Proceedings of KFIS Autumn Conference 2007 Vol. 17, No. 2.

검정에 관한 퍼지 검정력 함수의 성질

The Fuzzy Power Function of a Test

강만기[1], 정지영[2], 박영례[2], 최규탁[3]

[1] 부산시 진구 동의대학교 데이터정보학과
E-mail: mkkang@deu.ac.kr
[2] 부산시 진구 동의대학교 데이터정보학과 대학원
[3] 부산시 사상구 경남정보대학 경영학부

Abstract

We introduction some properties for fuzzy power function of performance of a test. First we define fuzzy type I error and type II error for the probability of the two types of error. And we show that an fuzzy error probability of one kind can only be reduced at cost of increasing the other fuzzy error probability.

Key Words : power function, type I, II error, fuzzy probability, agreement index.

1. Introduction

A discussion of the formulation of a fuzzy statistical hypothesis testing problem and the steps for solving it requires the introduction of a number of definitions and concepts. Before proceeding with this topic in its full generality, we develop its basic ideas in terms of a specific problem in which the chance behavior is governed.

When an investigation is aimed at establishing an assertion with substantive support obtained from the sample, the negation of the assertion is taken to be the fuzzy null hypothesis H_0 and the assertion itself is taken to be the fuzzy alternative H_1.

In testing a null fuzzy hypothesis H_0 against an fuzzy alternative H_1, our attitude is to uphold H_0 as true unless the data speak strongly against it, in which case, H_0 should be rejected in favor of H_1 by degree of acceptance and rejection. Falsely rejecting H_0 is viewed as more serious error than failing to reject H_0 when H_1 is true.

Thus, we introduction some properties for fuzzy power function of performance of a test. First we define fuzzy type I error and type II error for the probability of the two types of error. And we show that an fuzzy error probability of one kind can only be reduced at cost of increasing the other fuzzy error probability.

2. Fuzzy probability

We denote by $x=x(s)$ the possible outcome of an F-random experiment of subject s and will be call the sample point.

Definition 2.1 The sample of an F-random experiment of subject s is a pair $(\Omega(s), \phi)$, where

(1) $\Omega(s)$ is the set of all possible outcomes of the F-random experiment of subject s

(2) ϕ is a σ-field of subjects of $\Omega(s)$.

In general, call $\Omega(s)$ the valuation set of subject s, which includes all possible observation result or includes the observation range of subject s. Let $\boldsymbol{P}(s)$ be the collection of all subjects of $\Omega(s)$; then it is the σ-field of subjects of $\Omega(s)$. So, ($\Omega(s)$, $P(s)$) is the sample space of subject s.

We must note that now we do not have any probability measure on sample space (Ω, ϕ), next let us discuss how one can construct a probability on (Ω, ϕ) by the sample of F-random experiment, satisfying the Kolmogorov axiom.

The F-random experiment can be repeated under identical fuzzy condition by one and the same subject s, for $k=1,2,\ldots$; we denote by

$E_k = E_k(s)$ the kth trial of subject s, and by $x_k = x_k(s)$

the sample point which is a result of E_k; we say that $E(s) = \{E_k(s)\}$

For each $A \in \mathbf{P}$, let $\chi_A(\omega)$ be given by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \qquad (2.1)$$

That is, it is the characteristic function (or indicator function) of events A and for each $x_k \in \Omega_m$, $\chi_A(x_k)$ have provided the time information of the occurrence of events A which has some uncertainty from the randomness of samples, Ω_m. According to Proposition 2.1, we believe that all the samples are the same. To provide the time information of occurrence of events A under identical conditions, take the weighted mean of all these $\chi_A(x_k)$, $1 \le k \le m$, which have the same weight, that is, for each $A \in P$, let

$$f(A, m, s) = \frac{1}{m} \sum_{k=1}^{m} \chi_A(x_k) = \frac{1}{m} \sum_{k=1}^{m} \chi_A(x_k(s)) \qquad (2.2)$$

be the relative frequency of occurrence of events A in trials m and it will be called frequency, in short, later.

Proposition 2.1 For any fuzzy random experiment, let (Ω, P) be its common sample space. Then in a sequence of F-random experiment of subject s ($\in S$), the frequency $f(A, m, s)$, $m=1,2,\ldots$, has all the properties as follows. For each $A \in P$, and for each $m \ge 1$;

(1) $f(A, m+1, s) = (1/(m+1))[m f(A, m, s) + \chi_A(x_m(s))]$.

