Fuzzy Hypotheses Testing by Vague Response Data with Reflected Correlation

Man-Ki Kang *, Chang-Eun Lee *, Ji-Young Jung *, Gyu-Tag Choi **
* Dept. of Data Information Science, Dongeui University
(mkkang@dongeui.ac.kr)
**Business Management Division, Kyungnam College of Information and Technology

Abstract
We propose some properties of fuzzy p-value and fuzzy significance level to the fuzzy hypotheses testing for vague response data with reflected correlation in survey research. Appling the principle of agreement index, we suggest the methods for fuzzy hypothesis testing by fuzzy rejection region.

Key words: vague data, degree of acceptance and rejection, fuzzy hypotheses testing, agreement index.

1. Preliminaries

In the traditional approach to hypotheses testing all the concepts are precise and well defined for survey research data. However if we consider vagueness into observations, we would be faced that hypotheses are quitely new and interesting problems.

Kang and Lee(2000, 2002, 2003) defined fuzzy hypotheses membership function also they found the agreement index by area for fuzzy hypotheses membership function and membership function of fuzzy confidence interval, thus they obtained the results by the grade for judgement to acceptance or rejection for the fuzzy hypotheses

Now, we propose some properties of vague response data in survey research with reflected correlation and fuzzy hypotheses testing by critical region with agreement index for fuzzy number.

We considered the fuzzy hypothesis

$$H_{f.0}: \theta \simeq \psi, \quad \theta \in \Theta$$

constructed by a set

$$\{(H_0(\psi), H_1(\psi)) | \psi \in \Theta\}$$

with membership function $m_{H_f}(\psi)$ where Θ is parameter space.

A fuzzy number A in $\mathcal{R}$ is said to be convex if for any real numbers $x, y, z \in \mathcal{R}$ with $x \leq y \leq z$,

$$m_A(y) \geq m_A(x) \wedge m_A(z) \tag{1.1}$$

with $\wedge$ standing for minimum.

A fuzzy number A is called normal if the following holds

$$\bigvee_x m_A(x) = 1. \tag{1.2}$$

An γ-level set of a fuzzy number A is a set denoted by $[A]^\gamma$ and is defined by

$$[A]^\gamma = \{x | m_A(x) \geq \gamma, \ 0 < \gamma \leq 1\}. \tag{1.3}$$

An α-level set of fuzzy number A is a convex fuzzy set which is a closed bounded interval denoted by $[A]^\gamma = [A_l^\gamma, A_r^\gamma]$.

2. Fuzzy test statistics

Let A and B be fuzzy numbers in $\mathcal{R}$ and let $\odot$ be a binary operation defined in $\mathcal{R}$. Then the operation $\odot$ can be extended to the fuzzy numbers A and B by defining the relation(the extension principle).

Let $A, B \subset \mathcal{R}$, $\forall x, y, z \in \mathcal{R}$:

$$m_{A \odot B}(z) = \bigvee_{z = x \cdot y} (m_A(x) \wedge m_B(y)). \quad (2.1)$$

A modelling the fuzziness of data were described the fuzziness of a sample $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$. As a precise sample x of n precise realizations $x_i \in \mathcal{R}$ may be regarded as vector in $\mathcal{R}^n$, a fuzzy sample $\tilde{x}$ of n fuzzy realizations $\tilde{x}_i \in \mathcal{F}(\mathcal{R})$ may be regarded as a fuzzy vector.

A piecewise continuous function $m_{\tilde{x}}: \mathcal{R} \to [0,1]$, fuzzy number $\tilde{x}$ in $\mathcal{F}(\mathcal{R})$ if the family

$$[\tilde{x}]^\gamma = \{x \in \mathcal{R},\ m_{\tilde{x}}(x) \geq \gamma\}, \quad \forall \gamma \in (0,1]$$

has the following properties:

$[\tilde{x}]^\gamma$ is not empty for $\gamma = 1$ and closed, finite interval $[[\tilde{x}]^\gamma_l, [\tilde{x}]^\gamma_r]$ in $\mathcal{R}$ and simple connected, compact in $\mathcal{R}^n$.

Assuming that a precise sample x is given, the sample mean $\bar{x}$ is the image of the sample x under the function $f(x)$ given by $f(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Definition 2.1. Let $\tilde{x} \in \mathcal{F}(\mathcal{R}^n)$ be a fuzzy sample with $m_{\tilde{x}}(\cdot)$ and γ-cut representation $\{[\tilde{x}]^\gamma : \gamma \in (0,1]\}$, and let $f: \mathcal{R}^n \to Y$, $Y \subseteq \mathcal{R}$, be a real valued continuous mapping. The fuzzy image $\tilde{u} = f(\tilde{x})$ of the fuzzy number data $\tilde{x}$ under the mapping $f(\cdot)$ is defined by the following characterizing function :

$$m_{\tilde{u}}(u) = \sup_{x \in X_u} m_{\tilde{x}}(x),$$

$$X_u = \{x \in \mathcal{R}^n : f(x) = u\},\ \forall u \in \mathcal{R}. \quad (2.2)$$

with $\sup_{x \in X} m_{\tilde{x}}(x) := 0$, if X_u is empty.

