퍼지 및 지능시스템학회 논문지 2005, Vol. 15, No. 5, pp. 533-537

리커트 퍼지 척도에 대한 퍼지 가설검정

Fuzzy Hypotheses Testing of Likert Fuzzy Scale

강만기*, 이창은*, 최규탁**

Man-Ki Kang*, Chang-Eun Lee*, Gue-Tak Chio**

*동의대학교 자연과학대학 정보통계학과

**경남정보대학 산업시스템경영과

요 약

리커트 척도는 어떤 의도의 질문에 대한 동의를 수치적으로 나타내는 방법으로서 주로 5점 척도를 많이 사용하나 여러 번 질문에서 응답자체가 애매성을 가지고 있다. 본 연구에서는 3점 척도의 응답의 애매성을 5점 척도의 멤버쉽함수로 나타낸 후 한 항목의 응답의 95% 신뢰구간을 다시 멤버쉽함수로 하여 그 데이터를 퍼지가설검정하였다.

Abstract

A Likert scale is an often used questionnaire format. It requests respondents to specify their level of agreement to each of a list of statements. A typical question using a five-point Likert scale might make a statement. The results shows vague values. We have five-point fuzzy membership function by fuzzy valued three-point for the question and fuzzy hypothesis test the membership function by 95% confidence interval.

Key Words : Likert scale, confidence interval, fuzzy hypotheses testing, vague values

1. 서 론

리커트 척도란 대표적인 응답자 중심의 척도화 방법인 서열척도이며, 질문은 사실에 대한 판단보다는 개인의 가치를 묻는 것을 중심으로 여러 개의 문항으로 응답자의 태도를 측정하고 해당항목에 대한 측정치를 합산하여 평가 대상자의 태도점수를 얻어내는 척도이다.

리커트 척도의 구성절차로써는 첫 번째로 조사자가 연구하고 싶은 어떤 문제에 관한 긍정적(호의적) 문항 및 부정적(비호의적)문항들을 선정하고, 두 번째로 반응 카테고리(응답범주)를 작성(보통 5개 정도)한다. 세 번째로 많은 응답자에게 각 문항에 자기의 의견에 부합되는 카테고리에 체크하게 하여 각 문항에 대한 응답을 받아낸다. 네 번째로 각 문항에 대한 응답자의 반응을 점수로 환산(보통의 5점 척도일 경우 가장 우호적인 답을 5점 가장 비우호적인 답을 1점으로)한다. 다섯 번째로 문항간의 내적 일관성 및 상관성을 알아보고 식별능력이 있는 문항만을 선택한다. 마지막으로 척도로써 타당한 항목을 가지고 다시 응답자의 총점을 구하면 그것이 응답자의 태도 측정치이다.

각 항목에 대한 응답자들의 태도의 강도가 응답자들마다 다른데도 같은 값을 가질 수 있고 어떤 목적 변인을 설명하기 위해 여러 질문을 통합하여 분석하는데 이 작업과정에서 각 질문에 포함되어 있는 특성이 목적 변인을 설명하는데 있어 버려 지는 경향이 있다. 이러한 특성을 살리고 응답자 개인의 응답을 될 수 있으면 모두 취하는 방법으로 질문을 목적을 정보의 손실을 최소화시키면서 변인 값으로 변화시키는 과정에서 각 응답자별로 질적인 속성을 양적인 계열로 전환하여 측정하는 방법으로서 설문의 최소 문항으로서 최대의 효과를 나타내기 위하여 3점 척도의 애매한 응답을 퍼지화하여 5점 척도의 멤버쉽함수로 나타낸 후 한 항목을 n번 측정하여 평균값과 95% 신뢰구간을 퍼지수로 한 데이터들을 리커트 척도로 활용하여 내분비방법에 의하여 검정을 하였다. 제2장에서는 기본적인 계산에 필요한 퍼지확률통계량에 대하여 알아보고, 제3장에서는 퍼지리커트 척도의 구성을 보였으며, 제4장에서 그 예를 들었다.

