(19) OPTIMUMなCRITICAL REGIONに就て

奈良医大

山本 純恭

n 個の実験データ $x_1, x_2, \cdots\cdots, x_n$ のうち，例えば，x_1 は正規母集団 $N(\mu+\mu', \sigma)$ から抽出された random sample で，$x_2, x_3, \cdots\cdots, x_n$ は同一正規母集団 $N(\mu, \sigma)$ から抽出された size $n-1$ の random sample であるとき，従つて $x_1, x_2, \cdots\cdots, x_n$ の尤度函数が

$$(1) \quad p \equiv p(x_1, x_2, \cdots, x_n; \mu, \mu', \sigma)$$
$$= \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\left\{(x_1-\mu-\mu')^2+\sum_2^n (x_i-\mu)^2\right\}\right]$$

であるとき，parameter μ' に関する統計的仮説

$$(2) \quad H_0: \quad \mu' = 0,$$

即ち，等平均値仮説を検定するための optimum な critical region を求めよう。

この場合には parameter は三つあつて，これらの admissible value は

$$(3) \quad -\infty < \mu < +\infty, \quad -\infty < \mu' < +\infty, \quad 0 < \sigma < +\infty$$

であるから，仮説 H_0 は明らかに複合仮説である。

以下 σ が既知の場合（Case I）と未知の場合（Case II）にわけ，夫々の場合を μ' の admissible value が one sided 即ち $\mu' > 0$（又は $\mu' < 0$）なる場合（[a]）と，一般の場合即

ち $-\infty < \mu' < +\infty$, $\mu' \neq 0$ をとる場合（[b]）とにわけて [a] の場合には uniformly most powerful, [b] の場合には B 型の棄却域を求める：

Case 1　σ が既知の場合.

[a]　対立仮説 $\mu_1 > 0$ に対する最良棄却域 $\sigma = 1$ として一般性を失わない。

(4)
$$p_0 = p(x_1, x_2, \cdots\cdots, x_n\ ; \mu, 0, 1)$$
$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2}\sum_1^n (x_i - \mu)^2\right]$$

とおくと

(5)
$$\phi = \frac{\partial \log p_0}{\partial \mu} = n(\bar{x} - \mu) \qquad \phi' = \frac{\partial^2 \log p_0}{\partial \mu^2} = -n$$

であるから

(6)
$$\phi' = A + B\phi \qquad \text{(但し A, B は x に無関係)}$$

即ち Neyman Peason の條件[1]を満足している。

hyper surface

(7)
$$\phi = Const,$$

はこの場合は超平面

(8)
$$\bar{x} = Const$$

である。　求める Critical region と (8) との交りを $w(\bar{x})$ とすれば，それは

(9)
$$p_1 \geqq k(\bar{x})\, p_0$$

によって定められる。　但し

(10)

$$p_1 = p(x_1, x_2, \cdots, x_n; \mu, \mu', 1),$$

定数 $k(\bar{x})$ は critical region の size を α とすれば

(11) $$\int_{w(\bar{x})} \cdots \int p_0 dx_1 \cdots dx_n = \alpha \int_{\bar{x}=Const.} \cdots \int p_0 dx_1 \cdots dx_n$$

によつて定められる。

(9) より $\mu' > 0$ のとき $w(\bar{x})$ は

(12) $$x_1 \geqq k'(\bar{x}).$$

仮説 H_0 の下で x_1 と $\bar{x}$ の同時分布の密度は

(13) $$Const.\, e^{-\frac{n}{2(n-1)}(x_1-\bar{x})^2 - \frac{n}{2}(\bar{x}-\mu)^2}$$

よつて，$\bar{x}$ given のときの x_1 の Conditional distribution の密度は

(14) $$Const.\, e^{-\frac{n}{2(n-1)}(x_1-\bar{x})^2}$$

よつて $w(\bar{x})$ は

(15) $$x_1 - \bar{x} \geqq \lambda_{2\alpha}\sqrt{\frac{n-1}{n}}$$

が定められる。但し $\lambda_{2\alpha}$ は

(16) $$\frac{1}{\sqrt{2\pi}} \int_{\lambda_{2\alpha}}^{\infty} e^{-\frac{t^2}{2}} dt = \alpha$$

