② 多次元分布の正規回帰論

所員　小川潤次郎

積率行列 $\Phi = (\varphi_{ij})$ なる k 次元正規分布

$$p(\mathfrak{x}) = \frac{\sqrt{\Delta(\Phi)}}{(\sqrt{2\pi})^k} e^{-\frac{1}{2}\mathfrak{x}\Phi\mathfrak{x}^*} \tag{1}$$

但し　$\mathfrak{x} = (x_1, x_2, \text{--------}, x_k)$

に於て $k-r$ 箇の変数 $x_{r+1}, \text{-------}, x_k$ を固定したときの條件 P は確率分布は

$$\phi(x_1, \text{-----}, x_r) = \frac{P(\mathfrak{x})}{\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{r} P(\mathfrak{x})\,dx_1\cdots dx_r} \tag{2}$$

となる。ところが

$$\mathfrak{x}\Phi\mathfrak{x}^* = \sum_{i,j=1}^{k} \varphi_{ij} x_i x_j$$

$$= \sum_{u,v=1}^{r} \varphi_{uv}(x_u - L_u)(x_v - L_v) + \sum_{g,h=r+1}^{k} \psi_{gh} x_g x_h \tag{3}$$

但し　$L_u = \sum_{g=r+1}^{k} \beta_{ug} x_g$,　$u = 1, 2, \text{------} r$

の形に書くことが出来るから, β_{ug} , ψ_{gh} は次の式から定まる。[1]

$$\sum_{u=1}^{r} \varphi_{uv}\beta_{ug} = -\varphi_{vg}, \qquad \begin{matrix} v = 1, 2, \text{-----}, r \\ g = r+1, \text{------}, k \end{matrix} \tag{4}$$

$$\sum_{u,v=1}^{r} \varphi_{uv}\, \beta_{ug}\beta_{vh} + \psi_{gh} = \varphi_{gh},\ g h = r+1, \cdots, k \tag{5}$$

よって

$$\Phi^{(r)} = (\varphi_{uv}) \tag{6}$$

とおくと

$$p(x_1, \cdots, x_r) = \frac{\sqrt{\Delta(\Phi^{(r)})}}{(\sqrt{2\pi})^r} e^{-\frac{1}{2}\sum \varphi_{uv}(x_u - L_u)(x_v - L_v)} \tag{7}$$

となる。

(4)から定められた β_{ug} は $\mathcal{E}(x_u - \sum_g \beta_{ug} x_g)^2 = min$ として決定された回帰係数と一致することが分る。

今大きさ n の任意標本 $\xi_1, \xi_2, \cdots\cdots, \xi_n$ を縦に並べて

$$\begin{pmatrix} x_{11} & x_{12} & \cdots\cdots & x_{1k} \\ x_{21} & x_{22} & \cdots\cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots\cdots & x_{nk} \end{pmatrix} \tag{8}$$

とする。さうすれば

$$\eta_v = x_{vu} - \beta_{ur+1} x_{vr+1} \cdots\cdots - \beta_{uk} x_{vk},$$

$$u = 1, 2, \cdots, r;\ v = 1, 2, \cdots\cdots, n \tag{9}$$

とすれば $\eta_v = (\eta_{v1}, \eta_{v2}, \cdots, \eta_{vr}),\ v = 1, 2, \cdots\cdots, n$ (10)

は r 次元正規分布（7）からの任意標本であるから

$$H = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \tag{11}$$

とすれば $H^*H=(\sum_{v=1}^{n}\eta_{vu}\eta_{vv})$ は Wishart 分布

$$W_{n,r}(H^*H,\ \Phi^{(r)}) \qquad (12)$$

に従ふ。

次に β_{ug} を Sample から推定する際に

$$\sum_{v=1}^{n}(x_{vu}-b_{ur+1}x_{vr+1}------b_{uk}x_{vk})^2 \qquad (13)$$

を最小にするように b_{ug} を定める。

$$y_{vu}=x_{vu}-b_{ur+1}x_{vr+1}------b_{uk}x_{vk},\ u=1,2,\cdots,r;\ v=1,\cdots n \qquad (14)$$

として

$$\boldsymbol{\eta}_{.u}=\begin{Bmatrix}\eta_{1u}\\ \eta_{2u}\\ \vdots\\ \eta_{nu}\end{Bmatrix},\quad \boldsymbol{y}_{.u}=\begin{Bmatrix}y_{1u}\\ y_{2u}\\ \vdots\\ y_{nu}\end{Bmatrix},\quad u=1,2,\cdots,r \qquad (15)$$