접수일자 : 2005년 5월 30일
완료일자 : 2005년 8월 25일

2. 퍼지확률 통계량

정의 2.1. 전체집합 E상의 퍼지집합 A에 대하여 $\forall x \in E$일 때, 퍼지집합 A의 퍼지멤버쉽함수(fuzzy membership function)는

$$\mu_A(x) \in (0, 1]$$

이다.

여기서, 퍼지집합 $A \subset R$에 대하여 $A_\delta = [a_1^{(\delta)}, a_2^{(\delta)}]$이 $\delta \in [0,1]$에서의 δ-수준집합(δ level set)일 때,

$$A_\delta = \{x \mid \mu_A(x) \geq \delta\}, \quad \delta \in [0,1]$$

를 만족하면, 퍼지집합 A는 convex fuzzy set이 된다.

또 멤버쉽함수의 최대값이 1이 되는 원소x가 존재하는

퍼지집합 즉,

$$\vee_{n}\,\mu_A(x) = 1$$

이 되면 퍼지집합 A는 normal이 된다.

퍼지표본을 $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$으로 두면 퍼지평균 $\overline{\tilde{x}}$, 퍼지분산 $\tilde{s}^2$, 퍼지표준편차 $\tilde{s}$ 등은 일반통계와 유사한 방법으로 나타내며, 이러한 모든 값들은 일반통계에서는 다루지 않은 오차가 포함된 구간자료의 연산에 사용될 수 있다.

정의 2.2. 퍼지수 $\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n \subset R$가 각각 δ-level set, $\tilde{x}_{i_\delta} = [\,C_L(\tilde{x}_i)_\delta,\ C_R(\tilde{x}_i)_\delta]$ 을 가질 때, 퍼지표본평균은

$$\overline{\tilde{x}}_\delta = \left[\frac{1}{n}\sum_{i=1}^{n} C_L(\tilde{x}_i)_\delta,\ \frac{1}{n}\sum_{i=1}^{n} C_R(\tilde{x}_i)_\delta\right] \quad (2\text{-}1)$$

와 같이 얻을 수 있고, 퍼지표본분산과 퍼지표본표준편차 또한 유사한 방법으로 얻을 수 있다.

$$\tilde{s}^2_\delta = [\,C_L(g(\tilde{x}_{i_\delta})),\ C_R(g(\tilde{x}_{i_\delta}))] \quad (2\text{-}2)$$

$$\tilde{s}_\delta = [\,C_L(\sqrt{g(\tilde{x}_{i_\delta})}),\ C_R(\sqrt{g(\tilde{x}_{i_\delta})})] \quad (2\text{-}3)$$

여기서

$$g(\tilde{x}_{i_\delta}) = g(\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n) = \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{x}_i \ominus \overline{\tilde{x}})^2$$

이다.

이와 같은 퍼지통계량들을 이용하여 퍼지가설검정을 할 때 검정역과 신뢰구간을 구할 수 있다.

3. 퍼지 리커트 척도의 구성

3.1 리커트 5점척도의 fuzzy화

일반적인 5점 척도에 대한 응답의 애매성을 멤버쉽함수로 표현하면 다음 [그림1]로 나타낼 수 있다.

