回帰係数の推定について（Ⅱ）

竹 内 清

1

今回は Z. Hellwig[(1)] の 2 点法による回帰係数の推定の問題を中心として筆者独自の展開を試みることにする。

Hellwig によると，2 点法による回帰係数の推定は，つぎのように行なわれる。

確率変数 X および Y に関する 2 次元の母集団を考え，確率変数（X, Y）の 1 組の値（x, y）は，この母集団の各項に対応するものとする。母集団の回帰線は直線と想定される。すなわち，

$$\hat{y}=\alpha_{21}x+\beta_{20} \tag{1}$$

および

$$\hat{x}=\alpha_{12}y+\beta_{10} \tag{2}$$

ただし，α_{21}，α_{12}，β_{20} および β_{10} は母集団の回帰定数。この母集団から大きさ n の無作為標本が抽出される。すなわち，（x_i, y_i）（i=1, 2, …, n）がえられる。ここで標本平均

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

が計算される。ここで，この標本を，第 1 のグループとして，$\bar{x}$ より大きくない横座標 x をもった点のグループを構成し，第 2 のグループとしては，そ

(1) Z. Hellwig, *Linear Regresion and its Application to Economics,* 1963.

れ以外の点のグループとする。第2のグループに n_2 個の点があるとすると，第1のグループの点は $n-n_2=n_1$ 個である。かくして

$$\left.\begin{aligned} X_1&=X \mid X\leq\bar{x}, \\ Y_1&=Y \mid X\leq\bar{x}, \\ X_2&=X \mid X>\bar{x}, \\ Y_2&=Y \mid X>\bar{x} \end{aligned}\right. \tag{3}$$

つぎの統計量が計算される。

$$\bar{x}_1=\frac{1}{n-n_2}\sum x_1, \qquad \bar{x}_2=\frac{1}{n_2}\sum x_2$$

$$\bar{y}_1=\frac{1}{n-n_2}\sum y_1, \qquad \bar{y}_2=\frac{1}{n_2}\sum y_2$$

ついで

$$\begin{vmatrix} \bar{x}_1 & \bar{y}_1 & 1 \\ \bar{x}_2 & \bar{y}_2 & 1 \\ \bar{x} & \bar{y} & 1 \end{vmatrix}=0$$

という定理を媒介として，3つの点 $(\bar{x}_1, \bar{y}_1)$，$(\bar{x}_2, \bar{y}_2)$，$(\bar{x}, \bar{y})$ が一直線上にあることが導かれ，パラメーター α_{21} の推定値として，この直線の傾斜をとるわけである。そこでそれはつぎの3つの公式のうちの任意の1つによって表わされることになる。

$$a_{21}=\frac{\bar{y}_2-\bar{y}}{\bar{x}_2-\bar{x}} \tag{5}$$

$$a_{21}=\frac{\bar{y}-\bar{y}_1}{\bar{x}-\bar{x}_1} \tag{6}$$

$$a_{21}=\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \tag{7}$$

またパラメーター β_{20} の推定値としては，つぎの3つのものの任意の1つによって表わされる。

$$b_{20}=\bar{y}_2-a_{21}\bar{x}_2 \tag{8}$$

$$b_{20}=\bar{y}_1-a_{21}\bar{x}_1 \tag{9}$$

$$b_{20}=\bar{y}-a_{21}\bar{x} \tag{10}$$

パラメーター α_{12} および β_{10} には，同様にして求められる。

ついで彼は，上掲書では証明なしに，以上の 2 点法による推定値の性質を，定理として，つぎのものを与えている。(2)

定理 1 「回帰係数 a_{21} は母集団の回帰係数 α_{21} の一致推定量である。すなわち，

$$P_r(\,|\,a_{21}-\alpha_{21}\,|>\varepsilon\}\underset{n\to\infty}{\longrightarrow}0 \tag{11}$$ 」

定理 2 「回帰係数 α_{21} は母回帰係数 α_{21} の不偏推定量である。すなわち，

$$E(a_{21})=\alpha_{21} \tag{12}$$ 」

ついで 2 点法による回帰係数の推定量の有効性をみるために，つぎの定理を与えている。

定理 5 「

$$e(a_{21p}\mid x_1=u_1,\cdots,\ x_n=u_n)$$
$$=\frac{V(a_{21c})}{V(a_{21p})}\geq\frac{d_u^2}{s_u^2}\cdot\frac{n-1}{n} \tag{13}$$

