統計的 假說檢定에 있어서의 等側棄却域에 관한 考察

(A Study on the Two Equal Tail Critical Region for the Testing Statistical Hypothesis)

金 光 燮*

Abstract

In most introductory statistics courses and text, the two equal tail test is presented without justification.

In the present paper, the two equal tail critical region will be discussed in the light of unbiasedness with some test examples for the mean and the variance based on the random sample $X_1, X_2, \dots, X_n$ from $N(\mu, \sigma^2)$ using only elementary mathematics.

1. 緖 論

基礎統計學의 과정이나 敎科書에서는 正規分布를 따르는 母集團의 「파라메타」에 관한 推定 또는 假說檢査의 문제가 중요한 비중을 차지하고 있다. 최근, 推定問題 특히, 區間推定의 問題에 있어서 그 最小性이라든가 또는 주어진 信頼區間의 길이 (length)에 대하여 최대의 信頼度를 갖는 구간의 最適位置 등에 관한 활발한 초등적 硏究가 행하여지고 있다.[1), 2), 3), 4)] 그러나 이러한 초등적인 硏究가 統計的 假說에 대한 檢定의 問題에 있어서는 별로 활발치 못한 實情인 것 같다.

兩側檢定의 경우, 方法論 中心의 기초 통계학 과정에서는 관례적으로 等側檢定(two equal tail test)을 적용하고 있지만, 兩側檢定에서는 一樣最有力棄却域(uniformly most powerful critical region)[5), 6), 7)]이 存在하고 있지 않음이 잘 알려져 있음에도 불구하고 아무런 論議없이 棄却域을 설정하고 있다.

따라서 本 硏究에서는 一變數 正規母集團의 「파라메타」에 대한 兩側檢定의 간단한 事例로써 棄却域의 不偏性이라는 側面에서 等側棄却域을 초등적인 方法으로 評價해 보기로 한다.

2. 母平均 및 母分散의 檢定

* 亞洲大學校工科大學産業工學科 敎授

1) A. H. Bowker & G. J. Lieberman, Engineering Statistics (2nd ed.), Prentice-Hall, Inc., 1972, pp. 172 ~ 178.

2) J. D. Edward, "An Optimal Interval for Introductory Statistics," The American Statistican, Vol. 24, No. 3, June 1970, pp. 22 ~ 23.

3) E. N. Gottfried, "Distribution-Free Confidence Intervals," The American Statistian, Vol. 26, No. 1, Feb. 1972, pp. 39 ~ 41.

4) M. M. Desu, "Optimal Confidence Intervals of Fixed Width," The American Statistican, Vol. 25, No. 2, Apr. 1971, pp. 27 ~ 29.

5) H. D. Brunk, An Introduction to Mathematical Statistics (2nd ed.), Blaisdell Publishing Co., 1965, pp. 264 ~ 267.

6) R. V. Hogg & A. T. Craig, Introduction to Mathematical Statistics (4th ed.), Mcmillan Publishing Co., 1978.

7) S. S. Wilks, Mathematical Statistics, John Wiley & Son, Inc., 1962, pp. 394 ~ 398.

2・1 母分散이 既知일 때의 平均檢定의 例

$X_1, X_2, \cdots\cdots, X_n$이 一變數 正規母集團으로부터의 「랜덤 샘플」이라고 하자. 지금, 有意水準 α에서 母平均 μ에 대한 單純歸無假說 $H_0 : \mu = \mu_0$, 複合對立假說 $H_1 : \mu \neq \mu_0$에 관하여 兩側檢定을 한다면 檢定統計量 U_0는,

$$U_0 = \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma} \cdots\cdots (1)$$

단, $\bar{X} = \frac{1}{n}\sum_{i=0}^{n} X_i$

이며, 等側棄却域은,

$(-\infty,\ Z_{\alpha/2}$ and $Z_{1-\alpha/2},\ \infty) \cdots\cdots$ (2)