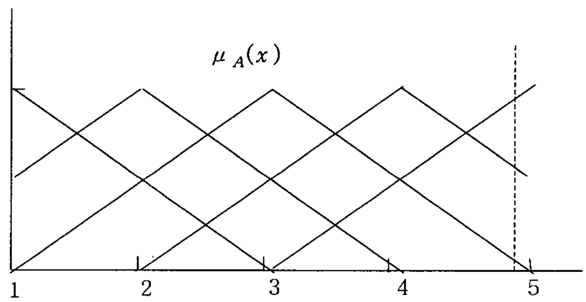

그림 1

각 membership함수에 대한 δ-level의 퍼지수는

$$[\tilde{x}_1]^\delta = [1,\ 3-2\delta] \qquad \delta \in (0,\ 1]$$

$$[\tilde{x}_2]^\delta = \begin{cases} [1,\ 4-2\delta] & \delta \in (0,\ 0.5] \\ [2\delta,\ 4-2\delta] & \delta \in (0.5,\ 1] \end{cases}$$

$$[\tilde{x}_3]^\delta = [2\delta+1,\ 5-2\delta] \qquad \delta \in (0,\ 1]$$

$$[\tilde{x}_4]^\delta = \begin{cases} [2\delta+2,\ 5] & \delta \in (0,\ 0.5] \\ [2\alpha+2,\ 6-2\alpha] & \delta \in (0.5,\ 1] \end{cases}$$

$$[\tilde{x}_5]^\delta = [2\delta+3,\ 5] \qquad \delta \in (0,\ 1]$$

3.2 3점척도를 fuzzy화한 5점 fuzzy 척도

질문의 항을 최소로 한 3점 척도에서 응답의 애매성을 멤버쉬함수로 표현하면 다음그림과 같이 5점 척도로 나타낼 수 있다.

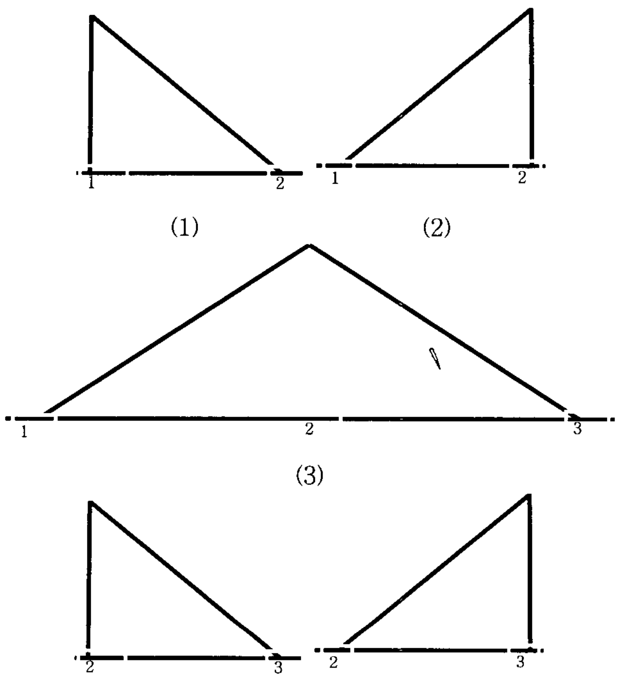

그림 2

각 항목을 선택했을 경우에 대한 멤버쉽함수와 δ-level fuzzy number는 다음과 같이 표현된다.

(1)을 선택한 경우:

$$\mu_A(x) = \begin{cases} -x+2, & (1,2) \\ 0, & \text{그외}, \end{cases} \quad (3\text{-}1)$$

$$[\tilde{x}]^\delta = [1, -\delta+2], \qquad \delta \in (0,1].$$

(2)를 선택한 경우:

$$\mu_A(x) = \begin{cases} x-1, & (1,2) \\ 0, & \text{그외}, \end{cases} \quad (3\text{-}2)$$

$$[\tilde{x}]^\delta = [\delta+1, 2], \qquad \delta \in (0,1].$$

(3)을 선택한 경우:

$$\mu_A(x)=\begin{cases} x-1, & (1,2) \\ -x+3, & (2,3), \end{cases} \tag{3-3}$$

$$[\tilde{x}]^{\delta}=\begin{cases} [\delta+1,2] \\ [2,-\delta+3]. \end{cases}$$