(2) $|f(A, m+1, s) - f(A, m, s)| \le 1/(m+1)$.
Note that $f(A, m+1, s)$ and x_{m+1} are unknown in (1) and (2)

(3) $0 \le f(A, m, d) \le 1$.

(4) $f(\Omega, m, s) = 1$.

(5) If $A_i \in P$ and $A_i A_j = \phi$, $i \ne j$; $i,j = 1,2,\ldots$, let $A = \bigcup_{i=1}^{\infty} A_i$,

then $f(A, m, s) = \sum_{i=1}^{\infty} f(A_i, m, s)$.

(6) If $A_1, A_2 \in P$ and $A_1 A_2 = \phi$, then

$$f(A_1 \cup A_2, m, s) = f(A_1, m, s) + f(A_2, m, s).$$

(7) $f(A^c, m, s) = 1 - f(A, m, s)$.

(8) $f(\phi, m, s) = 0$.

(9) If $A_1, A_2 \in P$ then

$$f(A_1 \cup A_2, m, s) = f(A_1, m, s) + f(A_2, m, s) - f(A_1 A_2, m, s).$$

(10) If $A_1, A_2 \in P$ and $A_1 \subseteq A_2$, then $f(A_1, m, s) \le f(A_2, m, s)$.

Let $S = \{s\}$ be an arbitrary set of subjects, for each $s \in S$, let $\Omega(s)$ be a valuation set of subject s, for each $m \ge 1$, in an observation duration, $\Omega_m(s)$, the relative frequency

$$f(A, m, s) = \frac{1}{m} \sum_{k=1}^{m} \chi_A(x_k(s)) \qquad (2.3)$$

is a set function that is determined by subject s with samples $\Omega_m(s)$. so we have

1. For set $\Omega(s)$ of subject $s (\in S)$.
2. Any σ-field ϕ in $\Omega(s)$.
3. Set function $f(A, m, s)$ is a normal measure on ϕ.

So every such triple ($\Omega(s), \phi, f(A, m, s)$) will be call a probability space according to the viewpoint of modern probability theory.

Corollary. Let (Ω, P, $P(\cdot, s)$) be an F-probability space, let ϕ be any σ-field of subsets of Ω. Then (Ω, ϕ, $P(\cdot, s)$) is an F-probability space, too.

Proof. Clearly $\phi \subseteq P$. $P(\cdot, s)$ is an F-probability on P, for each $A \in \phi$, then $A \in P$ and $P(A, s) = \lim_{m \to \infty} f(A, m, s)$.

(1) If ϕ is a σ-field, then $\Omega \in \phi$ and $P(\Omega, s) = 1$.

(2) For each $A \in \phi$, $A \in P$ and $0 \le P(A, s) \le 1$.

(3) Let be a sequence of disjoint

subsets of Ω that belong to ϕ, and

$$A=\bigcup_{i=1}^{\infty}A_i, \text{ then } A\in P \text{ and } A_i\in P,$$

$$i=1,2,\ldots, \quad P(A,s)=\sum_{i=1}^{\infty}P(A_i,s).$$

3. Agreement index

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 $\chi_H(x),\ x\in R$, we consider another membership function $\chi_A(x),\ x\in R$, which we call the agreement index of A with regard to H the ratio being defined in the following way;

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

as shown in Figure 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

$$\chi_{R_\gamma}(0)=\sup_\psi\{area\ (\chi_{H_\gamma}(\psi)\cap\chi_{T_\gamma}(\psi))\,/area\ \chi_{H_\gamma}(\psi)\} \quad (3.2)$$

