㉓ $Variance$の相等しい二次元正規分布の等平均仮説の検定に就て

奈良医大　山本純恭

二つの確率変数 X, Y の同時分布の密度が

$$(1)\quad f(x,y)=\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}e^{-\frac{1}{2\sigma^2(1-\rho^2)}\left\{(x-m_1)^2-2\rho(x-m_1)(y-m_2)+(y-m_2)^2\right\}}.$$

であるとき，大いさ n の任意標本

$$(2)\quad O_n:\ (x_1,y_1),\cdots\cdots\cdots,(x_n,y_n)$$

によつて帰無仮説 $m_1=m_2$ を検定するための Neyman-Peason の意味に於ける最良の Critical region を求める。

$$(3)\quad \xi=\frac{x+y}{2},\quad \eta=\frac{x-y}{2}$$
$$m=\frac{m_1+m_2}{2},\quad \delta=\frac{m_1-m_2}{2}$$

とおけば

$$(4)\quad f(x,y)=\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}e^{-\frac{1}{(1+\rho)\sigma^2}(\xi-m)^2-\frac{1}{(1-\rho)\sigma^2}(\eta-\delta)^2}$$

従つて O_n の Likelihood function は

$$L = \prod_{i=1}^{n} f(x_i, y_i)$$

(5)
$$= \left(\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\right)^n \exp\left[-\frac{1}{(1+\rho)\sigma^2}\sum_1^n(\xi_i - m)^2 - \frac{1}{(1-\rho)\sigma^2}\sum_1^n(\eta_i - \delta)^2\right].$$

この場合 Parameter は4つあつて，それ等の admissible value は

(6) $-\infty < m < +\infty$, $-\infty < \delta < +\infty$, $0 < \sigma < +\infty$
$-1 \leqq \rho \leqq +1$.

帰無仮説H： $m_1 = m_2$ 又 $\delta = 0$ であつて，明らかに複合仮説である。

ρ が未知の場合には Neyman-Peason の條件を満さないので，ここでは ρ を既知として取扱うことにする。

従つて unspecified parameter としては m, σ となる。

(7) $$\log L = Const. - 2n\log\sigma - \frac{1}{\sigma^2(1+\rho)}\sum_1^n(\xi_i - m)^2 - \frac{1}{\sigma^2(1-\rho)}\sum_1^n(\eta_i - \delta)^2$$

(8)
$$\begin{cases} \phi_1 \equiv \dfrac{\partial \log L}{\partial \delta} = \dfrac{2n}{\sigma^2(1-\rho)}(\bar{\eta} - \delta), \quad \text{但し} \quad \bar{\eta} = \dfrac{1}{n}\sum_1^n \eta_i, \\ \phi_2 \equiv \dfrac{\partial \log L}{\partial m} = \dfrac{2n}{\sigma^2(1+\rho)}(\bar{\xi} - m), \quad \text{但し} \quad \bar{\xi} = \dfrac{1}{n}\sum_1^n \xi_i, \\ \phi_3 \equiv \dfrac{\partial \log L}{\partial \sigma} = -\dfrac{2n}{\sigma} - \dfrac{2}{\sigma^3(1+\rho)}\sum_1^n(\xi_i - m)^2 - \dfrac{2}{\sigma^3(1-\rho)}\sum_1^n(\eta_i - \delta)^2, \\ \phi_{11} \equiv \dfrac{\partial^2 \log L}{\partial \delta^2} = -\dfrac{2n}{\sigma^2(1-\rho)}, \\ \phi_{22} \equiv \dfrac{\partial^2 \log L}{\partial m^2} = -\dfrac{2n}{\sigma^2(1+\rho)}, \end{cases}$$