(4)을 선택한 경우:

$$\mu_A(x)=\begin{cases} -x+3, & (2,3) \\ 0, & \text{그외}, \end{cases} \tag{3-4}$$

$$[\tilde{x}]^{\delta}=[2,-\delta+3].$$

(5)를 선택한 경우:

$$\mu_A(x)=\begin{cases} x-2, & (2,3) \\ 0, & \text{그외}, \end{cases} \tag{3-5}$$

$$[\tilde{x}]^{\delta}=[\delta+2,3].$$

위의 결과로 검정통계량의 모평균은 아래의 퍼지수로 주어진다.

$$[\frac{3}{5}\delta+1.4,\ -\frac{3}{5}\delta+2.6].$$

3.3 한 항목에 대한 멤버쉽함수

한 항목에 대한 K번 질문의 경우의 5점 척도에 대하여 l=Min(5점 점수), r=Max(5점 점수), $c=\frac{1}{k}\Sigma$(각 점수)로 두면 응답에 대한 퍼지수는 (l, c, r)로 되어 각 경우에 대한 멤버쉽함수는

$$\mu_A(x)=\begin{cases} \frac{1}{c-l}x-\frac{l}{c-l}, & l\le x\le c \\ -\frac{1}{r-c}x+\frac{r}{r-c}, & c\le x\le r \end{cases} \tag{3-6}$$

이며 δ-level set는

$$[\tilde{x}]^{\delta}=[(c-l)\delta+l,\ -(r-c)\delta+r]$$

이다.

(3-6)의 결과로 한 항목에 대한 응답 결과는 다음의 [그림 3], [그림4]와 같이 표현된다.

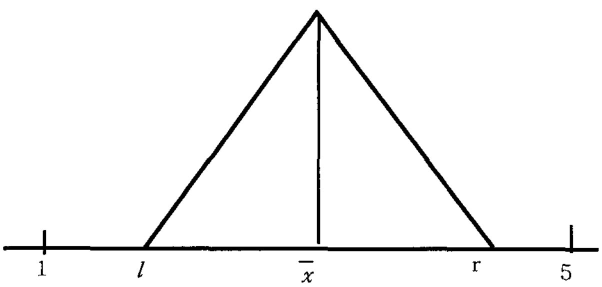

그림 3

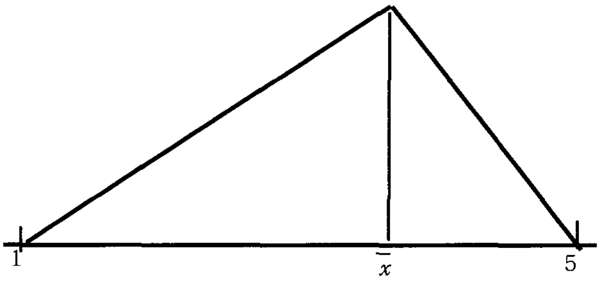

그림 4

일반적으로 5점 리커트 척도는 설문지의 한 질문에 대해 가장 우호적인 답을 5점에서부터 가장 비우호적인 답에 대해 1점까지 부여한다. 이러한 문항들 중 어떤 하나의 가치 척도를 파악하기 위해 하나의 문항보다는 신뢰도 검정을 통해 비슷한 여러 개의 문항에 대해 평균이나 합을 가지고 하나의 가치를 판단하게 되는데 우리는 이러한 경우 한 사람이 어떤 가치에 대해 판단하는 것이 절대적일 수 없을 것이라는데 착안하여 각 사람들의 어떤 가치에 대해 한 문항을 k번 측정한 평균값과 사분위범위(Q_1, Q_3)를 사용하여 [그림5]와 같이 표현할 수 있고, 95%신뢰구간을 *level* 0 의 한 구간으로 보고 삼각 퍼지수라 가정할 때 한 응답자에 대한 한 처지수는 [그림6]와 같다.