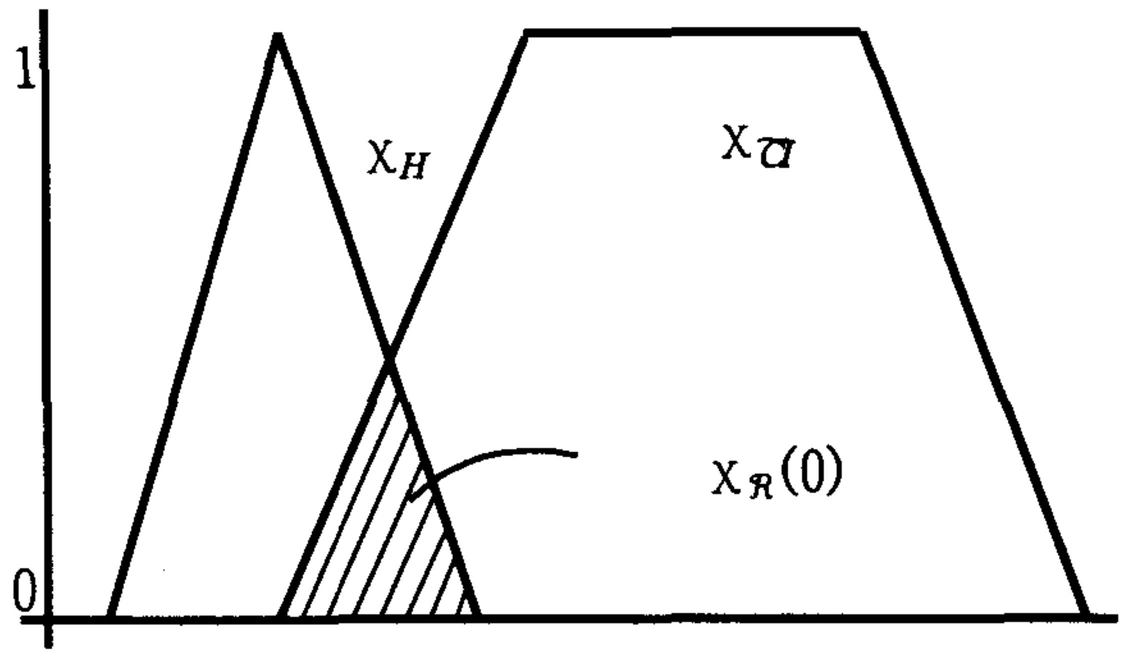
$$\chi_{R_\gamma}(1)=1-\chi_{R_\gamma}(0) \quad (3.3)$$

for the fuzzy hypothesis testing as Figure 1.

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

$$area(\chi_A(x)\cap\chi_H(x))=\int_{\gamma_0}^{\gamma_1}(A_r^{-1}(\gamma)-H_l^{-1}(\gamma))d\gamma \quad (3.4)$$

where A_r, A_l are right and left side line of $\chi_A(x)$, H_l is left side line of $\chi_H(x)$ and γ_0 is reliable degree and γ_1 is meeting point of $\chi_A(x)$ and $\chi_H(x)$.



[Figure 3.1]

4. Fuzzy power function

A test of the fuzzy null hypothesis is a course of action specifying the set if values of a fuzzy random variable X for which H_0 is to be rejected. The random variable whose value serves to determine the action is called the test statistic, and the set of its values for which H_0 is to be rejected is called the fuzzy rejection region of the test. A test of completely specified by a fuzzy test statistic and the fuzzy rejection region.

	UNKNOWN TRUE STATE OF NATURE	
TEST concludes	H_0 true $p<p_0$	H_0 false $(p>p_0)$
Do not reject H_0	Correct degree	Wrong (type Ⅱ error)
Reject H_0	Wrong (type Ⅰ error)	Correct degree

Fuzzy Type Ⅰ error : rejection degree of H_0 when H_0 is true.

Fuzzy Type Ⅱ error : failure degree to reject H_0 when H_1 is true.

The probabilities of the two types of error

$\tilde{\alpha}=P$[type Ⅰ error] = P[rejection of H_0 when H_0 is true]

$\tilde{\beta}=P$[type Ⅱ error] = P[not rejecting H_0 when H_1 is true]

The probability $\tilde{\alpha}$ depends on the particular value of the parameter in the range covered by H_0, whereas β depends on the value over the range covered by H_1.

$\gamma(p) = P$ [the test rejects H_0 when the true value of the parameter is p]

Under H_0, p is restricted to the range $p < p_0$, which is to the left of the middle vertical line in Figure 4.1. In this part of the graph, the rejection probability $\gamma(p)$ is, by definition, the same as the type I error probability $\tilde{\alpha}(p)$. Under H_1, the range of p is $p > p_0$, which is to the right of the middle vertical line. In this range, $1 - \gamma(p) = P$[retain H_0]= P[type II error]= $\beta(p)$. Thus the graph of the rejection probability curve $\gamma(p)$ of a test provides a complete picture of the performance of the test for all possible contingencies with regard to the true state of nature.

Comparing the power curves of the three tests, we observe two important features:

(a) In each case, the largest type I error probability $\tilde{\alpha}(p)$ occurs at $p \approx p_0$, which is the boundary between H_0 and H_1. For the purpose of controlling the type I error probability, it is therefore sufficient to pay attention to its magnitude at this boundary point.

(b) If one of any two tests has a smaller $\tilde{\alpha}(p)$, its $\beta(p)$ is larger than that of the other test. This shows that an error probability of one kind can only be reduced at the cost of increasing the other error probability.