ただし

a_{21c}＝最小二乗法による回帰係数の推定量

a_{21p}＝ 2 点法による回帰係数の推定量

$V(a_{21c})=V(_{21c}\mid x_1=a_1,\cdots,x_n=u_n)$

$V(a_{21p})=V(a_{21p}\mid x_1=u_1,\cdots,x_n=u_n)$

$s_u=\sqrt{\frac{1}{n}\sum(u-\bar{u})^2},\quad d_u=\frac{1}{n}\sum\mid u-\bar{u}\mid$

$u_1,u_2,\cdots\cdots,u_n$ は一定の定数。」

---

(2) これらの定理の証明は，
Z. Hellwig, "*Wyznaczanie parametrów regresji Iiniowej metoda dwóch punktów,*" *Zastosowania Matematyki,* vol. 3, 1956.
に与えられていると，上掲書には述べてあるが，筆者は未見。筆者独自の証明と展開を以下で述べる。

ついで $n\to\infty$ の場合の漸近性については，⒀式の右辺にスルッキーの定理を応用して，e が確率的に D_x^2/σ_x^2 に収斂することを述べている。[(3)] ただし

$$D_x=\int_{-\infty}^{\infty}|x-\mu_x|f_1(x)dx,\quad \sigma_x^2=\int_{-\infty}^{\infty}(x-\mu_x)^2 f_1(x)dx$$

もし母集団が正規型であれば，e は $2/\pi$ に確率的に収斂し，また確率変数 X が $N(\mu_x,\ \sigma_x^2)$ であれば，μ_x の推定量としてのメディアンの有効性は同じく $2/\pi$ に等しいことを与えている。さらに

定理４「確率変数

$$\frac{a_{21}-\alpha_{21}}{\sqrt{V(a_{21})}} \tag{14}$$

は，$n\to\infty$ のとき $N(0.1)$ に近づく」

以下これらの論点を中心として，筆者独自の展開することにしよう。

2

さて前節の定理２で与えられた，２点法による回帰係数の推定量 a_{21} が α_{21} の不偏推定量であることを，われわれの前回の問題と対応して，$(X,\ Y)$ がつぎのような２変量正規分布で与えられる場合を出発点として考察しよう。[(4)]

$$f(x,\ y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2-\right.\right.$$
$$\left.\left.-\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_x}+\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right] \tag{15}$$

まず a_{21} の不偏性を検討しよう。前回の結果からして

$$E(y\mid x)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x) \tag{16}$$

したがって

(3)　ここで彼は単に「スルッキーの定理を応用して……」と言っているだけであるが，ここで指しているスルッキーの定理とは，前回脚注(4)でふれた，スルッキーの論文の中で展開されている統計的漸近性に関する定理と考えられる。

(4)　ここでは前節(7)式を利用することにする。

$$\alpha_{21}=\rho\frac{\sigma_y}{\sigma_x} \tag{17}$$

これから

$$E(a_{21}\mid x)=\alpha_{21} \tag{18}$$

を証明するには

$$E(a_{21}\mid x)=\rho\frac{\sigma_y}{\sigma_x} \tag{19}$$

を導出すればよいことになる。ところで x を確定変数と考えると

$$E\left(\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1}\mid x\right)=\frac{1}{\bar{x}_2-\bar{x}_1}E(\bar{y}_2-\bar{y}_1\mid x)$$

$$=\frac{1}{\bar{x}_2-\bar{x}_1}\left\{E(\bar{y}_2\mid x)-E(\bar{y}_1\mid x)\right\} \tag{20}$$

ところで

$$E(\bar{y}_2\mid x)=\frac{1}{n_2}\sum E(y_2\mid x) \tag{21}$$

$$E(y_2\mid x)=\int_{-\infty}^{\infty}yf(y\mid x)dy$$

$$=\int_{-\infty}^{\infty}y\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}exp\left[-\frac{1}{2}\left\{\frac{y_x-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)}{\sigma_y\sqrt{1-\rho^2}}\right\}^2\right]dy \tag{22}$$

上式において

$$\frac{y_x-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)}{\sigma_y\sqrt{1-\rho^2}}=t$$

とおくと

$$dy=\sigma_y\sqrt{1-\rho^2}\,t$$

であるから

$$E(y_2 \mid x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\{\sigma_y\sqrt{1-\rho^2}\,t+\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)\}\,e^{-\frac{t^2}{2}}dt$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sigma_y\sqrt{1-\rho^2}\,te^{-\frac{t^2}{2}}dt+\frac{\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt$$

$$=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x) \tag{23}$$

したがって

$$E(\bar{y}_2 \mid x)=\frac{1}{n_2}\sum E(y_2 \mid x)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_2-\mu_x) \tag{24}$$

