Journal of the KSQC
Vol. 20, No. 1. June. 1992

1992년 6월 韓國品質管理學會誌 제20권 제1호

특정처리가 최우수 처리인가의 판정방법+
A Method for Detecting Whether a Specific Treatment is the Best

정 규 진*

Abstract

Consider a situation that a specific treatment is regarded as the best in comparing several treatments. In this case, a natural procedure to decide the best treatment is to choose the specific one which is proved to be better than each one of the others.

This study proposes a method for detecting the best by comparing two treatments from the above point of view. Since this method deals with comparing only two treatments, it is very simple to use and expected to be widely applicable.

1. 서 론

다수의 처리를 비교하고 최우수 처리를 판정하는 문제는 다중비교법(multiple comparisons procedure) 등에서 오랫동안 다루어진 문제이다. 특히 최우수 처리의 판정은 많은 경우 실험자의 최종 목적이라 할 수 있다.

여러가지 다중비교 기법 중에서 미지의 최우수 처리 판정에 적용될 수 있는 것으로는 Tukey(1953)과 Hsu(1981, 1984)의 방법 등이 있다.

K개의 처리를 비교할 때 각 처리효과를 $\theta_1, \cdots, \theta_k$라 하면 Tukey(1953)의 방법은 모든 $1 \leq i < j \leq k$에 대하여 $\theta_i - \theta_j$의 동시 신뢰구간(simultaneous confidence interval)을 제공한다. 이 방법은 k개의

* 한남대학교 응용통계학과 교수
+ 이 논문은 1991년도 문교부 지원 한국학술진흥재단의 지방대 육성 학술연구 조성비에 의하여 연구되었음.

처리에 대한 전체적인 구조를 밝히는데 그 목적이 있으므로 최우수 처리의 판정만을 고려한다면 그 판정이 어렵다는 측면에서 비효율적이라 할 수 있다.

이점에 착안하여 Hsu(1981, 1984)는 최우수 처리의 판정에 보다 효율적인 소위 최우수 처리와의 다중비교(multiple comparisons with the best)법을 제안하였다. 그의 방법은 $\theta_i - /\max_{i \neq j} \theta_j$, $i = 1, \cdots, k$,에 대한 동시 신뢰구간을 제공한다.

한편 Dunnett(1955)은 특정처리 — 예를 들어 1번 처리 — 가 기준처리(control treatment)로 간주되어질 때 기준처리와 나머지 처리들의 다중비교 방법으로 $\theta_1 - \theta_i$, $i = 2, \cdots, k$,의 동시 신뢰구간을 고려하였다. Tukey(1953)나 Hsu(1981, 1984)의 방법이 미지의 최우수 처리 판정에 이용될 수 있는 반면 이 방법은 특정처리가 최우수 처리인가를 판정하는 문제에 적용될 수 있다.

본 논문에서는 실험 전에 1번 처리가 최우수 처리로 예상되어지는 경우 과연 1번 처리가 최우수 처리인가를 판정하는 문제를 고려한다.

만약 θ값이 큰 처리를 보다 우수한 처리라 하면 1번 처리가 최우수 처리인가의 판정은 $\theta_1 - \max_{i \neq 1} \theta_i$에 대한 신뢰하한(lower confidence bound)을 구하는 문제로 또는 가설

$$H_0: \ \theta_1 \leq \max_{i \neq 1} \theta_i, \ \text{대} \ H_1: \ \theta_1 > \max_{i \neq 1} \theta_i, \quad \cdots (1.1)$$

에 대한 검정 문제로 다루어질 수 있다. 이 방법은 최우수 처리의 판정시 오직 하나의 신뢰하한만을 요하므로 전술한 방법들보다 효율적이라 하겠다.

최근에 Kim, Na and Han(1987)은 정규모집단에서 가설(1.1)에 대한 우도비검정(likelihood ratio test)을 유도하고 이에 기초한 신뢰하한을 제시하였다. 또 그들은 Dunnett(1955)의 방법과의 비교를 통하여 최우수 처리 판정에서의 효율성을 보였다.

본 논문에서는 가설(1.1)에 대한 검정방법으로 Berger(1981)의 검정법을 이용한다. 제2장 1절에서 설명되는 이 검정법은 Kim, Na and Han(1987)이 고려한 우도비검정에 비하면 그 유도가 매우 간단하여 여러 가지 모형에 적용이 쉬운 장점이 있다.