とおくと

$$\boldsymbol{\eta}_{.u}=\boldsymbol{y}_{.u}+\sum_{g}(b_{ug}-\beta_{ug})\boldsymbol{x}_g \qquad (16)$$

但し

$$\boldsymbol{x}_g=\begin{Bmatrix}x_{1g}\\ x_{2g}\\ \vdots\\ x_{ng}\end{Bmatrix},\quad g=r+1,-----,n \qquad (17)$$

ここで n 次元ユークリッド空間 R_n である直交座標系に於て (15) (17) を原点から引かれたベクトルと考へると，(13) を min. するためには

$$(\boldsymbol{x}_{r+1}\ \boldsymbol{y}_{.u})=------=(\boldsymbol{x}_k\ \boldsymbol{y}_{.u})=0 \qquad (18)$$

これは $\boldsymbol{x}_g$ の作る $k-r$ 次元部分空間 R_{k-r} に垂直な空間 R_{-k+r} 内に $\boldsymbol{y}_{.u}$

が含まれることを示す。よって座標軸を適当に回転させて、初めの $k-r$ 箇の軸が R_{k-r} 内にある如く出来る。

$$\xi_{\cdot u} = \begin{Bmatrix} \xi_{1u} \\ \xi_{2u} \\ \vdots \\ \xi_{nu} \end{Bmatrix} \tag{19}$$

として

$$\eta_{\cdot u} = C\,\xi_{\cdot u} \tag{20}$$

$$y_{\cdot u} = C \begin{Bmatrix} 0 \\ \vdots \\ 0 \\ \xi_{k-r+1\,u} \\ \vdots \\ \xi_{nu} \end{Bmatrix} \tag{21}$$

となるような直交行列 C がある。そしてこの C は x_q のみで定まるのであるから、すべての u に対して共通にとれる。よって

$$(y_{\cdot 1} \cdots y_{\cdot r}) = C \begin{Bmatrix} 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \\ \xi_{k-r+1\,1} & \xi_{k-r+1\,2} & \cdots & \xi_{k-r+1\,r} \\ \vdots & \vdots & & \vdots \\ \xi_{n1} & \xi_{n2} & & \xi_{nr} \end{Bmatrix} \tag{22}$$

よって今

$$Z_2 = \begin{Bmatrix} \xi_{k-r+1\,1} & \cdots\cdots & \xi_{k-r+1\,r} \\ \vdots & & \\ \xi_{n1} & \cdots\cdots & \xi_{nr} \end{Bmatrix} \tag{23}$$

とすれば

$$Y^*Y = Z_2{}^* Z_2 \tag{24}$$

又明かに

$$H^*H = Z^*Z \tag{25}$$

であるから $(\xi_{v1}, \xi_{v2}, \cdots, \xi_{vr})$ と $(\eta_{v1}, \eta_{v2}, \cdots, \eta_{vr})$ とは同一正規分布に従ふから Y^*Y は *Wishart* 分布

$$W_{n-k+r,\,r}\,(Y^*Y, \Phi^{(r)})$$

に従ふ。

又 (16), (20), (21) から

$$\sum_g (b_{ug} - \beta_{ug})\,\mathfrak{x}_g = C \begin{Bmatrix} \xi_{1u} \\ \vdots \\ \xi_{k-r\,u} \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \tag{25}$$

である

$$\sum_g (b_{ug} - \beta_{ug})(\mathfrak{x}_g\, \mathfrak{x}_h) = \sum_{v=1}^{r} \sum_{\mu=1}^{k-r} C_{v\mu}\, \xi_{\mu u}\, x_{vh} \tag{26}$$

今 $a_{gh} = (\mathfrak{x}_g\, \mathfrak{x}_h)$, として

$$A_{r+1} = (a_{gh}), \quad A_{r+1}^{-1} = (a^{gh}) \tag{27}$$

とおくと, (26) を解いて

$$b_{ug} - \beta_{ug} = \sum_{h=r+1}^{k} \sum_{v}^{r} \sum_{\mu}^{k-r} C_{v\mu}\, \xi_{\mu u}\, x_{vh}\, a^{gh} \tag{28}$$