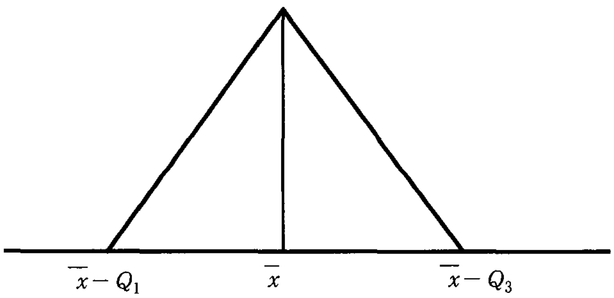

그림 5

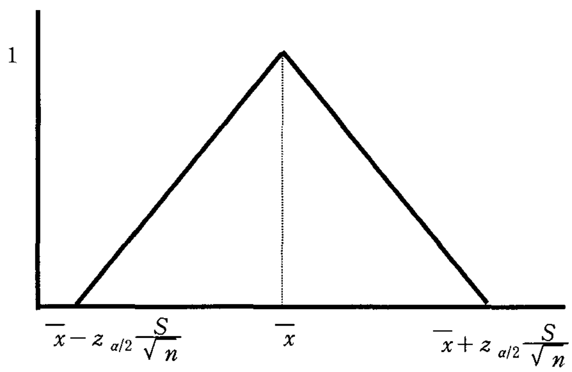

그림 6

각 개개인의 n개 항목 중 i번째 항목과 j번째 설문자에 대한 처치수의 멤버쉽함수는 다음과 같다.

$$\mu_A(x_{ij}) = \begin{cases} \dfrac{x_{ij} - \overline{x_i}}{z_{\alpha/2}\frac{s}{\sqrt{n}}} + 1, & (\overline{x_i} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{x_i}) \\ 1 + \dfrac{\overline{x_i} - x_{ij}}{z_{\alpha/2}\frac{s}{\sqrt{n}}}, & (\overline{x_i}, \overline{x_i} + z_{\alpha/2}\frac{s}{\sqrt{n}}) \end{cases}$$

$$i = 1, \cdots, n, \quad j = 1, \cdots, m. \qquad (3\text{-}7)$$

리커트 척도의 경우 응답자 개개인의 우호정도를 평균에 의해 표현하게 되는데 이 정도가 응답자 개개인에 따라 다르나 응답자 개개인의 우호성의 표현 범위는 모두 같게 된다. 그러나 위의 식과 같이 각 응답자의 신뢰구간을 구간으로 주는 퍼지 점수화를 하게 되면 같은 평균을 가지는 응답자라도 수준별 차이를 가지게 되므로 보다 정확한 응답자의 우호성 정도를 파악 할 수 있다.

그래서 그 응답이 매우 비우호적이라 하더라도 응답자 별로 수준의 차이가 있을 것이라 보고 보다 비우호적이면서 응답에 매우 비우호적으로 답할 경우에 대해서도 수준별 기대치를 주는 형식으로 점수를 위의 식과 같이 퍼지멤버쉽함수로 하였다.

각 자료에 대한 퍼지 수 $\mathcal{X}_i$의 δ-level 집합은 $\delta \in (0,1]$에서 다음과 같다.

$$[\mathcal{X}_i]^{\delta} = [[\overline{x}_i - z_{\alpha/2}\frac{s}{\sqrt{n}}]^{\delta}, [\overline{x}_i + z_{\alpha/2}\frac{s}{\sqrt{n}}]^{\delta}],$$

$$i = 1, \cdots, n \qquad (3\text{-}8)$$

로 된다.

4. 예 제

2004년 모 대학의 간호학과의 학위논문 발표에서 환자의 병원 치료비 등에 대한 보호자의 부담감을 알기위하여 설문자료로서 리커트 척도(5점)대한 55명의 남성에 대한 자료와 117명의 여성에 대한 자료를 다음 데이터로서 얻었다. 과연 남성(M)과 여성(F)의 보호자 부담감에 대한 차이가 존재하는가에 대한 결과를 퍼지 자료화하여 양측검정으로 분석한 결과 다음과 같다.

