

## 내분비에 의한 퍼지 가설 검정

# Fuzzy Test of Hypotheses by Rate of Internal Division

강만기<sup>†</sup> · 정지영

Man Ki Kang<sup>†</sup> and Ji Young Jung

Department of Data Information Science, Dong-Eui University

### Abstract

We propose some properties for fuzzy hypotheses testing by the principle of the rate of internal division on delta-levels. By the rate of internal division, we show that the acceptance and rejection degree for fuzzy the fuzzy hypotheses and reduce the spread of the fuzzy variance with average of the center and left or right spread of fuzzy number data.

**Key Words** : fuzzy number data, degree of acceptance and rejection, fuzzy hypotheses testing, rate of internal division.

## 1. Introduction

For the uncertainty data, fuzzy hypotheses testing method concerning the fuzzy expected value combing with other fuzzifications have shown to be very useful inferences about distributions of real-valued random variables[1].

We define the fuzzy null hypotheses membership function as

$$H_{f0}: m_{\theta} = m_{\theta_0} \text{ or } H_{f0}: m_{\theta} < m_{\theta_0}, \theta \in \Omega$$

where  $\Omega$  is parameter space. Kang and Seo defined an agreement index by area ratio for the fuzzy hypotheses membership function with respect to membership function of fuzzy critical region[2]. They obtained the results by the grade for the judgement of acceptance or rejection for the fuzzy hypotheses.

Now, we suggest some properties for a fuzzy hypothesis test by a rate of internal division with a reduced-spread fuzzy variance and covariance using an average of fuzzy number's center and spreads[3,4]. Thus, we illustrate the fuzzy hypothesis  $t$ -test by difference of hypothyroid samples for symptom light and heavy with a fuzzy degree of freedom.

This paper is organized as follows. In chapter 2, we suggest a reduced-spread of fuzzy variance. The

acceptance or rejection degree by rate of internal division for the fuzzy hypothesis was shown in Chapter 3. Finally, We illustrate the fuzzy hypothesis by thyroxine random sample data of hypothyroids difference of symptom light and heavy.

## 2. Reduced-spread of fuzzy variance

Let  $K_c(\mathfrak{R}^p)$  be the class of the non-empty compact convex subsets of  $\mathfrak{R}^p$ . We will consider the class of fuzzy sets

$$F_c(\mathfrak{R}^p) = \{U: \mathfrak{R}^p \rightarrow [0,1] \mid U^{(\delta)} \in K_c(\mathfrak{R}^p) \text{ for all } \delta \in [0,1]\} \quad (2.1)$$

where  $U^{(\delta)}$  stands for the  $\delta$ -level of  $U$  (i.e.  $U^{(\delta)} = \{x \in \mathfrak{R}^p \mid U(x) \geq \delta\}$ ) for all  $\delta \in (0,1]$ ,  $\delta$  is precision of data in statistical concept and  $U^{(0)}$  is the closure of the support of  $U$  [5].

The space  $F_c(\mathfrak{R}^p)$  can be endowed with the sum and the product a scalar based on Zadeh's extension principle. Let  $A$  and  $B$  be fuzzy numbers data in  $\mathfrak{R}$  and let  $\odot$  be a binary operation defined in  $\mathfrak{R}$ . Then the operation  $\odot$  can be extended to the fuzzy numbers  $A$  and  $B$  by defining the relation of extension principle as;

given  $A, B \subset \mathfrak{R}$ ,  $\forall x, y, z \in \mathfrak{R}$ :

$$m_{A^{(\delta)} \odot B^{(\delta)}}(z) = \bigvee_{z=x \odot y} (m_{A^{(\delta)}}(x) \wedge m_{B^{(\delta)}}(y)) \quad (2.2)$$

by  $\delta$ -level.

