한국지능시스템학회 논문지 2011, Vol. 21, No. 1, pp. 25-28
DOI : 10.5391/JKIIS.2011.21.1.25

균일최강력검정에 의한 가설의 퍼지 검정

Fuzzy Test of Hypothesis by Uniformly Most Powerful Test

강만기
Man Ki Kang

Dept. of Data Information Science, Dong-eui University, Busan 614-714, Korea

요 약

본 연구는 퍼지가설 검정을 위하여 퍼지수로 주어진 데이터에 대한 데이터의 조건을 제시하였고 면적비에 의한 동의지수법을 정의한 후 균일최강력 퍼지검정법을 제안하였다. 또한 균일최강력 퍼지 검정을 위하여 유의수준에 의한 신뢰한계를 제시하였다. 예증으로서 지수분포에서 얻은 랜덤샘플을 우도비에 의한 최강력 기각역을 구하여 동의지수법을 이용한 카이제곱분포의 퍼지검정법을 제시하였다.

키워드 : 동의지수, 퍼지기각역, 균일최강력검정, 신뢰한계, 우도함수.

Abstract

In this paper, we study some properties of condition for fuzzy data, agrement index by ratio of area and the uniformly most powerful fuzzy test of hypothesis. Also, we suggest a confidence bound for uniformly most powerful fuzzy test. For illustration, we take the most powerful critical fuzzy region from exponential distribution by likelihood ratio and test the hypothesis of χ^2-distribution by agreement index.

Key Words : agreement index, fuzzy critical region, uniformly most powerful test, confidence bound, likelihood function.

1. Introduction

In many real situations an uncertainty of data comes from randomness and fuzziness. Randomness models the stochastic variability of all possible outcomes of an experiment and fuzziness describes the vagueness of a given outcome.

From the random fuzzy samples, the fuzzy negation of the assertion is taken to be the fuzzy null hypothesis H_{f0} and the fuzzy assertion itself is taken to be the fuzzy alternative hypothesis H_{f1}([1],[2],[6]).

Kang, Choi and Lee[3] and Kang, Lee, and Han[4] defined fuzzy hypotheses membership function. Also, they found the agreement index by area ratio for the fuzzy hypotheses membership function against the membership function of a fuzzy critical region. Thus, they obtained the results by the grade for judgement of acceptance or rejection degree for the fuzzy hypotheses.

First, we define the condition of fuzzy data from a fuzzy random sample and an agreement index. Next, we show some properties of the uniformly most powerful fuzzy test for the fuzzy data.

The properties of acceptance fuzzy region and confidence bound by significance level was introduced in Chap. 4.

Finally, we give an example for the uniformly most powerful fuzzy test by exponential distribution.

The fuzzy hypothesis is constructed by a set

$$\{(H_{f0}(\theta), H_{f1}(\theta)) | \theta \in \Omega\} \tag{1.1}$$

with membership function $m_{H_f}(H_{f0}(\theta), H_{f1}(\theta))$, where θ is parameter vector and Ω is some sample space with fuzzy numbers.

The fuzzy hypothesis H_{f0} is referred to as the null fuzzy hypothesis while H_{f1} is referred to as the alternative fuzzy hypothesis.

We consider the fuzzy hypothesis

$$H_{f0}: \theta \simeq \theta_0 \text{ or } H_{f0}: \theta \prec \theta_0, \ \theta \in \Omega \tag{1.2}$$

where "$\simeq$" is desire as follows for crisp null hypothesis $H_0: \theta = \theta_0$ against to crisp alternative hypotheses

(a1) $H_1: \theta \neq \theta_0$

(a2) $H_1: \theta > \theta_0 + \Delta$ or $H_1: \theta < \theta_0 - \Delta$ for $\Delta > 0$

and "$\prec$" desire as follows for crisp null hypothesis $H_0: \theta < \theta_0$ against to crisp alternative hypotheses

(b1) $H_1: \theta = \theta_0$

(b2) $H_1: \theta > \theta_0 + \Delta$ for $\Delta > 0$.