단, Z는 標準化變數이며,

$P_r(Z \leqq Z_\alpha) = \alpha$

가 된다. 기각역(2)의 不偏性을 알아 보기 위하여 우선 檢定能力函數 $\beta(\mu)$를 구하면,

$$\begin{aligned}\beta(\mu) &= 1 - P_r\left[Z_{\alpha/2} \leqq \frac{\sqrt{n}(\bar{X}-\mu_0)}{\sigma} \leqq Z_{1-\alpha/2} \mid \mu\right] \\ &= 1 - P_r\left[Z_{\alpha/2} - \frac{\sqrt{n}(\mu_0-\mu)}{\sigma} \leqq \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \leqq Z_{1-\alpha/2} + \frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right] \\ &= 1 - \Phi\left[Z_{\alpha/2} + \frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right] - \Phi\left[Z_{1-\alpha/2} + \frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right] \cdots\cdots (3)\end{aligned}$$

이다. (3)式에서, $\beta(\mu_0) = \alpha$가 됨이 틀림없으며, 또,

$$\frac{d\beta}{d\mu} = -\frac{1}{\sigma}\left[\phi\left(Z_{\alpha/2} + \frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right) - \phi\left(Z_{1-\alpha/2} + \frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right)\right]$$
$$= 0$$

의 解는 $\mu = \mu_0$임을 알 수 있다. 이는 곧 $\beta(\mu_0) \leqq \beta(\mu)$인 것을 의미하고, 따라서 위의 假說檢定에서의 棄却域(2)는 不偏棄却域이 되기 위한 充分條件임을 알 수 있다. 逆으로 有意水準 α에 대한 임의의 兩側棄却域을, $(-\infty,\ Z_a$ and $Z_b,\ \infty)$라고 하면,

$\Phi(Z_a) - \Phi(Z_a) = 1 - \alpha \cdots\cdots (4)$

단, $Z_a < Z_b$ 임

이고,

$$\left.\frac{d\beta}{d\mu}\right|_{\mu=\mu_0} = \Phi(Z_a) - \Phi(Z_b) = 0 \cdots\cdots (5)$$

의 解는,

$Z_a = Z_b$ (or $Z_b = -Z_a$)

이다. 따라서 (4)式으로부터,

$Z_a = Z_{\alpha/2},\quad Z_b = Z_{1-\alpha/2}$

임을 알 수 있다.

이상에서, 一變數 正規母集團의 母平均檢定에 대한 兩側棄却域이 不偏棄却域이 되기 위한 必要하고도 充分한 條件은 等側棄却域임을 알 수 있다.

또, 有意水準 대신에 棄却域의 길이가 미리 前提된 條件下에서 檢定을 해야만 하는 현실인 경우를 생각할 수 있는데, 이때에도 等側棄却域이 最適棄却域이 됨을 쉽게 알 수 있다. 여기서 最適棄却域이라 함은 제 I 종의 過誤를 범할 確率을 최소로 하는 棄却域을 말한다.

이제 採擇域의 길이를 L이라고 하면 採擇域은 $[Z_a,\ Z_a + L]$이 되고 제 I 종의 過誤를 범할 確率 $W(Z_a)$는,

$W(Z_a) = \Phi(Z_a) - \Phi(Z_a + L) \cdots\cdots (6)$

이며, 式 (6)을 최소로 하는 Z_a는,

$$Z_a = -\frac{L}{2}$$

이다. 이제 $\frac{\alpha}{2} = \Phi\left(-\frac{L}{2}\right)$이라면, 正規分布曲線의 對稱性에 의하여,

$$1 - \frac{\alpha}{2} = \Phi\left(\frac{L}{2}\right)$$

이며, 式 (6)을 최소로 하는 棄却域은 式 (2)와 같아짐을 알 수 있다.

2・2 母分散 檢定의 例

2・1에서와 같이 X가 $N(\mu, \sigma^2)$을 따를 때, 有意水準 α에서 $H_0 : \sigma^2 = \sigma_0^2$, $H_1 : \sigma^2 \neq \sigma_0^2$에 대하여 검정하는 경우를 생각한다.