접수일자: 2012년 5월 20일

심사(수정)일자: 2012년 6월 20일

게재 확정일자 : 2012년 8월 14일

† 교신 저자

This work was supported by Dong-eui University Grant.(2012AA086)

Let  $(\Omega, A, P)$  be the probability space. A random fuzzy variable (RFV) is a mapping  $\chi: \Omega \rightarrow F_c(\mathbb{R}^p)$  so that the  $\delta$ -level functions  $\chi^{(\delta)}: \Omega \rightarrow K_c(\mathbb{R}^p)$ , defined so that  $\chi^{(\delta)}(\theta) = (\chi(\theta))^{(\delta)}$  for all  $\theta \in \Omega$ , are the random sets.

**Definition 2.1.** If  $\chi: \Omega \rightarrow F_c(\mathbb{R}^p)$  is a RFV such that  $E(\sup_{x \in \chi_0} \|x\|) < \infty$  with  $|\chi_0|(\theta) = \sup\{|x| | x \in \chi_0(\theta)\}$  for all  $\theta \in \Omega$ , then the *expected fuzzy value* (or *mean*) of  $\chi$  is the unique  $\tilde{E}(\chi) \in F_c(\mathbb{R}^p)$ . Thus we have Aumann's expectation of the random set  $\chi^{(\delta)}$  for all  $\delta \in [0, 1]$  as

$$(\tilde{E}(\chi))^{(\delta)} = \{E(X|P) | X: \Omega \rightarrow \mathbb{R}^p, X \in L^1(\Omega, A, P), X \in \chi^{(\delta)} \text{ a.s. } [P]\} \quad (2.3)$$

where  $L^1$  is a metrics for left, center and right spreads form origin "0" of fuzzy number data, respectively.

**Definition 2.2.** Given a probability space  $(\Omega, A, P)$ , if RFV  $\chi$  have  $E(|\chi|) < \infty$  then *fuzzy variance* of  $\chi$  is defined as

$$\sigma_{\chi}^2 = \text{Var}(\chi) = E([D_A(\chi, \tilde{E}(\chi))]^2) \quad (2.4)$$

where  $D_A^2(\cdot)$  are  $(c^2 + \min\{l^2, r^2\})/2$ ,  $c^2$  and  $(c^2 + \max\{l^2, r^2\})/2$  for  $\chi \ominus \tilde{E}(\chi) = [l, c, r]$  by Zadeh's extension principle.

A modeling of data were describe a random sample  $\chi = \{\chi_1, \chi_2, \dots, \chi_n\}$ . If we observe an object  $x_i$  at three times then we have minimum, median or maximum value (minimum, median or maximum error term data). Thus we organize a fuzzy number data to  $\chi_i = [\min(\chi_{ik}), \text{median}(\chi_{ik}), \max(\chi_{ik})] = [x_{il}, x_{ic}, x_{ir}]$ ,  $k = 1, 2, 3$ ,  $i = 1, 2, \dots, n$ .

We have the fuzzy sample mean as:

$$\begin{aligned} \bar{\chi}^{(\delta)} &= \left[ \frac{1}{n} \sum_{i=1}^n \chi_i \right]^{(\delta)} \\ &= \left[ \frac{1}{n} \sum_{i=1}^n x_{il}, \frac{1}{n} \sum_{i=1}^n x_{ic}, \frac{1}{n} \sum_{i=1}^n x_{ir} \right]^{(\delta)} \\ &= [\bar{x}_l, \bar{x}_c, \bar{x}_r]^{(\delta)}, \end{aligned} \quad (2.5)$$

by definition 2.1 for  $\delta$ -level.