によつて定められる。

よつて求める最良棄却域は

(17) $$x_1 \geqq \bar{x} + \lambda_{2\alpha}\sqrt{\frac{n-1}{n}}\,\sigma.$$

対立仮説として $\mu_1<0$ が admissible なときは明らかに

(18) $$x_1 \leqq \bar{x}-\lambda_{2\alpha}\sqrt{\frac{n-1}{n}}\,\sigma.$$

[b] 対立仮説 $\mu'\neq 0$ に対する B 型棄却域 $\sigma=1$ として一般性を失わない。

(19) $$p=p(x_1,x_2,\cdots\cdots,x_n;\mu,\mu',1)$$
$$=\left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2}\left\{(x_1-\mu-\mu')^2+\sum_2^n(x_i-\mu)^2\right\}\right]$$

とおくと

(20) $$\log p=\text{Const.}-\frac{1}{2}(x_1-\mu-\mu')^2-\frac{1}{2}\sum_2^n(x_i-\mu)^2$$

従って

(21) $$\begin{cases}\phi_1=\dfrac{\partial\log p}{\partial\mu'}=x_1-\mu-\mu', & \phi_{11}=\dfrac{\partial^2\log p}{\partial\mu'^2}=-1\\[2ex] \phi_2=\dfrac{\partial\log p}{\partial\mu}=n(\bar{x}-\mu)-\mu', & \phi_{22}=\dfrac{\partial^2\log p}{\partial\mu^2}=-n\\[2ex] \phi_{12}=\phi_{21}=-1\end{cases}$$

であるから,

(22) $$\begin{cases}\phi_{11}=A_0+A_1\phi_1+A_2\phi_2\\ \phi_{12}=B_0+B_1\phi_1+B_2\phi_2\\ \phi_{22}=C_0\qquad\quad+C_2\phi_2\end{cases}\quad\left(\begin{matrix}\text{但し } A_i,\ B_j,\ C_k \text{ は}\\ x \text{ に無関係}\end{matrix}\right)$$

その他 Neyman (2) の條件をすべて満していることが容易に検証出来る。

従って仮説 H_0 の下に於て

(23) $$k_1(\phi_2)<k_2(\phi_2)$$

があつて

$$(24)\quad \int_{k_1(\phi_2)}^{k_2(\phi_2)} p(\phi_1, \phi_2 \mid H_0)\, d\phi_1 = (1-\alpha)\int_{-\infty}^{\infty} p(\phi_1, \phi_2 \mid H_0)\, d\phi_1$$

$$(25)\quad \int_{k_1(\phi_2)}^{k_2(\phi_2)} \phi_1 p(\phi_1, \phi_2 \mid H_0)\, d\phi_1 = (1-\alpha)\int_{-\infty}^{\infty} \phi_1 p(\phi_1, \phi_2 \mid H_0)\, d\phi_1$$

を満すならば

$$(26)\quad \phi_1 \leqq k_1(\phi_2)\ ,\quad k_2(\phi_2) \leqq \phi_1$$

が求めるB型の棄却域である。

ϕ_2 given 即ち $\overline{x} = Const$ のときの ϕ_1 即ち x_1 の條件附分布の密度は(14)で與えられるから $x_1 - \overline{x} = y$ とすると(24)(25)の條件は

$$(27)\quad \int_{k_1'}^{k_2'} e^{-\frac{n}{2(n-1)}y^2} dy = (1-\alpha)\int_{-\infty}^{\infty} e^{-\frac{n}{2(n-1)}y^2} dy$$

$$(28)\quad \int_{k_1'}^{k_2'} y e^{-\frac{n}{2(n-1)}y^2} dy = 0$$

となる。 (28)より $k_1' = -k_2'$, 又(27)より

$$(29)\quad -k_1' = k_2' = \sqrt{\frac{n-1}{n}}\,\lambda_\alpha$$

が得られる。

但し λ_α は

$$(30)\quad \frac{1}{\sqrt{2\pi}}\int_{-\lambda_\alpha}^{\lambda_\alpha} e^{-\frac{t^2}{2}} dt = 1-\alpha$$