제2장 2절과 3절에서는 위치모수(location parameter) 및 척도모수(scale parameter) 모형에서의 검정법이 제시되고 그 예로써 정규모집단에서의 최대 평균과 최소 분산의 판정방법이 고려된다.

제3장에서는 맺음말로써 기타 모형에 대한 적용방법 등이 설명된다.

2. 최우수 처리의 판정

2.1. 판정방법

θ값이 큰 처리가 보다 우수한 처리일 때, 1번 처리가 최우수 처리인가를 판정하는 방법으로 가설 (1.1)에 대한 검정문제를 생각하자.

가설 (1.1)의 검정은 가설 H_0: 어떤 i에 대하여 $\theta_1 \leq \theta_i$ 대 H_1: 모든 $i = 2, \cdots, k$에 대하여 $\theta_1 \leq \theta_i$의 검정으로 표현될 수 있으므로 $i = 2, \cdots, k$에 대하여

$$H_{0i}: \ \theta_1 \leq \theta_i \ \text{대} \ H_{1i}: \ \theta_1 > \theta_i \quad \cdots\cdots\cdots\cdots\cdots (2.1)$$

의 검정을 고려하면 H_0가 기각되는 경우와 모든 H_{0i}가 기각되는 경우가 일치한다.

따라서 "모든 H_{0i}가 기각되면 H_0를 기각"하는 검정법을 사용하여 H_0 대 H_1의 검정을 보다 간단한 형태의 가설 H_{0i} 대 H_{1i}의 검정으로 대체할 수 있다. 이때 H_{0i} 대 H_{1i}의 검정법이 모두 수준(level) α $(0 \leq a \leq 1)$이면 H_0 대 H_1의 검정도 수준 α임을 쉽게 알 수 있다.

이와 같은 현상은 H_0 모수영역이 H_{0i} 모수영역들의 합(union)으로, H_1 모수영역이 H_{1i} 모수영역들의 공통부분(intersection)으로 표현될 때 일어

난다. 이런 형태의 가설에 대하여 Berger(1981)는 보다 일반적인 검정법을 제시하였고, Gutmann (1987)은 위치모수 모형에서 H_0 : $\min_i \theta_i \leq 0$대 H_1 : $\min_i \theta_i > 0$의 검정문제를 가설 H_{0i}를 고려하지 않고 다루었다.

H_0 대 H_1의 검정을 위하여 H_{0i} 대 H_{1i}의 검정법을 이용하는 아이디어는 $\theta_1 - \max_{i \neq 1} \theta_i$의 신뢰하한을 구하는 문제에도 그대로 적용된다. 즉, $\theta_1 - \theta_i$의 $100(1\text{-}\alpha)\%$ 신뢰하한을 L_i라 하면 $\theta_1 - \max_{i \neq 1} \theta_i$의 $100(1-\alpha)\%$ 신뢰하한은 $\min_{i \neq 1} L_i$가 된다.

물론 이와 같은 신뢰하한은 검정법과 신뢰구간의 관계를 이용하여 H_0 대 H_1의 검정법으로부터 직접 유도할 수도 있으나 이보다는 H_{0i} 대 H_{1i}의 검정법으로부터 $\theta_1 - \theta_i$의 신뢰하한을 계산하는 것이 상대적으로 쉽기 때문에 위에서 언급한 방법이 더 간편하다 하겠다.

2.2. 위치모수 모형

서로 독립인 연속확률변수 $Y_1, \cdots, Y_k$를 각각 위치모수(= 처리효과) $\theta_1, \cdots, \theta_k$의 추정량이라 하고 가설(2.1)에 대하여

"$Y_1 - Y_i > d_i$, $d_i > 0$, 이면 H_{0i}를 기각한다." ……………………(2. 2)

는 수준 α의 검정법을 고려하자.