이때 檢定統計量 χ_0^2는,

$$\chi_0^2 = \frac{(n-1)V}{\sigma_0^2} \cdots\cdots (7)$$

단, $V = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

이며, 等側棄却域은,

$(0,\ \chi^2_{n-1;\ \alpha/2}$ and $\chi^2_{n-1;\ 1-\alpha/2},\ \infty) \cdots\cdots (8)$

이다. 이때 檢定能力函數 $\beta(\sigma^2)$은,

$$\begin{aligned}\beta(\sigma^2) &= 1 - P_r\left[A \leqq \frac{(n-1)V}{\sigma_0^2} \leqq B \mid \sigma^2\right] \\ &= 1 - P_r\left[A\frac{\sigma_0^2}{\sigma^2} \leqq \frac{(n-1)V}{\sigma^2} \leqq B\frac{\sigma_0^2}{\sigma^2}\right]\end{aligned}$$

$$= 1-\int_{A\frac{\sigma_0^2}{\sigma^2}}^{B\cdot\frac{\sigma_0^2}{\sigma^2}} K(\chi^2)^{\frac{n-1}{2}} \cdot e^{-\frac{x^2}{2}} d\chi^2 \cdots\cdots (9)$$

단, $A=\chi^2_{n-1\,;\,\alpha/2}$

$B=\chi^2_{n-1\,;\,1-\alpha/2}$

$$K=\frac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)}$$

이다.

우선, $\beta(\sigma_0^2)=\alpha$임은 틀림없다. 그러나, $\sigma^2=\sigma_0^2$일 때 $\frac{d\beta(\sigma^2)}{d\sigma^2}$의 값은 零이 되지 않고, $\sigma^2=1\div\frac{n-1}{\sigma_0^2(B-A)}\log\frac{B}{A}$일 때 零이 된다.

따라서 이 경우에 있어 等側棄却域은 不偏棄却域이 아님을 알 수 있다.

또, 2・1에서와 같이 有意水準 대신에 採擇域의 폭이 L로 전제되는 조건하에서는 棄却域이,

$$(0,\ a \ \text{and}\ a+L,\ \infty) \cdots\cdots (10)$$

로 주어질 것이며, 제 I 종의 과오를 범할 確率 $W(a)$는,

$$W(a)=1-\int_a^{a+b} K(\chi^2)^{\frac{n-1}{2}-1}\ e^{-\frac{x^2}{2}}\,d\chi^2 \cdots\cdots (11)$$

이다. 式 (11)을 최소로 하는 a를 구하면,

$$a=Le^{-\frac{L}{n-1}} \Big/ 1-e^{-\frac{L}{n-1}}$$

이고 最適棄却域은,

$$\left(0,\ \frac{Le^{-L/(n-1)}}{1-e^{-L/(n-1)}}\ \text{and}\ \frac{Le^{-L/(n-1)}}{1-e^{-L/(n-1)}}+L,\ \infty\right) \cdots\cdots (12)$$

이다. 棄却域이 식 (12)로 설정되면, $\beta(\sigma_0^2)\leqq\beta(\sigma^2)$이 성립됨은 자명한 일이다.

3. 結 論

이상에서 一變數 正規母集團의 母平均 및 母分散에 대한 兩側棄却域의 不偏性 여부에 관하여 검토해 보았다.

즉, 母平均의 檢定에 있어서는 等側棄却域이 바로 最適棄却域이 되었으나, 母分散의 檢定에 있어서는 等側棄却域이 곧 不偏棄却域이 아님을 알 수 있었다.

따라서 우리가 一般的으로 等側棄却域을 採擇하는 것은 통계적인 理論에 바탕을 두었다기 보다는 慣例的인 면이 큰 것이라 말할 수 있다.

參 考 文 獻

1) Bowker, A. H. & G. J. Lieberman, Engineering Statistics (2nd ed.), Prentice-Hall, Inc., 1972.
2) Brunk, H. D., An Introduction to Mathematical Statistics (2nd ed.), Blaisdell Publishing Co., 1965.
3) Desu, M. M., "Optimal Confidence Intervals of Fixed Width," The American Statistican, Vol. 25, No. 2, Apr. 1971.
4) Edward, J. D., "An Optimal Interval for Introductory Statistics," The American Statistican, Vol. 24, No. 3, June, 1970.
5) Gottfried, E. N., "Distribution-Free Confidence Intervals," The American Statistican, Vol. 26, No. 1, Feb. 1972.
6) Hogg, R. V. & A. T. Craig, Introduction to Mathematical Statistics (4th ed.), Mcmillan Publishing Co., 1978.
7) Wilks, S. S., Mathematical Statistics, John Wiley & Son, Inc., 1962.