$$
\left\{
\begin{aligned}
\Phi_{33} &\equiv \frac{\partial^2 \log L}{\partial \sigma^2} = \frac{2n}{\sigma^2} + \frac{6}{\sigma^4(1-\rho)}\sum_1^n(\xi_i - m)^2 + \frac{6}{\sigma^4(1+\rho)}\sum_1^n(\eta_i-\delta)^2, \\
\Phi_{12} &\equiv \Phi_{21} = \frac{\partial^2 \log L}{\partial m \partial \delta} = 0 \\
\Phi_{13} &\equiv \Phi_{31} = \frac{\partial^2 \log L}{\partial \sigma \partial \delta} = -\frac{4n}{\sigma^3(1-\rho)}(\bar{\eta} - \delta) \\
\Phi_{23} &\equiv \Phi_{32} = \frac{\partial^2 \log L}{\partial \sigma \partial m} = -\frac{4n}{\sigma^3(1+\rho)}(\bar{\xi}_1 - m)
\end{aligned}
\right.
$$

とすれば，*linear relation*

$$
(9)\quad
\left\{
\begin{aligned}
\Phi_{11} &= A_{10} + A_{11}\Phi_1 + A_{12}\Phi_2 + A_{13}\Phi_3 \\
\Phi_{12} &= A_{20} + A_{21}\Phi_1 + A_{22}\Phi_2 + A_{23}\Phi_3 \\
\Phi_{13} &= A_{30} + A_{31}\Phi_1 + A_{32}\Phi_2 + A_{33}\Phi_3 \\
\Phi_{22} &= B_{220} \qquad\quad + B_{222}\Phi_2 + B_{223}\Phi_3 \\
\Phi_{23} &= B_{230} \qquad\quad + B_{232}\Phi_2 + B_{233}\Phi_3 \\
\Phi_{33} &= B_{330} \qquad\quad + B_{332}\Phi_2 + B_{333}\Phi_3
\end{aligned}
\right.
$$

が成立つ（A，B は ξ η に無関係）。 即ちこの場合は Neyman— Peason(1)(2) の條件を満している．

以下対立仮説が One-sided の場合と Both-sided の場合とにわけて論ずる。

Case 1. One-sided の場合

考える対立仮説 H_t を $\delta > 0$ として一般性を失わない。このとき求める critical region w と $\Phi_2^0 = \Phi_2(\delta=0) = \text{const}$, $\Phi_3^0 = \Phi_3(\delta=0) = \text{const}.$ との交わり $w(\Phi_2^0, \Phi_3^0)$ は，

(10) $L \geqq k(\Phi_2^0, \Phi_3^0) L_0$ 但し $L_0 = L(\delta=0)$

によつて定まる。即ち

(11) $e^{\frac{n\delta}{(1-\rho)\sigma^2}(2\bar{\eta}-\delta)} \geqq k(\Phi_2^0, \Phi_3^0)$, $(\delta>0)$.

仮説 H：$\delta = 0$ の下で

(12)
$$\Phi_0^0 = \frac{2n}{\sigma^2(1-\rho)}\bar{\eta} \quad , \quad \Phi_2^0 = \frac{2n}{\sigma^2(1+\rho)}(\bar{\xi}-m)$$
$$\Phi_3^0 = -\frac{2n}{\sigma} - \frac{2n}{\sigma^3(1+\rho)}(\bar{\xi}-m)^2 - \frac{2n}{\sigma^3(1-\rho^2)}\left\{(1-\rho)S_\xi^2 + (1+\rho)S_\eta^2 + (1+\rho)\bar{\eta}^2\right\}$$

但し $S_\xi^2 = \sum_1^n (\xi_i - \bar{\xi})^2/n$ $\quad S_\eta^2 = \sum_1^n (\eta_i - \bar{\eta})^2/n$

よつて (11) は

(13)
$$u \equiv \bar{\xi} - m = \text{const.}$$
$$v^2 \equiv (1-\rho)S_\xi^2 + (1+\rho)S_\eta^2 + (1+\rho)\bar{\eta}^2 = \text{const.}$$