<data>

ID	성별	문1	문2	…	문47	min	평균	max			평균	
1	2	2	3	…	1	1	2.49	4		2.28	2.49	2.70
2	2	2	3	…	1	1	1.83	4		1.61	1.83	2.05
3	1	3	2	…	4	1	2.94	5		2.70	2.94	3.17
4	2	3	3	…	3	2	3.15	5		2.95	3.15	3.35
5	2	4	3	…	1	1	2.49	4		2.25	2.49	2.73
6	1	4	2	…	1	1	2.79	5		2.43	2.79	3.15
⋮	⋮	⋮	⋮	⋱	⋮	⋮	⋮	⋮		⋮	⋮	⋮
171	2	3	3	…	3	1	3.04	5		2.81	3.04	3.27
172	1	3	4	…	4	2	2.98	4		2.80	2.98	3.16
173	2	4	4	…	4	2	3.40	4		3.23	3.40	3.58
174	2	3	3	…	2	2	3.21	5		3.01	3.21	3.42

한 항목에서 퍼지화된 자료의 평균은 $\delta \in (0,1]$에서 식(2-1)을 사용하여

$$\overline{\mathcal{M}} = (2.623 + 0.219\delta,\ 3.061 - 0.219\delta) \qquad (4\text{-}1)$$

$$\overline{\mathcal{F}} = (2.810 + 0225\delta,\ 3.260 - 0.225\delta) \qquad (4\text{-}2)$$

을 얻었으며,

남, 여의 분산은 $\delta \in (0,1]$에서 식(2-2)를 사용하여

$$\tilde{s}^2_M = (0.199\delta^2 - 0.006\delta + 0.138,\ 0.199\delta^2 - 0.515\delta + 0.647) \qquad (4\text{-}3)$$

$$\tilde{s}^2_F = (0.209\delta^2 - 0.047\delta + 0.101,\ 0.209\delta^2 - 0.592\delta + 0.646) \qquad (4\text{-}4)$$

로 나타났다.

그러므로 남성의 평균에 대한 95% 신뢰구간은 $\delta \in (0,1]$에서

$$\begin{cases} \delta = 0 & [2.525,\ 3.273] \\ \delta = 0.5 & [2.619,\ 3.126] \\ \delta = 1 & [2.690,\ 2.994] \end{cases} \qquad (4\text{-}5)$$

을 가진다.

퍼지 자료의 검정법은 신뢰구간과 검정통계량의 퍼지성에 따라 동의지수법 등 여러 가지가 있으나 검정통계량이 신뢰구간에 포함되어 귀무가설의 기각과 채택 정도는 [그림 6]과 등 같이 나타낸다. 또한 내분비를 사용한 검정법은 본 연구에서 정의4.1과 같이 정의하여 사용한다.

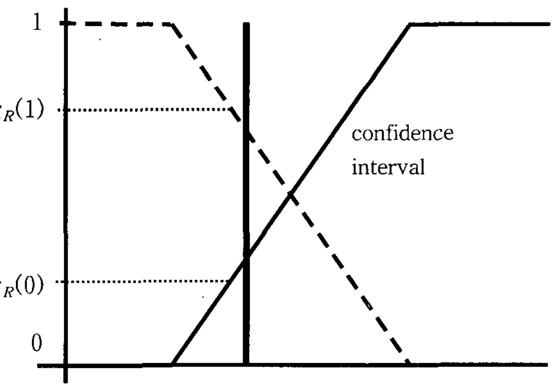

그림 7

정의 4.1. 퍼지 가설 H_f를 검정하기 위해 모수 θ에 대한 퍼지가설멤버쉽함수 $\mu_{H_f}(\theta)$와 퍼지 자료에 의해 구해진 퍼지 신뢰구간 $\mathcal{T}(\delta, \theta)$에 대하여 δ-level set이 존재하고 퍼지수 A의 δ-level set에 대한 길이를