If we have a fuzzy test with rejection region

$$X > \tilde{n}_i \qquad (4.1)$$

then rejection probabilities for fuzzy test is

$$r(p) = P(X > \tilde{n}) \qquad (4.2)$$

such as fuzzy number

$$[\, r^l_{n_i}(p),\, r^c_{n_i}(p),\, r^r_{n_i}(p)] \qquad (4.3)$$

in Figure 4.1.

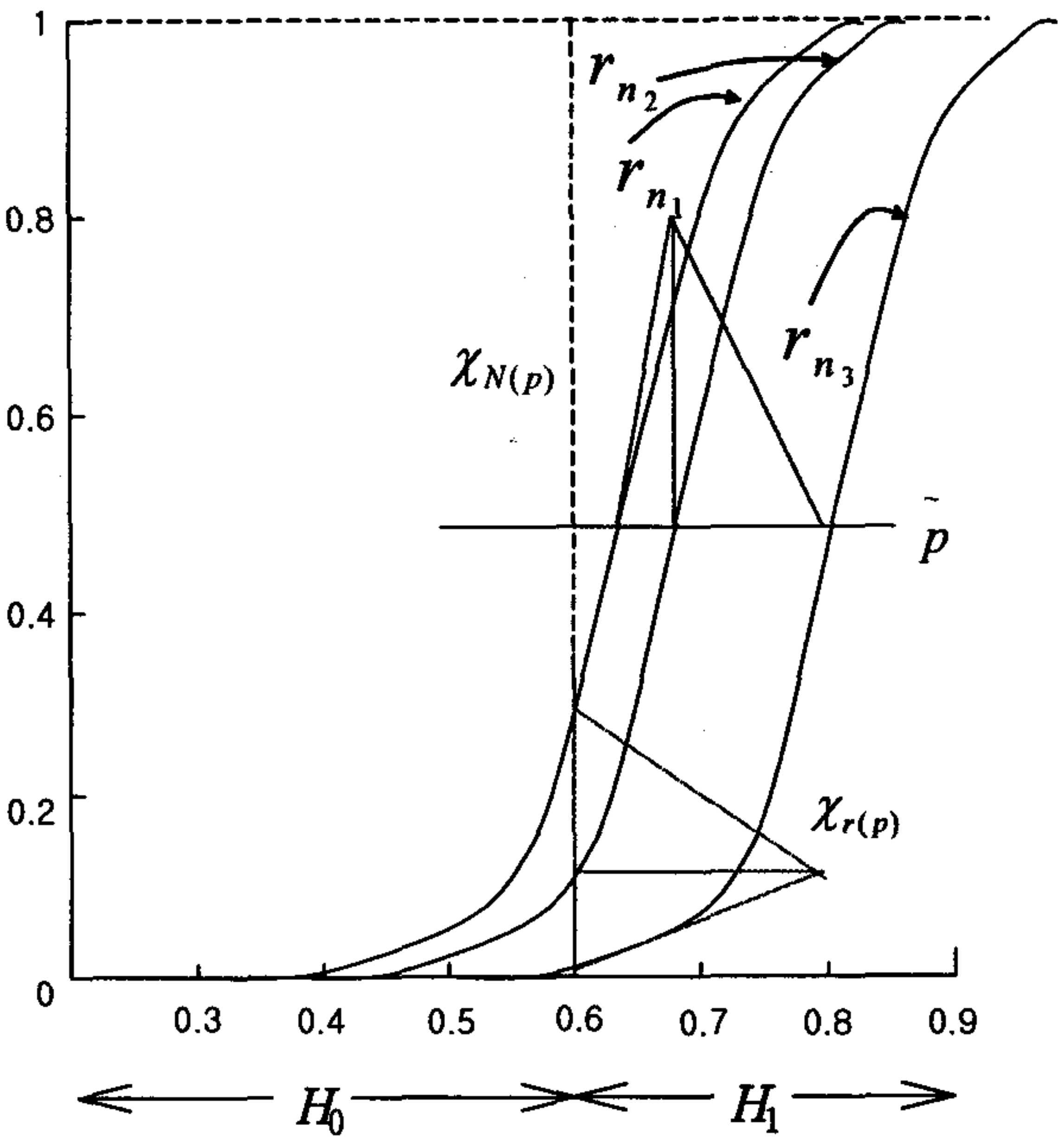


Figure 4.1 Power curves for the tests

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 Test 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.

[7] Z. Xia, Fuzzy Probability System: fuzzy probability system, *Fuzzy Sets and systems 120* , pp.469-486, 2001.