よって

$$\mathcal{E}(b_{ug} - \beta_{ug}) = 0 \tag{29}$$

又

$$\mathcal{E}\,(b_{ug}-\beta_{ug})^2=\sum_{h,h'}^{k}\sum_{u,u'}^{r}\sum_{\mu,\mu'}^{k-r}C_{\nu\mu}C_{\nu'\mu'}x_{\nu h}x_{\nu' h'}a^{gh}a^{gh'}\mathcal{E}(\xi_{\mu u}\xi_{\mu' u})$$

$$=\sum_{h,h'}\sum_{\nu,\nu'}\sum_{\mu'}C_{\nu\mu}C_{\nu'\mu'}x_{\nu h}x_{\nu' h'}a^{gh}a^{gh}\varphi^{uu}$$

$$\sum_{h,h'}\sum_{\nu,\nu'}\sum_{\mu}C_{\nu\mu}C_{\nu'\mu'}x_{\nu h}x_{\nu' h'}a^{gh}a^{gh'}\varphi^{uu}$$

$$=\sum_{h,h'}a_{hh'}a^{gh}a^{gh'}\varphi^{uu}=a^{gg}\varphi^{uu} \tag{30}$$

よって

$$\mathcal{E}\,(b_{ug}-\beta_{ug})^2=a^{gg}\varphi^{uu}$$

$$=\frac{A_{r.gg}}{A_r}\sigma^2_{u\cdot r+1\cdots\cdots k} \tag{31}$$

ところで $y_{\cdot u}$ の絶体値を求める

$$\sum_{\nu=1}^{n}y_{\nu u}^2=\sum_{\nu}x_{\nu u}^2-b_{ur+1}\sum_{\nu}x_{\nu u}x_{\nu r+1}\cdots\cdots-b_{uk}\sum_{\nu}x_{\nu u}x_{\nu k}$$

$$=a_{uu}-\sum_{g=r+1}^{k}b_{ug}a_{ug} \tag{32}$$

又 b_{ug} の決定方程式は

$$\sum_{g}b_{ug}a_{hg}=a_{hu} \tag{33}$$

であるから

$$b_{ug}=\frac{1}{A_r}\sum_{h}a_{hu}A_{r,hg} \tag{34}$$

これを (32) に代入して

$$\sum_{\nu}y_{\nu u}^2=a_{uu}-\frac{1}{A_r}\sum a_{ug}a_{uh}A_{r,hg}$$

$$= \frac{1}{A_r}\begin{vmatrix} a_{uu} & a_{ur+1} & \cdots & a_{uk} \\ a_{r+1u} & a_{r+1r+1} & \cdots & a_{r+1k} \\ \vdots & \vdots & & \vdots \\ a_{ku} & a_{kr+1} & \cdots & a_{kk} \end{vmatrix} \equiv \varrho_{u\cdot r+1\cdots k} \text{ とおく} \qquad (35)$$

よって

$$\varrho_{r+1,r+2\cdots k} = \frac{A_r}{A_{r\cdot r+1\cdot r+1}} \qquad (36)$$

こゝで

$$u_{u\cdot r+1} = \sqrt{\varrho_{r+1,r+2\cdots k}}\,(b_{ur+1} - \beta_{ur+1}) \qquad (37)$$

とおくと (25) より

$$u_{u\cdot r+1} = k_1\zeta_{1u} + \cdots\cdots + k_{k-r}\,\zeta_{k-ru} \qquad (38)$$

$\mathcal{E}(u_{u\cdot r+1}^2) = \sigma_{u\cdot r+1\cdots k}^2$ なることより

$$\sum k_v^2 = 1 \qquad (39)$$

よって

$$\mathcal{E}(u_{u\cdot r+1}u_{v\cdot r+1}) = \sum_v k_v^2\,\mathcal{E}(\zeta_u\zeta_v) = \mathcal{E}(\zeta_u\zeta_v) \qquad (40)$$

よって $u_{1\cdot r+1}, u_{2\cdot r+1}, \cdots, u_{r\cdot r+1}$ は $\zeta_{v1}, \zeta_{v2}, \cdots, \zeta_{vr}$ と同一の正規分布をする。

以上を綜合すると次のことが云へたことになる。

$x_{r+1}, \cdots\cdots, x_k$ を固定したときの條件附確率で考へると

H^*H は $W_{n-k+r,r}(H^*H, \Phi^{(r)})$

$\mathfrak{u}_{r+1} = (u_{1,r+1}, \cdots\cdots, u_{r,r+1})$ は $\dfrac{\sqrt{\Delta(\Phi^{(r)})}}{(\sqrt{2\pi})^r}\, e^{-\frac{1}{2}\mathfrak{u}_{r+1}\Phi^{(r)}\mathfrak{u}_{r+1}^*}$

に従ふふのであるが，これらは $x_{\nu r+1}, \cdots\cdots, x_{\nu k}$ に functionally independent であるから absolute prob でも成立する。

参　考　文　献

(1) S.S.Wilks, Math, Stat, 1943. p.168