Definition 2.2. From the fuzzy sample, the sample mean is a fuzzy number $\bar{\tilde{x}} \in \mathcal{F}(\mathcal{R})$ with characterizing function $m_{\bar{\tilde{x}}}(\cdot)$ given by

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

$$X_y = \left\{ x \in \mathcal{R}^n : \frac{1}{n}\sum_{i=1}^{n} x_i = y \right\},\ \forall y \in \mathcal{R} \quad (2.3)$$

The γ-cut representation is given by

$$[\bar{\tilde{x}}]^\gamma = \left[\min_{x \in [\tilde{x}]^\gamma} f(x),\ \max_{x \in [\tilde{x}]^\gamma} f(x) \right],$$
$$\forall \gamma \in (0,1] \quad (2.4)$$

where $f(x) = g(x_1, x_2, \cdots, x_n) = \frac{1}{n}\sum_{i=1}^{n}(x_i)$.

Definition 2.3. If $\tilde{x}$ in Definition 2.1 is a minimum rule of fuzzy sample, the γ-cut representation of fuzzy sample mean is given by

$$[\bar{\tilde{x}}]^\gamma = \left[\frac{1}{n}\sum_{i=1}^{n}[\tilde{x}_i]^\gamma_l,\ \frac{1}{n}\sum_{i=1}^{n}[\tilde{x}_i]^\gamma_r \right],\ \forall \gamma \in (0,1] \quad (2.5)$$

$[\tilde{x}_i]^\gamma = [[\tilde{x}_i]^\gamma_l, [\tilde{x}_i]^\gamma_r]$ is the γ-cut of the fuzzy data point $\tilde{x}_i$.

An approximate solution is based on the introduction of a *standardized fuzzy sample* $\tilde{z} = (\tilde{z}_1, \tilde{z}_2, \cdots, \tilde{z}_n)$ define by $\tilde{z}_i = \tilde{x}_i - \bar{\tilde{x}}$.

The γ-cut of each $\tilde{z}_i$ is given by

$$[\tilde{z}_i]^\gamma = [[\tilde{z}_i]^\gamma_l, [\tilde{z}_i]^\gamma_r]$$
$$= [[\tilde{x}_i]^\gamma_l - [\bar{\tilde{x}}]^\gamma_r, [\tilde{x}_i]^\gamma_r - [\bar{\tilde{x}}]^\gamma_l] \quad (2.6)$$

Now renumber that the sample variance s^2 of sample x is related to the second empirical moment of the standardized sample by $z = (x_1 - \bar{x}, x_2 - \bar{x}, \cdots, x_n - \bar{x})$ by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} z_i^2 . \quad (2.7)$$

A minimization rule fuzzy sample this approximation results in the following γ-cut representation of fuzzy sample variance and the fuzzy sample standard deviation, respectively:

$$[\tilde{s}^2]^\gamma = \left[\frac{1}{n-1}\sum_{i=1}^{n}[\tilde{z}_i^2]^\gamma_l,\ \frac{1}{n-1}\sum_{i=1}^{n}[\tilde{z}_i^2]^\gamma_r \right],$$
$$\forall \gamma \in (0,1] \quad (2..8)$$

$$[\tilde{s}]^\gamma = \left[\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[\tilde{z}_i^2]^\gamma_l},\ \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}[\tilde{z}_i^2]^\gamma_r} \right],$$
$$\forall \gamma \in (0,1] \quad (2.9)$$

where

$$[z_i^2]_l^\gamma = \begin{cases} ([\tilde{z}_i]_l^\gamma)^2, & [\tilde{z}_i]_l^\gamma \geq 0 \\ ([\tilde{z}_i]_r^\gamma)^2, & [\tilde{z}_i]_r^\gamma \leq 0 \\ 0, & oterwise \end{cases} \quad (2.10)$$

$$[z_i^2]_r^\gamma = \begin{cases} ([\tilde{z}_i]_r^\gamma)^2, & [\tilde{z}_i]_l^\gamma + [\tilde{z}_i]_r^\gamma \geq 0 \\ ([\tilde{z}_i]_l^\gamma)^2, & oterwise. \end{cases} \quad (2.11)$$

Thus, $[\tilde{z}_i]^\gamma = [[\tilde{z}_i]_l^\gamma, [\tilde{z}_i]_r^\gamma]$ are the γ-cuts of the components of the standardized fuzzy sample $\tilde{z}$.