If we let  $[l_i, l_i, l_i] < \chi_i \ominus \bar{\chi} < [r_i, r_i, r_i]$  for  $[l_i, c_i, r_i]$  then

$$[\min\{l_i^2, r_i^2\}, \min\{l_i^2, r_i^2\}, \min\{l_i^2, r_i^2\}] < (\chi_i \ominus \bar{\chi})^2$$

$$< [\max\{l_i^2, r_i^2\}, \max\{l_i^2, r_i^2\}, \max\{l_i^2, r_i^2\}], \quad i = 1, \dots, n. \quad (2.6)$$

However, in case of overlapping  $\chi_i$  by  $\bar{\chi}$ , we have

$$[0, 0, 0] < (\chi_i \ominus \bar{\chi})^2 < [\max\{l_i^2, r_i^2\}, \max\{l_i^2, r_i^2\}, \max\{l_i^2, r_i^2\}], \quad i = 1, \dots, n. \quad (2.7)$$

In order to reduce the spreads of fuzzy numbers for sample fuzzy variance, we have reduced-spread fuzzy variance in this paper as:

$$\begin{aligned} \tilde{S}^2 &\doteq \frac{1}{n-1} \sum_{i=1}^n \left( [(c_i^2 + \min\{l_i^2, r_i^2\})/2, c_i^2, (c_i^2 + \max\{l_i^2, r_i^2\})/2], \right. \\ &\quad \left. \text{for } l_i > 0 \text{ or } r_i < 0 \right. \\ &\quad \left. [(0 + c_i^2)/2, c_i^2, (c_i^2 + \max\{l_i^2, r_i^2\})/2], \text{ for } l_i \leq 0 \text{ and } r_i > 0 \right) \\ &= [S_l^2, S_c^2, S_r^2] \end{aligned} \quad (2.8)$$

where  $\doteq$  is analogously equal by definition 2.2.

### 3. Acceptance or rejection degree by rate of internal division

Let  $T$  is a test statistics by fuzzy random sample from sample space  $\Omega$ . Let  $\{P_\theta, \theta \in \Omega\}$  is a family of fuzzy probability distribution, where  $\theta$  is a parameter vector of  $\Omega$ . Choose a membership function  $m_T(x)$  of  $T$  whose value is likely to best reflect the plausibility of the fuzzy hypothesis being tested. Let us consider membership function  $m_C(x)$  of critical region  $C$ , we have a index of rate of internal division of  $m_T(x)$  which regard to  $m_C(x)$  by  $\delta$ -level.

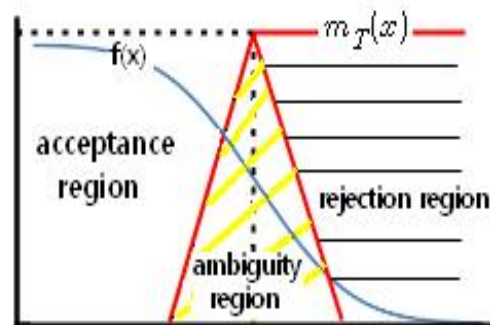


Fig. 3.1. Fuzzy test statistic

**Definition 3.1.** Let a fuzzy membership function  $m_T(x)$ ,  $x \in \mathbb{R}$ , we consider a degree of rate of internal division by  $\delta$ -level as;

$$\mathbb{D}_{ID} = \frac{T_r - k}{T_r - T_l} \text{ for all } k \in (T_l, T_r) \quad (3.1)$$

for the fuzzy number  $[T]^{(\delta)} = \{x | m_T(x) \geq \delta, 0 \leq \delta \leq 1\}$   
 $= [T_l, T_c, T_r]^{(\delta)}$ .

**Definition 3.2.** We define real-valued function  $\mathbb{D}(\cdot)^{(\delta)}$  by supremum grade of membership function for rejection or acceptance degree by rate of internal division by  $\delta$ -level as;

$$\mathbb{D}(0)^{(\delta)} = \sup_{\delta} \left\{ \frac{(C_r^{(\delta)} - C_l^{(\delta)}) \times \frac{1}{2} + (T_r^{(\delta)} - C_r^{(\delta)}) \times 1}{T_r^{(\delta)} - T_l^{(\delta)}} \right\} \quad (3.2)$$

for the fuzzy number  $[C]^{(\delta)} = \{x | m_C(x) \geq \delta, 0 \leq \delta \leq 1\}$   
 $= [C_l, C_c, C_r]^{(\delta)}$  and  $T_l^{(\delta)} < C_l^{(\delta)} < C_r^{(\delta)} < T_r^{(\delta)}$ ,

$$\mathbb{D}(1)^{(\delta)} = 1 - \mathbb{D}(0)^{(\delta)}, \quad (3.3)$$

for the fuzzy hypothesis testing, respectively as figure 3.1.