同様にして

$$E(\bar{y}_1 \mid x)=\frac{1}{n_1}\sum E(y_1 \mid x)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_1-\mu_x) \tag{25}$$

したがって

$$E\left(\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \mid x\right)=\frac{1}{\bar{x}_2-\bar{x}_1}\left[\left\{\mu_y+\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_2-\mu_x)\right\}-\left\{\mu_y+\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_1-\mu_x)\right\}\right]$$

$$=\rho\frac{\sigma_y}{\sigma_x}=\alpha_{21} \tag{26}$$

よって a_{21} が α_{21} の不偏推定量であることが証明された。

以上の結果は，$(x,\ y)$ が2変量正規分布をしていない場合でも，つぎのようにして求められるのであろう。

$$E\left(\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \mid x\right)=\frac{1}{\bar{x}_2-\bar{x}_1}\left\{E(\bar{y}_2 \mid x)-E(\bar{y}_1 \mid x)\right\} \tag{27}$$

ところで

$$E(\bar{y}_2 \mid x)=\frac{1}{n_2}\sum E(y_2 \mid x) \tag{28}$$

しかるに

$$E(y_2 \mid x)=\alpha_{21}x_2+\beta_{20} \tag{29}$$

$$\therefore E(\bar{y}_2 \mid x)=\frac{1}{n_2}\sum(\alpha_{21}x_2+\beta_{20})=\alpha_{21}\bar{x}_2+\beta_{20} \tag{30}$$

同様にして

$$E(\bar{y}_1 \mid x)=\alpha_{21}\bar{x}_1+\beta_{20} \tag{31}$$

$$\therefore E\left(\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \mid x\right)=\frac{1}{\bar{x}_2-\bar{x}_1}\left\{(\alpha_{21}\bar{x}_2+\beta_{20})-(\alpha_{21}\bar{x}_1+\beta_{20})\right\}=\alpha_{21} \tag{32}$$

したがって，$a_{21}=(\bar{y}_2-\bar{y}_1)/(\bar{x}_2-\bar{x}_1)$ は α_{21} の不偏推定量である。

つぎに α_{21} の分散 $\sigma^2(a_{21})$ を求めよう。

$$\begin{aligned}\sigma^2(a_{21})&=\sigma^2\left(\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \mid x\right)\\&=\frac{1}{(\bar{x}_2-\bar{x}_1)^2}\sigma^2(\bar{y}_2-\bar{y}_1 \mid x)\\&=\frac{1}{(\bar{x}_2-\bar{x}_1)^2}\left\{\sigma^2(\bar{y}_2 \mid x)+\sigma^2(\bar{y}_1 \mid x)\right\}\end{aligned} \tag{33}$$

ところで

$$\sigma^2(\bar{y}_2 \mid x)=\frac{1}{n_2}\sigma^2(y_2 \mid x) \tag{34}$$

ところで

$$\begin{aligned}\sigma^2(y_2 \mid x)&=\int_{-\infty}^{\infty}\{y_2-E(y_2 \mid x)\}^2 f(y \mid x)\,dy\\&=\int_{-\infty}^{\infty}y^2f(y \mid x)dy-E(y_2 \mid x)^2\\&=\int_{-\infty}^{\infty}y^2f(y \mid x)dy-\left\{\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)\right\}^2\end{aligned} \tag{35}$$

$$\int_{-\infty}^{\infty}y^2f(y \mid x)dy$$

$$=\int_{-\infty}^{\infty}y^2\frac{1}{\sqrt{2\pi}\,\sigma_y\sqrt{1-\rho^2}}\,exp\left[-\frac{1}{2}\left\{\frac{y_x-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)}{\sigma_y\sqrt{1-\rho^2}}\right\}^2\right]dy \tag{36}$$

$$\frac{y_x-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x_2-\mu_x)}{\sigma_y\sqrt{1-\rho^2}}=t$$

とおくと

$$\int_{-\infty}^{\infty} y^2 f(y \mid x) dy = \int_{-\infty}^{\infty} \sigma_y{}^2(1-\rho^2) t^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$+ \int_{-\infty}^{\infty} \left\{ \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x_2 - \mu_x) \right\}^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$= \sigma_y{}^2(1-\rho^2) + \left\{ \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x_2 - \mu_x) \right\}^2 \tag{37}$$