3. Fuzzy hypothesis test

Let x be a random sample from sample space Ω. Let $\{P_\Theta, \Theta \in \Theta\}$ be a family of fuzzy probability distribution, where Θ is a parameter vector and Θ is a parameter space.

Definition 3.1. Let a fuzzy membership function $m_H(x)$, $x \in R$, we consider another membership function $m_A(x)$, $x \in R$, which we call the agreement index of A with regard to H, the ratio being defined in the following way;

$$R(A, H) = (area(m_A(x) \cap m_H(x)) / area(m_A(x)) \in [0, 1] \quad (3.1)$$

as shown in Fig 3.1.

Definition 3.2. We define agreement index by real-valued function R_γ on Θ as the maximum grade membership function of acceptance or rejection is

$$m_{\tilde{R}_\gamma}(0) = \sup_\psi \{area\ (m_{H_\gamma}(\psi) \cap m_{T_\gamma}(\psi)) / area\ m_{H_\gamma}(\psi)\} \quad (3.2)$$

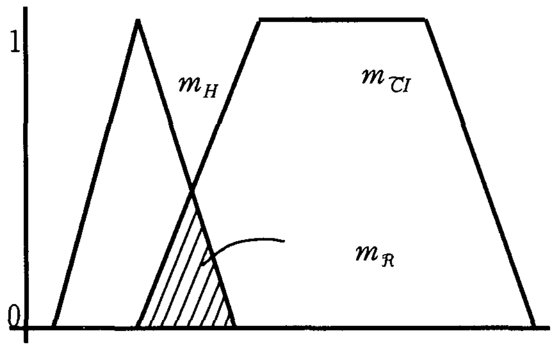
$$m_{\tilde{R}_\gamma}(1) = 1 - m_{\tilde{R}_\gamma}(0) \quad (3.3)$$

for the fuzzy hypothesis testing as [Figure 3.1]

Definition 3.3. In agreement index, we have the area by α-level as:

$$area(m_A(x) \cap m_H(x) = \int_{\alpha_0}^{\alpha_1} (A_r^{-1}(\alpha) - H_l^{-1}(\alpha)) d\alpha \Big/ \int_{\alpha_0}^{1} (A_r^{-1}(\alpha) - A_l^{-1}(\alpha)) d\alpha \quad (3.4)$$

where A_r, A_l are right and left side line of $m_A(x)$, H_l is left side line of $m_H(x)$ and α_0 is reliable degree and α_1 is meeting point of $m_A(x)$ and $m_H(x)$.



[Figure 3.1]

4. Example

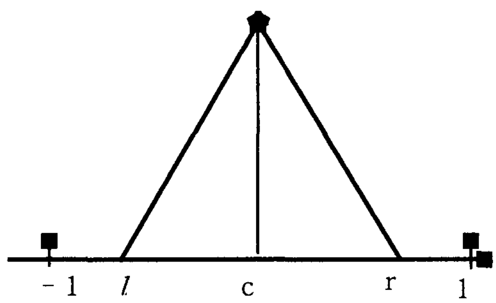
We have a vague response data as fuzzy number $[l, c, r]$, $0 \leq l \leq c \leq r \leq 1$, then the membership function is

$$\mu_A(x) = \begin{cases} \frac{1}{c-l}x - \frac{l}{c-l}, & l \leq x \leq c \\ -\frac{1}{r-c}x + \frac{r}{r-c}, & c \leq x \leq r \end{cases} \quad (4.1)$$

as [Figure 4.1].

By γ-level set, we have the fuzzy number

$$[\tilde{x}]^\gamma = [(c-l)\gamma + l, \ -(r-c)\gamma + r]. \quad (4.2)$$



[Figure 4.1]

For comparing two treatment, if we have independent random response samples fuzzy numbers from two population such as:

samples1($\tilde{x}$)

(-0.02, 0.01, 0.05), (0.00, 0.03, 0.05),
(-0.09, -0.07, -0.03), (0.02, 0.05, 0.06),
(-0.01, 0.00, 0.03), (-0.08, -0.04, -0.01),
(0.01, 0.05, 0.07), (-0.09, -0.06, -0.04),
(0.03, 0.07, 0.09), (-0.07, -0.05, -0.04)

samples2($\tilde{y}$)

(0.01, 0.03, 0.05), (0.00, 0.03, 0.04),
(0.00, 0.01, 0.02), (-0.01, 0.02, 0.02),
(-0.01, 0.00, 0.01), (0.01, 0.04, 0.05),
(0.01, 0.02, 0.05), (0.00, 0.04, 0.05),
(-0.01, 0.04, 0.07), (0.02, 0.05, 0.07)

With the objective of drawing a comparison between two populations or, synonymously, between two treatment, we examine the situation.