で定められる。

よつて求めるB型の棄却域は

(31) $$|x_1-\bar{x}| \geq \lambda_\alpha \sqrt{\frac{n-1}{n}}\,\sigma$$

である。

Case II. σ が未知の場合

[a] 対立仮説 $\mu' > 0$ に対する最良棄却域

(32) $$p_0 \equiv p(x_1, x_2, \cdots\cdots, x_n ; \mu, 0, \sigma) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_1^n (x_i-\mu)^2\right]$$

とおくと

(33) $$\log p_0 = const. - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2$$

(34) $$\begin{cases} \phi_1 = \dfrac{\partial \log p_0}{\partial \mu} = \dfrac{n}{\sigma^2}(\bar{x}-\mu), \quad \phi_1' = \dfrac{\partial \phi_1}{\partial \mu} = -\dfrac{n}{\sigma^2} \\ \phi_2 = \dfrac{\partial \log p_0}{\partial \sigma} = -\dfrac{n}{\sigma} + \dfrac{1}{\sigma^3}\sum_1^n (x_i-\mu)^2 \\ \phi_2' = \dfrac{\partial \log \phi_2}{\partial \sigma} = \dfrac{n}{\sigma^2} - \dfrac{3}{\sigma^4}\sum_1^n (x_i-\mu)^2 \end{cases}$$

であるから

(35) $$\begin{cases} \phi_1' = A_1 + B_1\phi_1 \\ \phi_2' = A_2 + B_2\phi_2 \end{cases} \qquad (A_i,\ B_j \text{ は } x \text{ に無関係})$$

Neyman Peason の條件[3]を満足しているから，求める best な棄却域 w と $\phi_1 = const.$，$\phi_2 = const.$ 即ち，$\bar{x} = const.$ $s^2 = const.$ との交りを $w(\bar{x}, s^2)$とすれ

ば，それは

(36) $$p_1 \geqq k(\overline{x}\ s^2)\ p_0$$

によつて定められる。 但し

(37) $$p_1 = p(x_1, \text{------}, x_n ; \mu, \mu', \sigma)$$
$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp.\left[-\frac{1}{2\sigma^2}\left\{(x_1-\mu-\mu_1)^2+\sum_{i=2}^{n}(x_i-\mu)^2\right\}\right]$$

又定数 $k(\overline{x}, s^2)$ は，w の size を α とすれば

(38) $$\int_{w(\overline{x}, s^2)}\cdots\int p_0 \prod_{i=1}^{n} dx_i = \alpha \int_{\overline{x}, s^2 = Const}\cdots\int p_0 \prod_{i=1}^{n} dx_i$$

によつて定められる。

(37) より $\mu' > 0$ のとき $w(\overline{x}, s^2)$ は

(39) $$x_1 \geqq k'(\overline{x}, s^2)$$

で定まる。

$\overline{x}$，s^2 given のとき，x_1 の條件附分布を見出さねばならないが（39）は又

(40) $$\tau \equiv \frac{x_1-\overline{x}}{s} \geqq k''(\overline{x}, s^2)$$

としてよいから，最良棄却域は

(41) $$\tau \geqq \tau_{2\alpha} \quad 或は \quad x_1 - \overline{x} \geqq \tau_{2\alpha} s.$$

但し $\tau_{2\alpha}$ は

(42) $$\frac{1}{\sqrt{\pi(n-1)}}\ \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)} \int_{\tau_{2\alpha}}^{\sqrt{n-1}}\left(1-\frac{x^2}{n-1}\right)^{\frac{n-4}{2}} dx = \alpha$$

によつて定められる。 又自由度 $n-2$ の t－分布の 2α %

point を $t_{2\alpha}$ とすれば

$$(43) \qquad \tau_{2\alpha} = \frac{t_{2\alpha}\sqrt{n-1}}{\sqrt{n-2+t_{2\alpha}^{2}}}$$

として求められる。 よって(41)は又

$$(44) \qquad x_1 \geqq \overline{x} + \tau_{2\alpha} s \quad \text{又は} \quad x_1 \geqq \overline{x} + \frac{t_{2\alpha}\sqrt{n-1}}{\sqrt{n-2+t_{2\alpha}^{2}}} s$$

対立仮説として *admissible* な μ' が $\mu' < 0$ なるときは全く同様に

$$(45) \qquad x_1 \leqq \overline{x} - \tau_{2\alpha} s \quad \text{又は} \quad x_1 \leqq \overline{x} - \frac{t_{2\alpha}\sqrt{n-1}}{\sqrt{n-2+t_{2\alpha}^{2}}} s$$

である.