なる條件の下で

(14) $\bar{\eta} \geqq k_1(u, v^2)$

によつて定められる。 よつて $u, v^2 = \text{const}$ なる條件の下で $\bar{\eta}$ の分布を求めよう。

$\bar{\xi}$, $\bar{\eta}$, S_ξ^2, S_η の同時分布の *probability element* は,

(15)
$$\text{const.}(S_\xi^2)^{\frac{n-3}{2}}(S_\eta^2)^{\frac{n-3}{2}} e^{-\frac{n}{(1+\rho)\sigma^2}(\bar{\xi}-m)^2 - \frac{n}{(1-\rho)\sigma^2}\bar{\eta}^2 - \frac{n}{\sigma^2(1+\rho)}S_\xi^2 - \frac{n}{\sigma^2(1-\rho)}S_\eta^2} d\bar{\xi}\, d\bar{\eta}\, dS_\xi^2 d.$$

$$= const.(S_\xi^2)^{\frac{n-3}{2}}(S_\eta^2)^{\frac{n-3}{2}} e^{-\frac{n}{(1+\rho)\sigma^2}(\bar{\xi}-m)^2-\frac{n}{\sigma^2(1-\rho)}v^2} d\bar{\xi}\, d\bar{\eta}\, dS_\xi^2\, dS_\eta^2$$

u, v^2, $\bar{\eta}$, S_ξ^2 に移れば

$$\frac{\partial(u, v^2, \bar{\eta}, S_\xi^2)}{\partial(\bar{\xi}, S_\eta^2, \bar{\eta}, S_\xi^2)} = 1+\rho$$

であるから

$$(16)\quad f(u, v^2, \bar{\eta}, S_\xi^2) = const.(S_\xi^2)^{\frac{n-3}{2}}\left(\frac{1}{1+\rho}v^2 - \frac{1-\rho}{1+\rho}S_\xi^2 - \bar{\eta}^2\right)^{\frac{n-3}{2}}$$

$$\times \exp\left\{-\frac{n}{\sigma^2(1+\rho)}u^2 - \frac{n}{\sigma^2(1-\rho^2)}v^2\right.$$

S_ξ^2 を integrate out して $\bar{\eta}$, u, v^2 の同時分布の密度を求めると

$$(18)\quad f(u, v^2, \bar{\eta}) = const.(v^2-(1+\rho)\bar{\eta}^2)^{\frac{2n-4}{2}} e^{-\frac{n}{\sigma^2(1+\rho)}u^2-\frac{n}{\sigma^2(1-\rho)}v^2}$$

$u = const.$ $v^2 = const.$ なる條件の下で $\bar{\eta}$ の分布の密度は

$$(19)\quad const.(v^2-(1+\rho)\bar{\eta}^2)^{\frac{2n-4}{2}},\quad (-v \leqq \sqrt{1-\rho^2}\,\bar{\eta} \leqq v).$$

従つて変換

$$(20)\quad \bar{\eta} = \frac{v}{\sqrt{1+\rho}}\frac{Z}{\sqrt{1+Z^2}} \qquad -\infty < Z < +\infty$$

によつてZに移れば，$1:1$ の対応で monotone だから，求める region は $Z \geqq k_2(u, v^2)$ でありZの條件附分布の密度は

$$(21)\quad const.\frac{1}{(1+Z^2)^{\frac{2n-1}{2}}}$$

更に

(22) $$t = \sqrt{2n-2}\, Z$$

とすればtは自由度 $2n-2$ の Student のt-分布に従い，求める region は

(23) $$\frac{1}{\sqrt{2n-2}\,B\left(\frac{1}{2},\frac{2n-1}{2}\right)}\int_{-\infty}^{k_3}\frac{dt}{\left(1+\frac{t^2}{2n-2}\right)^{\frac{2n-1}{2}}}$$

$$= (1-\alpha)\frac{1}{\sqrt{2n-2}\,B\left(\frac{1}{2},\frac{2n-1}{2}\right)}\int_{-\infty}^{+\infty}\frac{dt}{\left(1+\frac{t^2}{2n-2}\right)^{\frac{2n-1}{2}}}$$

によつて決定される $k_3 = t_{2\alpha}$ を用いると

(24) $$t \geqq t_{2\alpha}$$

となる。

tを explicit に求めると

(25) $$t = \sqrt{(1+\rho)(n-1)}\,\frac{\bar{x}-\bar{y}}{\sqrt{m_{20}-2\rho m_{11}+m_{02}}}$$

