퍼지자료에 관한 퍼지가설의 통계적 검정

On statistical testing for fuzzy hypotheses with fuzzy data

최 규 탁
경남정보대학 공업경영학과 교수

이 창 은, 강 만 기
동의대학교 전산통계학과 교수(mkkang@hyomin.dongeui.ac.kr)

ABSTRACT

We propose fuzzy statistical test of fuzzy hypotheses membership function with fuzzy number data. Finding the maximum grade of the meeting point for fuzzy hypotheses membership function and membership function of confidence interval. By the maximum grade, we obtain the results to acceptance or reject for the test of fuzzy hypotheses.

I. Introduction

Watanebe and Imaizumi[3] were suggested fuzzy statistical test for fuzzy hypotheses. But their idea of fuzzy hypotheses testing method are based on fuzzy hypothesis membership function. Other hand, Grzegorzewski[4] proposed fuzzy test for testing statistical hypotheses with vague data. He suggested a measure of fuzzyness of the considered fuzzy test and also discussed the robustness of that test. On the other hand, our aim is to treat fuzzy number data $\tilde{x}_i$ ($i=1,2,\cdots,n$) and fuzzy hypotheses contains in uncertainty. From the fuzzy number data, we seek the fuzzy test statistic with fuzzy sample mean and variance. We define fuzzy hypotheses membership function and obtain the fuzzy confidence interval from fuzzy test statistic, also find the maximum grade of the meeting for fuzzy hypotheses membership function and membership function of fuzzy confidence interval, thus we obtain the results by the grade for judgement to acceptance or reject for the fuzzy hypotheses.

II. Descriptive statistics for fuzzy data

We would like to illustrate how methods from descriptive statistic can be generalized for fuzzy data. In descriptive statistics a sample $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ is described by same characteristic values such as the sample mean $\bar{\tilde{x}}$, the sample variance $\tilde{s}^2$, or sample standard deviation $\tilde{s}$. All these value can be extended to a fuzzy sample $\tilde{x}$ as out lined below.

Corollary 1. Let $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ be a fuzzy sample with characterizing function $f_{\tilde{x}}(\cdot)$. The sample mean is a fuzzy number $\bar{\tilde{x}} \in F(R)$ with characterizing function $f_{\bar{\tilde{x}}}(\cdot)$ given by

$$f_{\bar{\tilde{x}}}(y') = \sup_{\tilde{x} \in X_y} f_{\tilde{x}}(x),$$

$$X_y = \left\{ \tilde{x} \in R^n : \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i = y \right\}, \quad \forall y \in R \tag{2-1}$$

The δ-cut representation is given by

$$C(\bar{\tilde{x}})_\delta = [\ \min_{\tilde{x} \in C(\tilde{x})\delta} \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i,\ \max_{\tilde{x} \in C(\tilde{x})\delta} \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i], \quad \forall \delta \in (0,1] \tag{2-2}$$

corollary 2. If $\tilde{x}$ in **corollary 1** is a minimum rule fuzzy sample mean is given by

$$C(\bar{\tilde{x}})_\delta = [\ \frac{1}{n} \sum_{i=1}^{n} C_L(\tilde{x}_i)_\delta,\ \frac{1}{n} \sum_{i=1}^{n} C_U(\tilde{x}_i)_\delta], \quad \forall \delta \in (0,1] \tag{2-3}$$

$C(\tilde{x}_i)_\delta = [C_L(\tilde{x}_i)_\delta, C_U(\tilde{x}_i)_\delta]$, is the δ-cut of the fuzzy data point $\tilde{x}_i$.

corollary 3. Let $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ be a fuzzy Sample with characterizing function $f_{\tilde{x}}(\cdot)$. The Sample variance and the sample standard deviation are fuzzy numbers $\tilde{S}^2 \in F(R)$ and $\tilde{S} \in F(R)$ with $supp(\tilde{S}^2) \subseteq R^+$ and $supp\,\tilde{S} \subseteq R^+$, respectively, with characterizing function given by

$$f_{(\tilde{S}^2)}(y) = \sup_{\tilde{x} \in X_Y} f_{\tilde{x}}(\tilde{x}),$$

$$X_Y = \left\{ \tilde{x} \in R^n : \frac{1}{n-1} \sum_{i=1}^{n} (\tilde{x}_i - C(\bar{\tilde{x}})_\delta)^2 = y \right\} \tag{2-4}$$

$$f_{\tilde{S}}(y) = \sup_{\tilde{x} \in X_Y} f_{\tilde{x}}(\tilde{x}),$$

$$X_Y = \left\{ \tilde{x} \in R^n : \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (\tilde{x}_i - C(\bar{\tilde{x}})_\delta)^2} = y \right\} \tag{2-5}$$