접수일자 : 2010년 8월 5일
완료일자 : 2011년 1월 30일

A fuzzy data A of a random sample in R is said to be convex if for any real numbers $x, y, z \in R$ with $x \leq y \leq z$,

$$m_A(y) \geqq m_A(x) \wedge m_A(z) \qquad (1.3)$$

with $\wedge$ standing for minimum.

A fuzzy random sample data A is called normal if the following holds

$$\bigvee_x m_A(x) = 1 \qquad (1.4)$$

A $\delta-$level set of a fuzzy random sample data A is a fuzzy number denoted by $[A]^\delta$ and is defined as

$$[A]^\delta = \{x | m_A(x) \geqq \delta, 0 < \delta \leqq 1\} \qquad (1.5)$$

A $\delta-$level set of sample fuzzy number A is also a convex fuzzy set which is a closed, bounded interval and denoted by $[A]^\delta = [A_l^\delta, A_r^\delta]$.

2. Agreement index

Let X be a random variable by random fuzzy sample from samples fuzzy space Ω and $\{P_\theta,\ \theta \in \Omega\}$ is a family of fuzzy probability distributions, where θ is a parameter vector of Ω.

Choose a membership function $m_\Theta(x)$ whose value is likely to best reflect the plausibility of the fuzzy hypothesis being tested, where x is observation of X.

Let us consider the membership function $m_C(x)$ of critical region C, which we will call the agreement index of $m_\Theta(x)$ with regard to $m_C(x)$([3],[4]).

Definition 2.1. Let a fuzzy membership function $m_\Theta(x)$, $x \in R$, and consider another membership function $m_C(x)$, $x \in R$, then we have an agreement index by the ratio being defined in the following way;

$$m_{AI(\theta)} = \frac{area(m_\Theta(x) \cap m_C(x))}{area(m_C(x))} \in [0,1]. \qquad (2.1)$$

Definition 2.2. We define a grade membership function of rejection or acceptance degree by agreement index to a real-valued function $0 \leq R_\delta(\cdot) \leq 1$ by $\delta-$level as respectively,

$$\Re_\delta(0) = \sup_\theta \left\{ \frac{\text{area } (\text{m}_{\Theta_\delta}(\text{x}) \cap \text{m}_{\text{C}_\delta}(\text{x}))}{\text{area } \text{m}_{\text{C}_\delta}(\text{x})} \right\}, \qquad (2.2)$$

$$\Re_\delta(1) = 1 - \Re_\delta(0) \qquad (2.3)$$

for the fuzzy test of hypothesis.

3. Uniformly most powerful fuzzy test

Let random fuzzy sample $X_1, X_2, \cdots, X_n$ have joint pdf(probability density function) $f(x_1, \cdots, x_n)$, consider a critical region C.

Definition 3.1. For $\theta \in \Omega$, consider a hypotheses of the form $H_{f0} : [\theta]_p^{(\delta)} \in \Omega_0$ versus $H_{f1} : [\theta]_p^{(\delta)} \in \Omega - \Omega_0$, where $[\theta]_p^{(\delta)}$ is any one point of θ by $\delta-$level and $\Omega_0 \subset \Omega$, then we have fuzzy hypotheses of the form

$$H_{f0} : \theta \in \Omega_0 \text{ versus } H_{f1} : \theta \in \Omega - \Omega_0 \qquad (3.1)$$

by resolution identity.

If a fuzzy test is the most powerful against fuzziness for every possible value in a composite alternative, then it will be called the uniformly most powerful fuzzy test.

Definition 3.2. A critical region C^*, and the associated test, are said to be fuzzy uniformed most powerful(UMP)of size significant level α if

$$Max_{[\theta]_p^{(\delta)} \in \Omega_0} P_{C^*}([\theta]_p^{(\delta)}) = \alpha \qquad (3.2)$$

and

$$P_{C^*}([\theta]_p^{(\delta)}) \geq P_C([\theta]_p^{(\delta)}) \qquad (3.3)$$

for all $\theta \in \Omega - \Omega_0$ and all critical region C and size α.