Assuming that the measurements consistence a fuzzy samples from the population then we have fuzzy sample mean as:

$$\bar{\tilde{x}} = [-0.03 + 0.029\gamma,\ 0.023 - 0.024\gamma],$$

$$\bar{\tilde{y}} = [0.002 + 0.026\gamma,\ 0.043 - 0.015\gamma]$$

fuzzy sample variance as:

$$\tilde{s}_x^2 = [0.000477 - 0.00123\gamma + 0.003362\gamma^2,\ 0.010411 - 0.0109\gamma + 0.003094\gamma^2]$$

$$\tilde{s}_y^2 = [0.002446 - 0.00407\gamma + 0.001862\gamma^2,\ 0.00182 - 0.00373\gamma + 0.002151\gamma^2]$$

By using

$$\tilde{s}_p^2 = \frac{1}{m+n-2}\{(m-1)\tilde{s}_x^2 + (n-1)\tilde{s}_y^2\} \quad (4.3)$$

we have pooled sample mean

$$\tilde{s}_p^2 = [0.000161 - 0.00265\gamma + 0.002612\gamma^2,\ 0.006115 - 0.00731\gamma + 0.002623\gamma^2]$$

If we have $\gamma=1$, and the level of significance $\alpha = 0.1$, employ the t-test

$$\tilde{t} = \frac{|\tilde{x} - \tilde{y}|}{\tilde{s}_p\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m+n-2) \quad (4.1)$$

$[1.718, 1.718] < t(0.05;18)=1.734$

thus we accept the alternative hypothesis $H_{f.0}$.

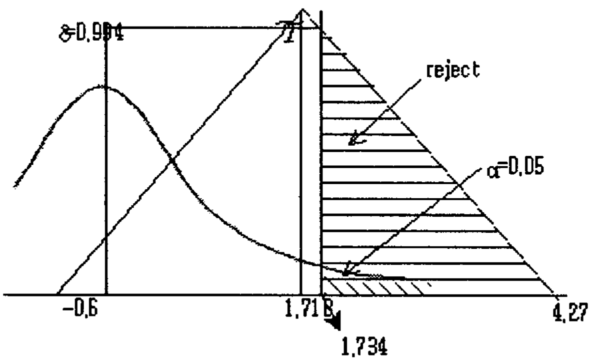
If we have $\gamma=0.0$, from the t-test

$\tilde{t}=[-0.600, 4.270]$ and $t(0.05;18)=1.734$,

By γ-level, if we have $\gamma_1 = 1$

$$\frac{\int_{1.734}^{4.27}(-0.392t+1.673)\,dt}{\int_{-0.6}^{1.718}(0.339t+0.417)\,dt + \int_{1.718}^{4.27}(-0.392t+1.673)\,dt} = 0.4584$$

Thus, we reject to alternative hypothesis $H_{f.1}$ degree of $m_{R_\gamma}(1) = 0.4584$ by [Figure 4.2]



[Figure 4.2]

References

[1] P. X. Gizegorzewski, Testing Hypotheses with vague data, *Fuzzy Sets and Systems*. 112 , pp.501-510, 2000.

[2] M. K. Kang, G. T. Choi and C. E. Lee, On Statistical for Fuzzy Hypotheses with Fuzzy Data, *Proceeding of Korea Fuzzy Logic and Intelligent System Society Fall Conference*, Vol. 10, Num. 2, 2000.

[3] M. K. Kang, G. T. Choi and S. I .Han, A Bayesian Fuzzy Hypotheses testing with Loss Function, *Proceeding of Korea Fuzzy Logic and Intelligent Systems Society Fall Conference*, Vol 13, Num. 2, 2002.

[4] M. K. Kang, Lee, C. E. and Han, S. I. . Fuzzy Hypotheses Testing for Hybrid Numbers by Agreement Index, *Far East Journal of Theoretical Statistics*. ,10(1), 2003.

[5] S. F. Schnatter, S. On Statistical inference for Fuzzy Data whit Application to Des Crispy Statistics. *Fuzzy Sets and Systems*. 50. pp 143-165. 1992.

[6] N. Watanabe, T. Imaizumi, A Fuzzy Statistical Test of Fuzzy Hypotheses, *Fuzzy Sets and systems* 53 , pp.167-178, 1993.