The " $\frac{1}{2}$ " and "1" for equation (3.2) are ambiguity decision and exactly decision within fuzzy rejection region.

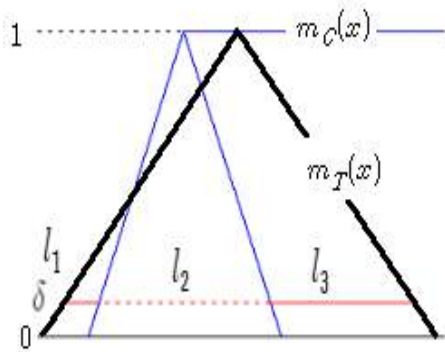


Fig. 3.2. Test statistics and critical region

For example, let  $l = T_r^{(\delta)} - T_l^{(\delta)}$ ,  $l_1 = C_r^{(\delta)} - T_l^{(\delta)}$ ,  
 $l_2 = T_r^{(\delta)} - C_l^{(\delta)}$  and  $l_3 = T_r^{(\delta)} - C_r^{(\delta)}$ , if  $l = l_1$  then  
 $\mathbb{D}(0)^{(\delta)} = 0$  by  $\delta = \delta_2$   
 and

$$\mathbb{D}(0)^{(\delta)} = \sup_{\delta} \left\{ \frac{l_1 \times 0 + l_2 \times \frac{1}{2}}{l} \right\} \quad (3.4)$$

by  $\delta = \delta_1$  as figure 3.1.

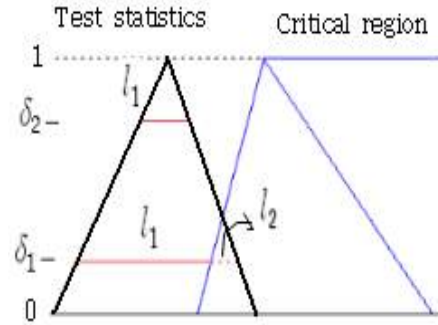


Fig. 3.3. Test statistics and critical region by  $\delta$ -level

### 4. Illustration

We can illustrate the fuzzy hypotheses test by rate of internal division on fuzzy  $t$ -test for the difference of thyroxine of hypothyroid sample data for symptom light and heavy by an experiment.

For the fuzzy hypothesis

$$H_{f0}: m_{\mu_1} = m_{\mu_2}, \quad (4.1)$$

we have fuzzy  $t$ -test statistics from [Table 4.1] as;

$$\begin{aligned} \tilde{t} &= ((\tilde{x}_1 \ominus \tilde{x}_2) \ominus (\mu_1 \ominus \mu_2)) \oslash \sqrt{\frac{\tilde{s}_1^2}{n_1} \oplus \frac{\tilde{s}_2^2}{n_2}} \\ &= (([53.4, 56.4, 59.3] \ominus [35.3, 37.4, 40.3]) \ominus [0, 0, 0]) \\ &\quad \oslash \sqrt{\frac{[10.5, 14.2, 18.3]}{9} + \frac{[26.4, 31.0, 36.7]}{7}} \\ &= [0.9, 1.0, 2.0] \end{aligned}$$

From the thyroxine of hypothyroid sample data, we have sample size  $n_1 = 9$  and  $n_2 = 7$ .