그러면 H_0 대 H_1에 대한 수준 α 검정법은 다음과 같이 표현된다 ;

"$Y_1 > \max_{i \neq 1} (Y_i + d_i)$이면 H_0를 기각한다." ……………………(2. 3)

만약 Y_i의 연속인 분포함수(c.d.f.)를 $F(y-\theta_i)$, $i=1, \cdots, k$,라 하면 상수 d_i는 $(Y_1-\theta_1) - (Y_i-\theta_i)$의 분포의 상위 $100\alpha\%$ 백분위수이고 이는 i와는 무관한 값(=d)임을 알 수 있다.

이 경우 처리 1과 $Y_j = \max\{Y_2, \cdots, Y_k\}$에 대응되는 처리 j를 비교하여 $Y_1 - Y_j > d$이면 H_0를 기각한다. 따라서 1번 처리가 최우수 처리인가의 판정은 단지 θ_1이 θ_j보다 큰가를 단측검정함으로써 가능해진다.

또, 분포함수 F가 연속이므로 검정법 (2.2)와 (2.3)은 실제로는 크기(size) α인 검정이고, Berger (1981)의 3.1절 또는 Gutmann(1987)의 정리 2.1에 의하면 검정법 (2.3)은 균일 최강력 단조검정(uniformly most powerful monotone test)임을 보일 수 있다.

한편 검정법 (2.2)로부터 얻어지는 $\theta_1 - \theta_i$의 $100(1-\alpha)\%$ 신뢰하한이 $Y_1 - Y_i - d_i$이므로 관심모수 $\theta_1 - \max_{i \neq 1} \theta_i$에 대한 $100(1-\alpha)\%$ 신뢰하한은 다음과 같이 주어진다 ;

$$\theta_1 - \max_{i \neq 1} \theta_i \geq Y_1 - \max_{i \neq 1} (Y_i + d_i) \quad \cdots\cdots\cdots(2.4)$$

이제 서로 독립인 정규모집단 $N(\theta_i, \sigma^2)$에서 크기 n_i인 표본을 뽑았을 때 표본평균을 Y_i, n_i의 합을 N, Y_i와 독립이고 자유도 $\nu = N-k$인 χ^2분포를 따르는 σ^2의 합동추정량을 S^2, 그리고 자유도 ν인 t분포의 상위 $100\alpha\%$ 백분위수를 $t\alpha(\nu)$라 하자. 단, $i=1, \cdots, k$.

그러면 검정법 (2.2)에서 $d_i = t_\alpha(\nu) S\sqrt{1/n_1 + 1/n_i}$이므로 검정법 (2.3), 또는 신뢰구간 (2.4)로부터

$$Y_1 > \max_{i \neq 1} \{Y_i + t\alpha(\nu) S\sqrt{1/n_1 + 1/n_i}\} \quad \cdots\cdots\cdots(2.5)$$

이면 유의수준 α에서 θ_1을 가장 크다고 판정한다. 이때 표본크기가 n으로 모두 같으면 $d = t\alpha(\nu) S\sqrt{2/n}$이 되어 $Y_1 - \max_{i \neq 1} Y_i > d$이면 θ_1을 가장 크다고 판정한다.

Kim, Na and Han(1987)은 우도비검정을 통하여 이와 동일한 결과를 얻었으며 상수 d가 1장에서 언급한 Dunnett(1955)등의 방법에서 사용되는 상수보다 작아서 최우수 처리의 판정에 보다 효율

적임을 보였다.

2.3. 척도모수 모형

이 절에서는 k개의 처리효과가 척도모수 $\theta_1, \cdots, \theta_k$로 표현되고 θ값이 가장 작은 처리가 최우수 처리이며 1번 처리가 최우수 처리일 것으로 예상되는 경우를 고려한다. 가장 큰 척도모수의 판정에 관심이 있을 때에도 비슷한 방법으로 다루어진다.

이 경우 관심있는 가설은

$$H_0 : \theta_1 \geq \min_{i \neq 1} \theta_i \text{ 대 } H_1 : \theta_1 < \min_{i \neq 1} \theta_i$$

으로 표현되고 이 가설의 검정법은

$$H_{0i} : \theta_1 \geq \theta_i \text{ 대 } H_{1i} : \theta_1 < \theta_i$$

의 검정법으로부터 얻어진다. $i=2, \cdots, k$.

서로 독립이고 연속인 확률연수 Y_i가 척도모수 θ_i, $i=1, \cdots, k$,의 추정량이라 하자. 또 상수 c_i, $0 < c_i < 1$,에 대하여 H_0 대 H_1의 수준 α 검정법으로

"$Y_1 < C_iY_i$ 이면 H_{0i}를 기각한다."