(35)式と(37)式から

$$\sigma^2(y_2 \mid x) = \sigma_y{}^2(1-\rho^2) \tag{38}$$

したがって

$$\sigma^2(\bar{y}_2 \mid x) = \frac{\sigma_y{}^2(1-\rho^2)}{n_2} \tag{39}$$

同様にして

$$\sigma^2(\bar{y}_1 \mid x) = \frac{\sigma_y{}^2(1-\rho^2)}{n_1} \tag{40}$$

(39)式と(40)式の結果を(33)式へ代入して

$$\sigma^2(a_{21} \mid x) = \frac{\sigma_y{}^2(1-\rho^2)}{(\bar{x}_2 - \bar{x}_1)^2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \tag{41}$$

これが，求める a_{21} の分散となる。

3

つぎに第1節の定理5の(13)式に対応する関係をわれわれの場合について求めておこう。回帰係数 α_{21} の最小二乗推定量は

$$\sigma^2(a_{21c}) = \frac{\sigma_y{}^2(1-\rho^2)}{\Sigma(x-\bar{x})^2} \tag{42}$$

したがって，2点法による回帰係数の推定量 a_{21p} の効率は，次式で与えられるであろう。

$$e(a_{21p} \mid x)=\frac{\sigma^2(a_{21c})}{\sigma^2(a_{21p})}$$

$$=\frac{\sigma_y{}^2(1-\rho^2)}{\Sigma(x-\bar{x})^2}\Bigg/\frac{\sigma_y{}^2(1-\rho^2)}{(\bar{x}_2-\bar{x}_1)^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$=\frac{(\bar{x}_2-\bar{x}_1)^2\, n_1 n_2/n}{\Sigma(x-\bar{x})^2} \tag{43}$$

上式の右辺の分子に関してはつぎのような関係が導かれるであろう。すなわち

$$\left\{\frac{1}{n}\Sigma \mid x-\bar{x} \mid\right\}^2=\left[\frac{1}{n}\{\Sigma(\bar{x}-x_1)+\Sigma(x_2-\bar{x})\}\right]^2$$

$$=\frac{1}{n^2}\left\{n_1(\bar{x}-\bar{x}_1)+n_2(\bar{x}_2-\bar{x})\right\}^2$$

$$=\frac{1}{n^2}\left\{n_1\left(\frac{n_1\bar{x}_1+n_2\bar{x}_2-n\bar{x}_1}{n}\right)+n_2\left(\frac{n\bar{x}_2-n_1\bar{x}_1-n_2\bar{x}_2}{n}\right)\right\}^2$$

$$=\frac{1}{n^2}\left\{\frac{n_1n_2(\bar{x}_2-\bar{x}_1)}{n}+\frac{n_1n_2(\bar{x}_2-\bar{x}_1)}{n}\right\}^2$$

$$=\frac{4\,n_1{}^2\,n_2{}^2}{n^4}(\bar{x}_2-\bar{x}_1)^2 \tag{44}$$

ところで $n>(n-1)$ であり，かつ

$$\frac{n}{2}\,\frac{n}{2}\geq n_1 n_2 \tag{45}$$

$$\therefore \frac{n_1n_2}{n}(\bar{x}_2-\bar{x}_1)^2\geq\frac{4n_1{}^2n_2{}^2}{n^3}(\bar{x}_2-\bar{x}_1)^2>\left\{\frac{1}{n}\Sigma \mid x-\bar{x} \mid\right\}^2(n-1) \tag{46}$$

したがって

$$e(a_{21p} \mid x)=\frac{\sigma^2(a_{21c})}{\sigma^2(a_{21p})}$$

$$=\frac{(\bar{x}_2-\bar{x}_1)^2\, n_1 n_2/n}{\Sigma(x-\bar{x})^2}$$

$$>\frac{\left\{\frac{1}{n}\Sigma \mid x-\bar{x} \mid\right\}^2(n-1)}{\Sigma(x-\bar{x})^2}$$

$$=\frac{d_x{}^2}{s_x{}^2}\cdot\frac{n-1}{n} \tag{47}$$

よって第1節の定理5が証明された。ただし不等号の場合だけが成立する。ところで，すでにふれたスルッキーの定理を上式の右辺に適用すると，e は確率的に D_x^2/σ_x^2 に収斂することがつぎのようにして証明される。ただし