From Definition 3.2, critical fuzzy region C^* for θ defines a uniformly most powerful fuzzy test(UMPFT) of size significance level α by resolution identity.

The theory of UMP one-side fuzzy tests can be applied to the problem of obtaining a lower and upper bound for fuzzy-valued parameter θ.

For example, θ is the breaking strength of a new alloy, some toxicity of a drug or the fuzzy probability have an error term[5] of an undesirable fuzzy event.

4. Confidence bounds

Since $\theta_0(X)$ will be a function of the observed random fuzzy samples x_i, $i = 1, 2, \cdots$, it cannot be required to fall below θ with uncertainty. One selects a number $1-\alpha$ is *significance lever* and restricts attention to bounds θ satisfying

$$P_\theta\{\theta_0([X]_p^{(\delta)}) < [\theta]_p^{(\delta)}\} > 1 - \alpha \qquad (4.1)$$

for any one point of θ.

The function $[\theta_0(X)]_p^{(\delta)}$ is lower *confidence bound* than the function $\theta_0(X)$ if called a lower fuzzy *confidence bound* for θ at significance level $1-\alpha$; the infimum of the left-hand side of the equation (4.1), which in practice will be near equal to $1-\alpha$, is called *significance coefficient* of θ.

For example, the fuzzy probability of θ falling below any $\theta' < \theta$ should be infimum as $P_\theta\{\theta_0(X) < \theta'\}$.

Thus, for all $\theta' < \theta$ subject to equation (4.1) is a uniformly most lower fuzzy confidence bound for θ at significance level $1-\alpha$.

If we have *confidence interval*[2] $S([X]_p^{(\delta)})$ for θ, a family of fuzzy subset $S(X)$ of the parameter space Ω is said to constitute a family of confidence fuzzy sets at significance level $1-\alpha$ if

$$P_\theta\{\theta \in S(X)\} > 1-\alpha \qquad (4.2)$$

for all $\theta \in \Omega$.

A lower fuzzy confidence bound corresponds to the special case that $S(x)$ is a one-side interval $S(x)=\{\theta : \theta(x) < \theta\}$ where x is observation of X.

Theorem 4.1. (1) For each $\theta \in \Omega$, let $A(\theta)$ be an *acceptance fuzzy region* of a *significance level* α fuzzy testing for

$$H_{f0}: \theta \simeq \theta_0 \text{ versus } H_{f1}: \theta \simeq \theta_1,$$

where $\theta_0 \in \Omega_0$ and $\theta_1 \in \Omega - \Omega_0$. For each sample fuzzy point x, let $S(x)$ denote the fuzzy set of parameter values as

$$S(x)=\{\theta : x \in A(x), \theta \in \Omega\}, \quad (4.3)$$

then $S(x)$ is a family of *confidence fuzzy sets* for θ at significance level $1-\alpha$.
(2) If $A(\theta_0)$ is a UMP for fuzzy testing H_{f0} at significance level α against the alternatives H_{f1}, then $S(X)$ minimizes the fuzzy probability

$$P_\theta\{\theta' \in S(X)\} \text{ for all } \theta \in \Omega_0 \qquad (4.4)$$

among all level $1-\alpha$ families of confidence fuzzy sets for θ.

Proof. (1) By definition of $S(x)$, $[\theta]_p^{(\delta)} \in S([x]_p^{(\delta)})$ if and only if $[x]_p^{(\delta)} \in A([\theta]_p^{(\delta)})$, and all level $1-\alpha$ family of confidence fuzzy sets for θ by resolution identity, we have

$$P_\theta\{\theta \in S(X)\} = P_\theta\{X \in A(\theta)\} > 1-\alpha \qquad (4.5)$$