는 검정법을 고려하자. 단, $i=2, \cdots, k$.

그러면 H_0 대 H_1에 대한 수준 α 검정법은

"$Y_1 < \min_{i \neq 1} (C_iY_i)$이면 H_0를 기각한다."

로 표현된다.

이때 Y_i의 분포함수를 $F(y/\theta_i)$, $i=1, \cdots, k$,라 하면 상수 C_i는 $(Y_1/\theta_1)/(Y_i/\theta_i)$의 분포의 상위 100α% 백분위수로 i와는 무관한 값$(=c)$이 되어 $Y_1/\min_{i \neq 1} Y_i < c$이면 1번 처리를 최우수 처리로 판정한다.

또, θ_1/θ_i의 $100(1-\alpha)$% 신뢰상한이 $Y_1/(C_iY_i)$이므로 $\theta_1/\min_{i \neq 1}\theta_i$의 $100(1-\alpha)$% 신뢰상한은

$$\theta_1/\min_{i \neq 1}\theta_i \leq Y_1/\min_{i \neq 1}(C_iY_i)$$

으로 주어진다.

서로 독립인 정규모집단 $N(\mu_i, \sigma_i^2)$에서 크기 n_i 표본을 뽑았을때 표본분산을 S_i^2, $i=1, \cdots, k$,라고 하고 $F_\alpha(m, n)$이 자유도 m과 n인 F분포의 하위 100α% 백분위수를 나타낸다고 하자. 그리면 σ_1 또는 σ_1^2이 최소인가의 판정방법은 σ_i와 Y_i를 각각 σ_i^2과 S_i^2, $i=1, \cdots, k$,으로 대체함으로써 얻어지며 이때 $H_{0i} : \sigma_1^2 \geq \sigma_i^2$ 대 $H_{1i} : \sigma_1^2 < \sigma_i^2$의 검정에서 $C_i = F_\alpha(n_1-1, n_i-1)$이 된다. 따라서

$$S_1^2 < \min_{i \neq 1} \{F_\alpha(n_1-1, n_i-1)\, S_i^2\}$$

이면 모집단 1의 분산 σ_1^2이 가장 작다고 판정하고 만약 표본크기가 n으로 모두 같으면

$$S_1^2 < F_\alpha(n-1, n-1) \min_{i \neq 1} S_i^2$$

이거나 또는

$$S_1^2 / \min_{i \neq 1} S_i^2 < F_\alpha(n-1, n-1) \quad \cdots\cdots\cdots\cdots (2.6)$$

일때 σ_1^2을 최소 분산으로 판정한다.

정규분포의 분산에 대한 다중비교 기법은 평균에 대한 기법과 매우 유사하게 얻어지며 이들 중 1장에서 언급한 Tukey나 Hsu 그리고 Dunnett형태의 기법(Hochberg and Tamhane(1987), 10장 2.1절 참조) 등이 최우수 처리의 판정에 이용될 수 있다.

최우수 처리의 판정에 촛점을 맞춘다면 Dunnett형태의 기법이 가장 효율적이고 이에 의하면 표본크기가 모두 n일 경우 σ_1^2/σ_i^2, $i=2, \cdots, k$,의 $100(1-\alpha)$% 동시 신뢰상한이

$$\sigma_1^2 / \sigma_i^2 \leq S_1^2 / (G\ S_i^2)$$

로 주어진다. 여기서 상수 G는 $\min_{i \neq 1} (S_i^2/\sigma_i^2)/(S_1^2/\sigma_1^2)$의 분포의 하위 100α% 백분위수이고 그 값은 Gupta and Sobel(1962)에서 찾아볼 수 있다.

이 동시 신뢰상한을 이용하면 최소 분산의 판정에 사용할 수 있는 $100(1-\alpha)$% 신뢰상한은 다음과 같이 얻어진다 :

$$\sigma_1^2 / \min_{i \neq 1} \sigma_i^2 \quad S_1^2 / (G \min_{i \neq 1} S_i^2)$$

여기서 상한이 1보다 작으면, 즉

$$S_1^2 / \min_{i \neq 1} S_i^2 < G \quad \cdots\cdots (2.7)$$

이면 σ_1^2을 최소 분산으로 판정한다.