$$D_x=\int_{-\infty}^{\infty}|x-\mu_x|f_1(x)dx$$

$$\sigma_x^2=\int_{-\infty}^{\infty}(x-\mu_x)^2 f_1(x)dx$$

もとの母集団がわれわれの想定しているような2変量正規分布をするときは，つぎの結果がえられる。

$$\begin{aligned}D_x&=\int_{-\infty}^{\infty}|x-\mu_x|f_1(x)dx\\&=\int_{-\infty}^{\mu_x}(\mu_x-x)f_1(x)dx+\int_{\mu_x}^{\infty}(x-\mu_x)f_1(x)dx\end{aligned}\tag{48}$$

$$\begin{aligned}\int_{-\infty}^{\mu_x}(\mu_x-x)f_1(x)dx&=\mu_x\int_{-\infty}^{\mu_x}f_1(x)dx-\int_{-\infty}^{\mu_x}xf_1(x)dx\\&=\frac{\mu_x}{2}-\left(-\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)=\frac{\sigma_x}{\sqrt{2\pi}}\end{aligned}\tag{49}$$

同様にして

$$\begin{aligned}\int_{\mu_x}^{\infty}(x-\mu_x)f_1(x)dx&=\int_{\mu_x}^{\infty}xf_1(x)dx-\int_{\mu_x}^{\infty}\mu_x f_1(x)dx\\&=\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)-\frac{\mu_x}{2}=\frac{\sigma_x}{\sqrt{2\pi}}\end{aligned}\tag{50}$$

$$\therefore D_x=\int_{-\infty}^{\infty}|x-\mu_x|f_1(x)dx=\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\sigma_x}{\sqrt{2\pi}}=\frac{2\sigma_x}{\sqrt{2\pi}}\tag{51}$$

したがって

$$\begin{aligned}e(a_{21p}\mid x)=\frac{\sigma^2(a_{21c})}{\sigma^2(a_{21p})}\underset{n\to\infty}{\longrightarrow}\frac{D_x^2}{\sigma_x^2}&=\left(\frac{2\sigma_x}{\sqrt{2\pi}}\right)^2\Big/\sigma_x^2\\&=\frac{2}{\pi}\end{aligned}\tag{52}$$

よって第１節の定理５の後で述べられた結果が証明された。

つぎに a_{21} が α_{21} の一致推定量であることを証明しよう。

ε>0 が与えられたものとしよう。チェビシェフの不等式から

$$P_r(\,|\,a_{21}-a_{21}\,|<\varepsilon)\geqq 1-\frac{\sigma_y{}^2(1-\rho^2)}{\varepsilon^2(\bar{x}_2-\bar{x}_1)_2}\left(\frac{1}{n_{\mathrm{I}}}+\frac{1}{n_2}\right) \tag{53}$$

上式右辺において

$$\frac{\sigma_y{}^2(1-\rho^2)}{\varepsilon^2(\bar{x}_2-\bar{x}_1)^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\leq\frac{\sigma_y{}^2(1-\rho^2)}{\varepsilon^2\left\{\frac{1}{n}\Sigma\,|\,x-\bar{x}\,|\right\}^2}\frac{n}{n_1n_2} \tag{54}$$

n_1n_2 は，$n>4$ において，$n_i\neq1$ のとき

$$n<n_1n_2 \tag{55}$$

したがって $n>4$ で $n_i\neq1$ のとき

$$\frac{n}{n_1n_2}<1$$

$$\therefore n_1n_2=k(n_1n_i)n \tag{56}$$

$k(n,\ n_i)$ は，n および n_i に依存する定数で，$k(n,\ n_i)>1$。$k(n,\ n_i)$ は，$|\,n_2-n_1\,|$ の一定値に対して，n が大きくなるにつれて増大し確率的に

$$\lim_{n\to\infty} k(n,\ n_i)\to\infty \tag{57}$$

したがって

$$\frac{\sigma_y{}^2(1-\rho^2)}{\varepsilon^2(\bar{x}^2-\bar{x}_1)^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\leq\frac{\sigma_y{}^2(1-\rho^2)}{k(n,\ n_i)\varepsilon^2\left\{\frac{1}{n}\Sigma\,|\,x-\bar{x}\,|\right\}^2} \tag{58}$$

また

$$\lim_{n\to\infty}\left\{\frac{1}{n}\Sigma\,|\,x-\bar{x}\,|\right\}^2\to\frac{2\sigma_x{}^2}{\pi} \tag{59}$$