(2) If $S^*(x)$ is any other family of confidence sets at level $1-\alpha$,
if

$$A^*([\theta]_p^{(\delta)}) = \left\{[x]_p^{(\delta)} : [\theta]_p^{(\delta)} \in S^*([x]_p^{(\delta)})\right\}$$

then

$$P_\theta\left\{[X]_p^{(\delta)} \in A^*([\theta]_p^{(\delta)})\right\} = P_\theta\left\{[\theta]_p^{(\delta)} \in S^*([X]_p^{(\delta)})\right\} > 1-\alpha$$

so that $A^*(\theta_0)$ is the *acceptance fuzzy region* of a level α test for H_{f0}. It follows from the property of $A(\theta_0)$ that for any θ of $\Omega - \Omega_0$,

$$P_\theta\left\{X \in A^*(\theta_0)\right\} > P_\theta\left\{X \in A(\theta_0)\right\} \qquad (4.6)$$

and hence that

$$P_\theta\left\{\theta \in S^*(X)\right\} > P_\theta\left\{\theta_0 \in S(X)\right\} \qquad (4.6)$$

The equivalence equation (4.5) shows the structure of the fuzzy confidence sets $S(x)$.

We exhibit the values for which the hypothesis is accepted $(\theta \in S(x))$ and those for which it is rejected $(\theta \in S(x))$.

5. Illustration

Consider a fuzzy random sample of size n from an exponential distribution,

$$X_i \sim \text{EXP}(\theta), \; i=1,\cdots,n. \qquad (5.1)$$

We have likelihood function to reject H_{f0} if

$$L(x,\theta_0,\theta_1) = \frac{\theta_0^{-n}\exp(-\sum x_i/\theta_0)}{\theta_1^{-n}\exp(-\sum x_i/\theta_1)} < k \qquad (5.2)$$

where k is such that $P[L(x;\theta_0,\theta_1) < k \mid \theta_0] \simeq \alpha$, under $\theta \simeq \theta_0$.

Now,

$$P[\ln L(x;\theta_0,\theta_1) < \ln k|\theta_0] = P[\sum X_i(\frac{1}{\theta_1}-\frac{1}{\theta_0}) < \ln((\theta_0/\theta_1)^n k)|\theta_0]$$

so that

$$P[X \in C^*|\theta_0] = P[\sum X_i > k_1|\theta_0] \qquad (5.3)$$

where $k_1 = \ln((\theta_0/\theta_1)^n k)/(\frac{1}{\theta_1}-\frac{1}{\theta_0})$.

Thus, the most powerful critical fuzzy region has the form

$$C^* = \left\{(x_1,\cdots,x_n)|\sum x_i > k_1\right\} \qquad (5.4)$$

Notice that the under $H_{f0}: \theta \simeq \theta_0$, we have $2\sum x_i/\theta_0 \sim \chi^2(2n)$ by the property of exponential distribution and χ^2 −distribution so that $k_1 = \theta_0 \chi^2_{1-\alpha}(2n)/2$ would give a fuzzy critical region of size α.

The most powerful test[4] of size α of $H_{f0}: \theta \simeq \theta_0$ versus $H_{f1}: \theta \simeq \theta_1$, when $\theta_1 > \theta_0$, we can reject H_0 by agreement index of the equation (2.2) and (2.3) if

$$2n\bar{x}/\theta_0 = 2\sum x_i/\theta_0 > \chi^2_{1-\alpha}(2n) \qquad (5.5)$$

This is does not depend on the value of θ_1. Thus we can find a UMP fuzzy test of $H_{f0}: \theta \simeq \theta_0$ versus $H_{f1}: \theta > \theta_1$ for $\theta_1 > \theta_0$.

Thus, we can seek fuzzy power function by χ^2 −cdf(cumulative distribution function), with degree of freedom $\nu = 2n$ as

$$P(\theta) = 1 - F[(\theta_0/\theta)\chi^2_{1-\alpha}(\nu);\nu] \qquad (5.6)$$

Since $P(\theta)$ is the increasing function of θ,

$$sup_{\theta < \theta_0} P(\theta) \simeq P(\theta_0) = \alpha,$$

and the fuzzy test is also a UMP fuzzy test of significance size α for the composite fuzzy hypotheses

$$H_{f0} : \theta < \theta_0 \text{ versus } H_{f1} : \theta > \theta_1.$$

Similarly, a UMP fuzzy test of either

$$H_{f0} : \theta \simeq \theta_0 \text{ or } H_{f0} : \theta > \theta_0$$

versus $H_{f1} : \theta < \theta_0$ is to reject H_{f0} if $2n\overline{x}/\theta < \chi^2_{\alpha}(\nu)$.