이제 (2.6)과 (2.7)에서 σ_1^2을 최소분산으로 판정하는데 사용되는 임계치 $F_\alpha(n-1,\ n-1)$와 G를 비교하여 보자. 두 값이 같아지는 K=2일 때를 제외하면 $\alpha=0.05$와 0.01, 그리고 모든 $k>2$와 n에 대하여 $F_\alpha(n-1,\ n-1) > G$가 된다. 따라서 우리의 방법이 Dunnett의 방법보다 최소 분산의 판정을 쉽게함을 알 수 있다.

3. 맺는 말

제 2장에서 우리는 위치 및 척도모수 모형하에서 1번 처리가 최우수 처리 인가를 판정하는 문제를 다루었다. 그러나 이 기법은 특별한 제약 조건 없이 단지 두 처리의 비교만을 필요로 하므로 다른 모형에도 쉽게 적용될 수 있다.

예를 들어서 k개의 이항비율(binomial proportion) $P_1, \cdots, P_k$, 중 P_1이 가장 큰가를 유의수준 α에서 판정하고 싶으면 이항분포의 정규근사(normal approximation)를 이용한 정규검정(normal test), 또는 Fisher(1935)의 정밀검정(exact test) 등을 이용하여 P_1이 P_i, $i=2, \cdots, k$, 보다 큰지를 유의수준 α에서 검정한다. 이때 P_1을 모든 P_i보다 크다고 할 수 있으면 P_1을 최대 비율로 판정한다.

마찬가지로 k개의 위치모수 $\theta_1, \cdots, \theta_k$중 θ_1이 최대인가를 판정하는 비모수적 방법은 θ_1이 θ_i보다 큰가에 대한 Wilcoxon(1945)의 순위합검정(rank sum test)을 이용하여 또는 $\theta_1-\theta_i$에 대한 Wilcoxon 신뢰하한 등을 통하여 얻어질 수 있다. 이때의 결과는 Kim, Na and Han(1987)에서 찾아볼 수 있다.

이상의 논의에서 보듯이 2.1절에서 제안된 방법은 사용이 간편하고 최우수 처리의 판정 면에서 효율적이며 특히 θ_1과 θ_i의 대소를 비교하는 기법(검정법 또는 신뢰구간 등)이 알려져 있는 문제에 널리 적용이 가능하다는 장점이 있다.

참 고 문 헌

1. Berger, R. L. (1981), "Multiparameter Hypothesis Testing and Acceptance Sampling", *Technometrics*, 24, 295-300.
2. Dunnett, C. W. (1955), "A Multiple Comparisons Procedure for Comparing Several Treatments with a Control", *Journal of the American statistical Association*, 50, 1096-1121.
3. Fisher, R. A. (1935), "The Logic of Inductive Inference", *Journal of the Royal statistical society*, series A, 98, 39-54.
4. Gupta, S. S. and Sobel, S. (1962), "On the Smallest of Several Correlated F-Statistics", *Biometrika*, 49, 509-523.
5. Gutmann, S. (1987), "Tests Uniformly More Powerful than Uniformaly Most Powerful Monotone Tests", *Journal of statistical planning and Inference*, 17, 279-292.
6. Hochberg, Y. and Tamhane, A. C. (1987), *Multiple Comparison Procedures*, John Wiley, New York.
7. Hsu, J. C. (1981), "Simultaneous Confidence Intervals for All Distances from the 'Best'", *Annals of Statistics*, 9, 1026-1034.
8. Hsu, J. C. (1984), "Constrained Two-sided Simultaneous Confidence Intervals for Multiple Omparisons with the Best", *Annals of Statistics*, 12, 1136-1144.
9. Kim, W. C., Na, J. H. and Han, K. S. (1987), "Testing Whether a Specific Treatment is Better Than the Others", *Journal of the Korean Society for Quality Control*, 15(2), 38-49.
10. Tukey, J. W. (1953), *The Problem of Multiple Comparisons*, Mimeographed Monogaph.
11. Wilcoxon, F. (1945), "Indivisual Comparisons by Ranking Methods", *Biometrics*, 1, 80-83.