By

$$\tilde{w}_1 = \frac{\tilde{s}_1^2}{n_1} = \frac{[10.5, 14.2, 18.3]}{9}$$

$$\text{and } \tilde{w}_2 = \frac{\tilde{s}_2^2}{n_2} = \frac{[26.4, 31.0, 36.7]}{7},$$

we have degree of freedom as;

$$\begin{aligned} \nu &= (w_1 \oplus w_2)^2 \oslash \left( \frac{w_1^2}{n_1 - 1} + \frac{w_2^2}{n_2 - 1} \right) \\ &= [4.8, 10.1, 20.8] \\ &\simeq [5, 10, 21]. \end{aligned}$$

Some significance level  $\alpha = 0.2$  and degree of freedom [5, 10, 21], we have ambiguity rejection region [1.3, 1.4, 1.5] and fuzzy rejection region [1.3, 1.4,  $\infty$ ].

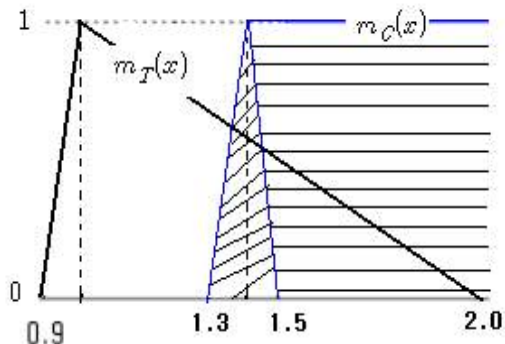


Fig. 4.1.  $t$ -test statistics and rejection region

Table 4.1. Thyroxine data of hypothyroid

symptom light(nmol/l)	symptom heavy(nmol/l)
[32,34,37]	[4,5,8]
[43,45,49]	[7,8,10]
[46,49,52]	[16,18,19]
[53,55,58]	[21,24,26]
[55,58,61]	[57,60,64]
[57,59,62]	[80,83,87]
[56,60,63]	[62,64,70]
[58,62,66]	
[81,86,90]	

By definition 3.2, we have rejection degree

$$\mathbb{D}^{(\delta=0)}(0) = \sup_{\delta=0} \left\{ \frac{(1.5-1.3) \times \frac{1}{2} + (2.0-1.5) \times 1}{2.0-0.9} \right\} = 0.56$$

for the hypothesis equation (4.1).

If the data are crisp cases, that is the precision  $\delta=1$ , then the fuzzy statistics is revolve to ordinary statistics. Thus we have rejection degree  $\mathbb{D}^{(\delta=1)}(0)=0$  by definition 3.2 for the hypothesis equation (4.1) as <Figure 4.1>.

### References

[1] P. X. Gizgorzewski, "Testing hypotheses with vague data," *Fuzzy Sets and Systems*, vol. 112, pp. 501-510, 2000.

[2] M. K. Kang, and H. A. Seo, "Fuzzy Hypothesis Test by Poisson Test for Most powerful Test," *Journal of Korean Institute of Intelligent Systems*, vol 19, no. 6, 809-813, 2009.

[3] M. K. Kang, "Fuzzy test of hypothesis by uniformly most powerful test," *Journal of Korean Institute of Intelligent Systems*, vol. 21, no. 1, pp. 25-28, 2011.

[4] M. K. Knng, "Corrlelation test by reduced-spread of fuzzy variance," *Communications of the Korean Statistical Sociate*, vol. 19, no. 1, pp. 147-155, 2012.

[5] A. Colubi, "Statistical inference about the means of fuzzy random variables: Applications to the analysis of fuzzy-and real-valued data," *Fuzzy Sets and Systems*, vol. 160, pp. 344-356, 2009.

### 저 자 소 개



**강만기 (ManKi Kang)**

1990년 : 동아대학교 이학박사  
1989년~현재 : 동의대학교 데이터정보학과 교수

관심분야 : 퍼지데이터통계처리  
E-mail : mkkang@deu.ac.kr



**정지영 (Ji Young Jung)**

현재 : 동의대학교 대학원 정보통계학과 박사과정

관심분야 : 퍼지데이터처리  
E-mail : jjy8255@hanmail.net