The associated fuzzy power function is

$$P(\theta) = F[(\theta_0/\theta)\chi^2_{\alpha}(\nu);\nu] \quad (5.7)$$

For example, if we observed some lifetime x_i of 40 electrical parts, and it was conjectured that those observations might be exponentially distributed with fuzzy mean number time $\theta = [85,90,95]$ months.

Suppose in a particular application, that the part will be unsuitable if the mean is less than $\theta = [85,90,95]$.

We carry out a significance level size $\alpha = 0.05$ for fuzzy test of

$$H_{f0} : \theta > \theta_0 \text{ versus}$$
$$H_{f1} : \theta < \theta_0.$$

For the sample x_i, if we observed the$\theta_0(x)$ is fuzzy mean number as $\overline{x} = [66.5,70.2,72.5]$ months by three times of random experiment observations, then consequently

$$\begin{aligned} 2n\overline{x}/\theta &= (80)([66.5,70.2,72.5])/[85,90,95] \\ &= [56,62.4,68] > 60.39 = \chi^2_{0.05}(80), \end{aligned} \quad (5.8)$$

where 0.05 is significance level.

By agreement index of the equation (2.2), we have rejection degree $R(0) = 0.78$ for the fuzzy null hypothesis H_{f0} at the $\alpha = 0.05$ level of significance.

Suppose that we wish to know type II error $P[TII]$, since the mean is $\theta_0(x) = [43,45,47]$ months with fuzzy number, according to equation (5.6),

$$\begin{aligned} &P([43,45,47]) = F([109.3,120.78,132.8];80) \\ &= [0.9835,0.9978,0.9998] > P_{\theta}\{\theta_0 \in S(X)\} \end{aligned}$$

by χ^2-distribution. The type II error is given by

$$\begin{aligned} P[TII] &= 1 - F([109.3,120.78,132.8];80) \\ &= [0.0002,0.0022,0.0165]. \end{aligned}$$

6. Conclusions

We proposed some properties of a fuzzy hypothesis test for the uniformly most powerful test by agreement index, and suggested a confidence bound for the uniformly most powerful fuzzy test.

For illustration, we have the most powerful critical fuzzy region from exponential distribution of observed some lifetime for electrical parts by likelihood ratio.

Given the fuzzy lifetime mean, we test the hypothesis using χ^2−distribution by agreement index.

Thus, we show a fresh method for testing hypotheses of fuzzy data.

References

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

[2] M. K. Kang and C. E. Lee, "Fuzzy Confidence Interval of Linear Regression for Slop", *Far East Journal of Theoretical Statistics*. 8(1), 2002.

[3] M. K. Kang, Y. R. Park, 'Fuzzy Binomial Proportion Test by Agreement Index", *Journal of Korean Institute of Intelligent Systems,* Vol 19, Num. 1,pp.19-24, 2009.

[4] M. K. Kang, A. H. Seo, "Fuzzy Hypothesis Test by Poisson Test for Most powerful Test", *Journal of Korean Institute of Intelligent Systems,* Vol 19, Num. 6, pp.809-813, 2009.

[5] W. Trutschnig, "A Strong Consistency Result for Fuzzy Relative Frequencies Interpreted as Estimator for the Fuzzy-Value Probability", *Fuzzy Sets and Systems,* 159, pp 259-269, 2008.

[6] Z. Xia, "Fuzzy Probability System: fuzzy probability space(1)", *Fuzzy Sets and systems,* 120, pp.469-486, 2001.

저 자 소 개

강만기(ManKi Kang)
1990년 : 동아대학교, 이학박사
1989∼현재 : 동의대학교 데이터정보학과 교수

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