但し

$$\bar{x} = \frac{1}{n}\sum_{1}^{n} x_i \quad , \quad \bar{y} = \frac{1}{n}\sum_{1}^{n} y_i$$

(26) $$m_{20} = \frac{1}{n}\sum_{1}^{n}(x_i-\bar{x})^2,\ m_{11} = \frac{1}{n}\sum_{1}^{n}(x_i-\bar{x})(y_i-\bar{y}),\ m_{02} = \frac{1}{n}\sum_{1}^{n}(y_i-\bar{y})^2$$

即ち (25) は自由度 $2n-2$ の Student のt-分布に従い，(24) は size 100α % の uniformly most powerful critical region である。

対立仮説が $\delta < 0$ なるときは明らかに (24) は

(24)' $$t \leqq -t_{2\alpha}$$

となる。

特に $\rho = 0$ の場合は (25) は

$$(25)' \qquad t = \sqrt{n-1}\ \frac{\bar{x} - \bar{y}}{\sqrt{m_{20} + m_{02}}} \qquad d.f \quad 2n-2 .$$

となり, R.A. Fisher の t 即ち

$$t = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}\ \frac{\bar{x} - \bar{y}}{\sqrt{n_1 S_1^2 + n_2 S_2^2}} \qquad d.f \quad n_1 + n_2 - 2$$

に於て $n_1 = n_2 = n$ なる場合に相当する。

Case 2. Both-sided の場合

このときは求める critical region は

$$(27) \qquad \bar{\eta} \leqq C_1(u, v^2), \quad C_2(u, v^2) \leqq \bar{\eta}$$

であつて (25) の t に移れば

$$(28) \qquad t \leqq C_1' \ , \quad C_2' \leqq t$$

であり, C_1', C_2' は

$$(29) \qquad (1-\alpha)\int_{-\infty}^{\infty}\left(1+\frac{t^2}{2n-2}\right)^{-\frac{2n-1}{2}} dt = \int_{C_1'}^{C_2'}\left(1+\frac{t^2}{2n-2}\right)^{-\frac{2n-1}{2}} dt$$

$$(30) \qquad (1-\alpha)\int_{-\infty}^{\infty} t\left(1+\frac{t^2}{2n-2}\right)^{-\frac{2n-1}{2}} dt = \int_{C_1'}^{C_2'} t\left(1+\frac{t^2}{2n-2}\right)^{-\frac{2n-1}{2}} dt$$

によつて定められる. (30) より $C_1' = -C_2'$ が得られ (29) より $C_2' = t_\alpha$ ($d.f\ 2n-2$ の t-分布の α % point) が得られる。

即ち Neyman の所謂 B-型 の critical region は自由度 $2n-2$ の t-分布の α% point を用いて

$$(31)\quad |t| = \sqrt{(1+\rho)(n-1)}\,\frac{|\bar{x}-\bar{y}|}{\sqrt{m_{20}-2\rho m_{11}+m_{02}}} \geq t_\alpha$$

によつて えられる。

参 考 文 献.

(1) Neyman, J. & Peason, E. S. (1933): On the problem of the most efficient tests of statistical hypotheses. Phil. Trans. Roy. Soc. Lon. A. 231.

(2) Neyman, J. (1935): Sur La Vérification des Hypothèses Statistiques Composées. Bull. Soc. Math. France. 63.

附 記： 阪大の小川先生の敎室で Neyman-Peason の，仮說檢定論の勉強をしながらこゝで取扱つた問題を演習の積りで書いて見たのが本年三月のことである。

先生に御校閲を願つたところ Hsu, C. T. の "On samples from a normal bivariate population" を見よとの御注意をいたゞいた。 Hsu は本文 (31) の結果を Likelihood ratio から導いている。

ここでは仮說が one-sided な場合には uniformly most powerful にない both-sided の場合には B 型になることを示したことになる。