The δ-cut representations of $\tilde{S}^2$ and $\tilde{S}$ are given by

$$C((\tilde{s}^2))_\delta(y) = [\min_{\tilde{x} \in C(\tilde{x})\delta} g(\tilde{x}),\ \max_{\tilde{x} \in C(\tilde{x})\delta} g(\tilde{x})], \forall \delta \in (0,1], \tag{2-6}$$

$$C(\tilde{s})_\delta(y) = [\min_{\tilde{x} \in C(\tilde{x})\delta} \sqrt{g(\tilde{x})},\ \max_{\tilde{x} \in C(\tilde{x})\delta} \sqrt{g(\tilde{x})}], \forall \delta \in (0,1], \tag{2-7}$$

Where

$$g(\tilde{x}) = g(\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n) = \frac{1}{n-1} \sum_{i=1}^{n} (\tilde{x}_i - \frac{1}{n} \sum_{j=1}^{n} \tilde{x}_j)^2$$

III. Fuzzy hypotheses with fuzzy test statistic

Let $\tilde{x}$ be a random sample from sample space Ω and $\{P_\theta, \theta \in \Theta\}$ be a family of fuzzy probability, where θ is a parameter and Θ is a parameter space. For each $\psi \in \Theta$, we can consider a family of hypothesis $\{(H_0(\psi), H_1(\psi)) | \psi \in \Theta\}$. We introduce the fuzzy hypothesis as a fuzzy subset.

Definition 1. The fuzzy hypothesis H_f is a fuzzy subset of $\{(H_0(\phi), H_1(\phi)) | \phi \in \Theta\}$ with fuzzy hypothesis membership function $\chi_{H_f}((H_0(\phi), H_1(\phi))$.

We set with normal and convex

$$\chi_{H_f}(\phi) \equiv \chi_{H_f}((H_0(\phi), H_1(\phi)) \quad (3\text{-}1)$$

for simplicity.

The fuzzy null hypothesis and the fuzzy alternative hypothesis can be defined as follows.

Definition 2. The fuzzy null hypothesis $H_{f,0}$ is a fuzzy subset of Θ with a membership function $\chi_H(\phi)$. The fuzzy alternative hypothesis $H_{f,1}$ is a fuzzy subset of Θ and defined by the equation

$$H_{f,1} = \overline{H}_{f,0} \cap [\{ \bigcup_{\delta \in (0,1]} \delta(\bigcup_{\{\phi | \chi_H(\phi) \geq 0\}} \Theta_{K,\phi})\}] \quad (3\text{-}2)$$

where $\delta(A)$ stands for the fuzzy set whose membership function is product of a scalar δ and the characteristic function of a set A. The first term of the right hand side corresponds to the negation of the null hypothesis.

For example, the fuzzy hypothesis H_f can be interpreted as a hypothesis "$\theta \simeq \phi$", which is obtained by adding fuzzyness to an ordinary null hypothesis $H_0: \theta = \phi_0$, we note that $H_f: \theta \simeq \phi_0$ for such a fuzzy hypothesis.

Now we shall develop a product for testing the fuzzy hypothesis H_f. We assume the existence of a fuzzy test statistic $\tilde{T}(\phi)$ and critical region $K(\alpha, \phi)$ for a level of significance α. Define the real-valued function r_α on Θ by

$$r_\alpha(\phi) = (d_H(\tilde{T}(\phi)), K(\alpha, \phi)) / || \tilde{T}(\phi) || \quad (3\text{-}3)$$

where $d_H(A, B) = \sup\{d(a, B); a \in A\}$ is the Hausdorff separation of A from B and $||A|| = \sup\{||a||: a \in A\}$.

Using membership function $\chi_{\tilde{T}}(\phi)$ of $\tilde{T}(\phi)$, we also define the fuzzy hypothesis membership function $\chi_{R\alpha}$ on $\{0,1\}$ as follows:

$$\chi_{R\alpha}(0) = \sup_\phi \{\chi_H(\phi) \wedge \chi_{\tilde{T}}(\phi)\} \quad (3\text{-}4)$$

$$\chi_{R\alpha}(1) = 1 - \chi_{R\alpha}(0) \quad (3\text{-}5)$$

Let R_α denotes the fuzzy subset of an entire set $\{0,1\}$ defined by $\chi_{R\alpha}$, since $\{0,1\}$ corresponds {"accept", "reject"}, the value $\chi_{R\alpha}(1)$ and $\chi_{R\alpha}(0)$ are equal to the grades of the judgements that the hypothesis is reject and is not rejected respectively. We show the statistical properties of our testing method.