[b] 対立仮説 $\mu' \neq 0$ に対するB型棄却域

$$(46) \qquad \log p = Const. - n\log\sigma - \frac{1}{2\sigma}\left\{(x_1-\mu-\mu')^2 + \sum_{2}^{n}(x_i-\mu)^2\right\}$$

であるから

$$(47) \quad \begin{cases} \phi_1 = \dfrac{\partial \log p}{\partial \mu'} = \dfrac{1}{\sigma^2}(x_1-\mu-\mu') \\ \phi_2 = \dfrac{\partial \log p}{\partial \mu} = \dfrac{n}{\sigma^2}\left(\overline{x}-\mu-\dfrac{\mu'}{n}\right) \\ \phi_3 = \dfrac{\partial \log p}{\partial \sigma} = -\dfrac{n}{\sigma} + \dfrac{1}{\sigma^3}\left\{(x_1-\mu-\mu')^2 + \sum_{2}^{n}(x_i-\mu)^2\right\} \\ \phi_{12} = \phi_{21} = \dfrac{\partial^2 \log p}{\partial \mu \partial \mu'} = -\dfrac{1}{\sigma^2} \\ \phi_{13} = \phi_{31} = \dfrac{\partial^2 \log p}{\partial \mu' \partial \sigma} = -\dfrac{2}{\sigma^3}(x_1-\mu-\mu') \\ \cancel{\phi_{23} =}\ \phi_{32} = \dfrac{\partial^2 \log p}{\partial \mu \partial \sigma} = -\dfrac{2n}{\sigma^3}\left(\overline{x}-\mu-\dfrac{\mu'}{n}\right) \end{cases}$$

$$\begin{cases} \phi_{11} = \dfrac{\partial^2 \log p}{\partial \mu'^2} = -\dfrac{1}{\sigma^2} \\ \phi_{22} = \dfrac{\partial^2 \log p}{\partial \mu^2} = -\dfrac{n}{\sigma^2} \\ \phi_{33} = \dfrac{\partial^2 \log p}{\partial \sigma^2} = \dfrac{n}{\sigma^2} - \dfrac{3}{\sigma^4}\left\{ (x_1-\mu-\mu')^2 + \sum_2^n (x_i-\mu)^2 \right\} \end{cases}$$

よつて

$$(48)\begin{cases} \phi_{11} = A_{10} + A_{11}\phi_1 + A_{12}\phi_2 + A_{13}\phi_3 \\ \phi_{12} = A_{20} + A_{21}\phi_1 + A_{22}\phi_2 + A_{23}\phi_3 \\ \phi_{13} = A_{30} + A_{31}\phi_1 + A_{32}\phi_2 + A_{33}\phi_3 \\ \phi_{22} = B_{220} + B_{222}\phi_2 + B_{223}\phi_3 \\ \phi_{23} = \phi_{32} = B_{230} + B_{232}\phi_2 + B_{233}\phi_3 \\ \phi_{33} = B_{330} + B_{332}\phi_2 + B_{333}\phi_3 \end{cases}$$

（但し A, B は何れも x に無関係）

を満足している．又 Neyman の條件を若干拡張したもの[(4)]をすべて満すことが験証出來る。従つて，帰無仮説 H_0 の下で ϕ_1, ϕ_2, ϕ_3 の同時分布の密度函数を $p(\phi_1, \phi_2, \phi_3)$ とすれば

$$(49)\quad k_1(\phi_2, \phi_3) < k_2(\phi_2, \phi_3)$$

があつて

$$(50)\quad \int_{k_1(\phi_2,\phi_3)}^{k_2(\phi_2,\phi_3)} p(\phi_1, \phi_2, \phi_3 | H_0)\, d\phi_1 = (1-\alpha)\int_{-\infty}^{\infty} p(\phi_1, \phi_2, \phi_3 | H_0)\, d\phi_1$$

$$(51)\quad \int_{k_1(\phi_2,\phi_3)}^{k_2(\phi_2\phi_3)} \phi_1 p(\phi_1, \phi_2, \phi_3 | H_0)\, d\phi_1 = (1-\alpha)\int_{-\infty}^{\infty} \phi_1 p(\phi_1, \phi_2, \phi_3 | H_0)\, d\phi_1$$