$$L(A)_{\delta} = A_R^{(\delta)} - A_L^{(\delta)} \qquad (4\text{-}6)$$

정의하면, δ-level 에서 귀무가설의 채택정도를

$$R_{(\alpha,\delta)}(0)=\begin{cases}\dfrac{L(H_f^{(\delta)}\cap T^{(\delta)})_\delta}{L(H_f^{(\delta)})_\delta} & ,\ \theta|r_\alpha(\theta)=0\neq\varnothing \\ 0 & ,\ \theta|r_\alpha(\theta)=0=\varnothing\end{cases} \qquad (4\text{-}7)$$

로 두며, 기각 정도는

$$R_{\alpha,\delta}(1)=1-R_{\alpha,\delta}(0) \qquad (4\text{-}8)$$

이다.

내분비를 이용한 가설검정을 실시하면 식 (4-7)에 의해 δ-level이 0.9인 경우는

$$R_{\alpha,0.9}(0)=\frac{0.088-0.015}{0.088}=0.830$$

이 구해진다

그러므로 $\delta=0.9$ 에서 가설을 채택할 정도는 0.830이 되어 [그림 8]로 표현된다.

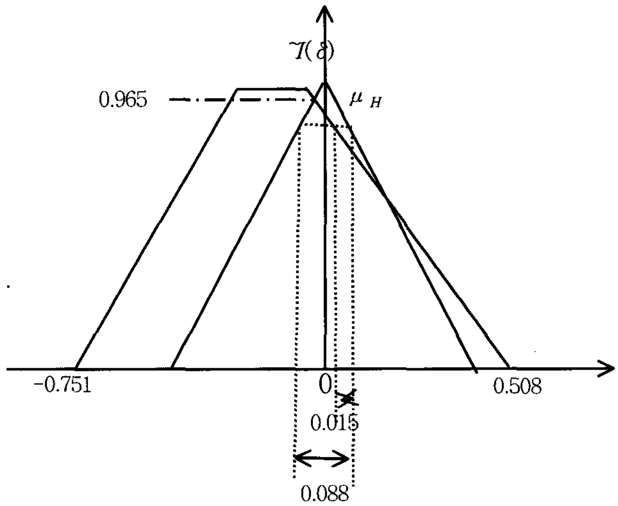

그림 8

참 고 문 헌

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

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

[3] M. K. Kang, J. R. Choi and E. S. Bae, Some properties of fuzzy sample correlation coefficient with fuzzy data, *Far East J. Theo. Stat.* 4(1)(2000), 165-177.

[4] 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)

[5] C. E. Lee, M. K. Kang and G. T. Choi and, Testing of Fuzzy Likert Scale, *Proceeding of Korea Fuzzy Logic and Intelligent Systems Society Fall Conference,* Vol 14, Num. 2, (2004)

[6] F. Li, C. Wu, J. Qiu and L. Su, Platform type fuzzy number a separability of fuzzy number space, *Fuzzy Sets and systems* 117(2000), 347-353.

[7] D. Pill and K. Peter, *Metric Space of Fuzzy Sets,* World Scientific, Singapore, 1994.

[8] N. Watanabe, T. Imaizumi, A Fuzzy Statistical Test of Fuzzy Hypotheses, *Fuzzy Sets and system* 53 (1993), pp.167-178.

저 자 소 개

강만기(Man-Ki Kang)
현 동의대학교 정보통계학과 교수

관심분야 : Vague data processing
Phone : 051-890-1482
E-mail : mkkang@deu.ac.kr

이창은(Chang-Eun Lee)
현 동의대학교 정보통계학과 강사

관심분야 : 퍼지 이론

최규탁(Gyu-Tag Choi)
현 경남정보대 산업시스템경영과 교수

관심분야 : 퍼지선형계획법