Theorem . If θ is a element of Θ then the grade of judgements of fuzzy hypothesis

$$P(\{\phi \mid \tilde{T}(\phi) \sqsubset K(\alpha, \phi)\}: \theta) \leq \alpha \quad (3\text{-}6)$$

is $r_\alpha(\phi)$, where $P(\cdot\,; \theta)$denotes the probability under the distribution P_Θ.

<proof> We can easily proved from (3-3)

IV. Application and example

We consider testing the fuzzy hypothesis $H_f: \theta \simeq \theta_0$ constructed by a set $\{(H_0(\phi), H_1(\phi)) | \phi \in \Theta\}$ and we gave fuzzy hypothesis membership function $\chi_H(\phi)$ such that $\chi_H(\theta_0) = 1$, where $H_0: \theta = \phi$, $H_1: \theta \neq \phi$. If there exists a fuzzy test statistic $\tilde{T}(\phi)$ from by (2-3) and (2-7) such as

$$\tilde{T}(\psi)_\delta = \frac{C(\tilde{\bar{x}}) - \theta}{\frac{C((\tilde{s})_\delta(y)}{\sqrt{n}}} \qquad (4\text{-}1)$$

then we can test the fuzzy hypothesis H_f . Suppose that $K(\alpha, \psi)$ does depend on ψ then we have fuzzy confidence interval $\overline{K(\alpha, \psi)} = (-k\alpha, k\alpha)$ for some positive number $k\alpha$, that is

$$P(\tilde{T}(\psi) \sqsubset (-k\alpha, k\alpha): \psi) \doteq 1-\alpha \qquad (4\text{-}2)$$

This implies that the condition $\tilde{T}(w) < \overline{k(\alpha, \psi)}$ is equivalents to the inequality

$$C(\tilde{\bar{x}})_\delta - k\alpha \frac{C(\tilde{S})_\delta(y)}{\sqrt{n}} \le \psi \le C(\tilde{\bar{x}})_\delta + k\alpha \frac{C(\tilde{S})_\delta(y)}{\sqrt{n}} \qquad (4\text{-}3)$$

There fore we have

$$\chi_R(0) = \sup_\psi \{\chi_H(\psi) \wedge \chi_{\tilde{T}}(\psi)\} \qquad (4\text{-}4)$$

which means that the grade of acceptance of H_f is equal to the maximum grade of $\chi_R(0)$ within the confidence interval of the parameter θ.

Example. We consider the fuzzy hypotheses

$$H_f: \theta \simeq 3.24$$

and we define the membership function for near equal to "3.2" as

$$\chi_H(\psi) = \begin{cases} 50(\psi - 3.22) & \psi \le 3.24 \\ 50(3.26 - \psi) & \psi > 3.24 \end{cases}$$

also we have that the observations yields n=5, sample mean with

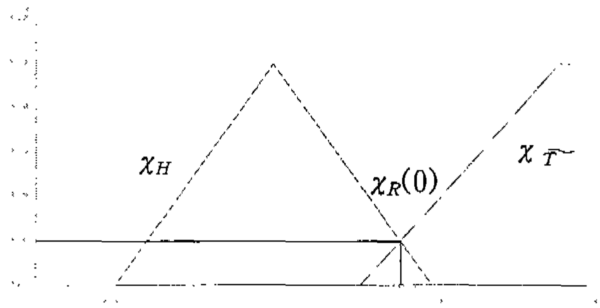
$C(\tilde{\bar{x}})_\delta = [0.01\delta + 4.34,\ 4.36 + 0.01\delta]$,

sample variance $\tilde{s}^2$ with

$C((\tilde{s}^2))_0$=[1.454333333, 1.538777778]

and confidence interval with $t_{0.05}(5)$=2.015 as [3.251051315 5.448948685], $\delta = 0$ and [3.2762544285.423745572], $\delta = 1$ from (4-3). The maximum value of $\chi_R(0)$ is 0.24 in the confidence interval from [Figure 1].

Thus we obtain the fuzzy test results R=0.24/0 + 0.76/1. The results say that the grade of rejection is 0.76 for $\alpha = 0.05$.



[Figure 1]

V. References

[1] D. Pill and K. Peter, *Metric space of fuzzy sets*, World Scientific, 1994.

[2] F. Li and C. Wu etc., Platform type fuzzy number a separability of fuzzy number space, *Fuzzy Sets and Systems*. 117(2000), 347-353.

[3] N. Watanabe and T. Imaizumi, A fuzzy statistical test of fuzzy hypotheses, *Fuzzy Sets and Systems*. 53(1993), 167-178.

[4] P. X. Gizegorzewski, Testing hypotheses with vague data, *Fuzzy Sets and Systems*. 112(2000), 501-510.

[5] S. Fruhwirth-Schnatter, On statistical inference for fuzzy data with application to des criptive statistics, *Fuzzy Sets and Systems*. 50(1992), 143-165.