を満足するならば

(52) $\phi_1 \leqq k_1(\phi_2, \phi_3)$ 及び $k_2(\phi_2\ \phi_3) \leqq \phi_1$

が求めるB型の棄却域である。

又 ϕ_2, ϕ_3 given のときの ϕ_1 の條件附確率分布の密度 $p(\phi_1 | \phi_2 \phi_3 ; H_0)$ とすれば (50) (51) 式は

(53) $$\int_{k_1}^{k_2} p(\phi_1 | \phi_2, \phi_3 ; H_0)\, d\phi_1 = 1-\alpha$$

(54) $$\int_{k_1}^{k_2} \phi_1\, p(\phi_1 | \phi_2, \phi_3 ; H_0)\, d\phi_1 = (1-\alpha)\int_{-\infty}^{\infty} \phi_1 p(\phi_1 | \phi_2 \phi_3 ; H_0)\, d\phi_1$$

となる。この場合には條件

(55) $\phi_2 = Const.$ $\phi_3 = Const.$

は條件

(56) $\bar{x} = Const.$ $s^2 = Const.$

と同値であるから，この條件の下で ϕ_1 の確率分布を求めなければならないが，そのためには x_1，從って又

(57) $$\tau = \frac{x_1 - \bar{x}}{s}$$

の分布を求めればよい。而してその標本分布の密度は

$$\frac{1}{\sqrt{(n-1)\pi}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)} \left(1 - \frac{x^2}{n-1}\right)^{\frac{n-4}{2}} \quad (|x| < \sqrt{n-1})$$

である。條件 (54) から $k_2 = -k_1$ を得て，更に條件 (53) から

(59) $k_2 = -k_1 = \tau_\alpha$

が得られる。但し τ_α は，

$$(60) \qquad \frac{2}{\sqrt{(n-1)\pi}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)} \int_{\tau_\alpha}^{\sqrt{n-1}} \left(1-\frac{x^2}{n-1}\right)^{\frac{n-4}{2}} dx = \alpha$$

によつて求められる。 又自由度 $n-2$ の t－分布の α % point t_α を用いると、

$$(61) \qquad \tau_\alpha = \frac{t_\alpha \sqrt{n-1}}{\sqrt{n-2-t_\alpha^2}}$$

として定められる。
従つて求めるB型の棄却域は

$$|\tau| \geq \tau_\alpha \quad \text{又は} \quad |x_1 - \bar{x}| \geq \tau_\alpha s$$

又は

$$|x_1 - \bar{x}| \geq \frac{t_\alpha \sqrt{n-1}}{\sqrt{n-2+t_\alpha^2}} s$$

である。

こゝに取扱つたのは所謂 "two sample problem" の最も簡単な場合である。 sample size 2以上の random sample については，分散が等しい場合には等平均値仮説の test には t－test が optimum であることが知られている。

又既に指摘(5)したように t－test と Thompson の τ-test とは同値である。 こゝでは t－test が使えない場合即ち唯一つの標本と他の標本とを比較して等平均値仮説を test する場合にも τ－test が optimum であることを示したことになる。

最後に御来阪以来研究に指導と種々の便宜を與えて下さつた，小川先生に感謝の意を表する。

参考文献

[1.] Neyman ; J. & Peason ; E.S. (1933) : On the problem of the most efficient test of statistical hypotheses. Phil. Trans. of the Roy. Soc. London. A. 231. p 289.

[2] Neyman ; J. (1935). Sur la vérification des hypothéses statistiques composées. Bull. Soc. Math. France, 63. pp. 246−266.

[3] Neyman ; J. & Peason ; E.S. loc. cit.

[4] Neyman ; J. & Peason ; E.S. loc. cit. 及 Sato ; R : (1937): Contribution to the theory of testing statistical composite hypothesés. (未刊) 又は"統計的複合仮説の理論"統計数理研究 Vol. I. No. 1. (1941) p. 38.

[5] 小川潤次郎，山本純恭 : Thompson の Rejection Test の Efficiency に就て，"数学"未刊.