beta = 0.1 \). Hence,

\[
E \left( N | H_0 \text{ is true} \right) = 21.325, \quad \text{and we must take } n = 21,
\]

whereas \( E \left( N | H_1 \text{ is true} \right) = 28.22 \), thus we take

\[
n = 28.
\]

Example 4.4 Let \( X \) be a RV from the \( U \sim (0, \theta) \) population, i.e.,

\[
f(x|\theta) = \frac{1}{\theta} \quad 0 \leq x \leq \theta, \quad \theta > 0,
\]

and \( H_i(\theta) \) s are our trapezoidal fuzzy parameters with membership functions given by:
for $a,b \geq 0$.

Let $a = b = 1$, $x_0 = 5$ and $x_i = 3$. If $\alpha = \beta = 0.05$, then $E(N|H_0 \text{ is true}) = 8.75$, and we must take $n = 9$, whereas $E(N|H_i \text{ is true}) = 9.11$, thus we take $n = 9$.

**5. Conclusions**

In this paper, an new approach for sequential test of fuzzy hypotheses based on fuzzy hypotheses for one-sample and two-sample when the available data are crisp, is presented. As for this paper, it sound the introduced method is very simple and applicable in the statistics and other sciences.

Extension of the proposed method to test the variance, correlation and parameters of linear models (regression models), design of experiment is a potential area for the future work. Furthermore, we can construct sequential test of fuzzy hypotheses based on intuitionistic fuzzy hypotheses or fuzzy data for the parameters of interest.

**6. References**