であるから，$n\to\infty$ とき

$$1-\frac{\sigma_y{}^2(1-\rho^2)}{\varepsilon^2(\bar{x}_2-\bar{x}_1)^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\to1 \tag{60}$$

$$\therefore \lim_{n\to\infty} P_r(\,|\,a_{21}-\alpha_{21}\,|<\varepsilon)\to 1 \tag{61}$$

したがって a_{21} は α_{21} の一致推定量であることが証明された。

4

つぎに $\sigma^2(a_{21}\,|\,x)$ の観測値をそれぞれの期待値でおきかえた結果をみてみよう。

$$(\bar{x}_2-\bar{x}_1)^2\approx\{E(\bar{x}_2)-E(\bar{x}_1)\}^2 \tag{62}$$

ところで

$$E(\bar{x}_2)\approx\int_{\mu_x}^{\infty}xf_1(x)dx$$

$$=\int_{\mu_x}^{\infty}x\frac{1}{\sqrt{2\pi}\sigma_x}exp\left\{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right\}dx \tag{63}$$

$$\frac{x-\mu_x}{\sigma_x}=t$$

とおくと

$$dx=\sigma_x dt$$

であるから

$$\int_{\mu_x}^{\infty}xf_1(x)dx=\int_0^{\infty}(\sigma_x t+\mu_x)\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2} \tag{64}$$

同様にして

$$E(\bar{x})\approx-\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right) \tag{65}$$

$$\therefore(\bar{x}_2-\bar{x}_1)^2\approx\left\{E(\bar{x}^2)-E(\bar{x}_1)\right\}^2\approx 4\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)^2 \tag{66}$$

つぎに

$$E(n_2)=n\int_{\mu_x}^{\infty}f_1(x)dx=\frac{n}{2} \tag{67}$$

同様にして

$$E(n_1)\approx\frac{n}{2} \tag{68}$$

したがって

$$\sigma^2(a_{21})\approx\frac{\sigma_y^2(1-\rho^2)}{4\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)^2}\left(\frac{2}{n}+\frac{2}{n}\right)$$

$$=\frac{\sigma_y^2(1-\rho^2)}{\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)^2 n} \tag{69}$$

ところで，最小二乗法によった回帰係数の推定量 a_{21c} を同様に展開すると

$$\sigma^2(a_{21c})=\frac{\sigma_y^2(1-\rho^2)}{\Sigma(x-\bar{x})^2}\approx\frac{\sigma_y^2(1-\rho^2)}{(n-1)\sigma_x^2} \tag{70}$$

このような観点から，$\sigma^2(a_{21p})$ に対する $\sigma^2(a_{21c})$ の比によって $\sigma^2(a_{21p})$ の有効性をみると，つぎのようになる。

$$a(a_{21p})=\frac{\sigma^2(_{21c})}{\sigma^2(_{21p})}\approx\frac{\sigma_x^2(n-1)}{\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)^2 n} \tag{71}$$

$$\doteqdot\frac{\sigma_x^2}{\left(\frac{\sigma_x}{\sqrt{2\pi}}+\frac{\mu_x}{2}\right)^2} \tag{72}$$

5

つぎに 2 点法による回帰係数の推定量を用いて，母回帰係数 $\beta=\rho\sigma_y/\sigma_x$ に関する検定および信頼限界の設定についての問題を中心に考察することにしよう。2 変量 (X, Y) に関する仮定は第 1 節に与えられた通りとする。

すでに導いたように，2 点法による回帰係数の推定量 b は

$$E(b)=\beta=\rho\frac{\sigma_y}{\sigma_x}$$

$$\sigma^2(b)=\frac{\sigma_y^2(1-\rho^2)}{(\bar{x}_2-\bar{x}_1)^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

$$=\frac{n\sigma_y^2(1-\rho^2)}{n_1n_2(\bar{x}_2-\bar{x}_1)^2}$$

なる正規分布をする。したがって

$$(b-\beta)/\sigma(b)=\sqrt{\frac{n_1n_2}{n}}\frac{(\bar{x}_2-\bar{x}_1)}{\sigma_y\sqrt{1-\rho^2}}(b-\beta) \tag{73}$$

は，規準正規分布 $N(0, 1)$ をすることになる。もし σ_y および ρ が既知の場合は，上式によって母回帰係数 β に関する検定および信頼限界を設定することができる。

ところで，σ_y および ρ は，一般には未知である。もし標本の大きさ n が大きいときは，σ_y を s_y で，また ρ を r でおきかえることにより，

$$\sqrt{\frac{n_1n_2}{n}}\frac{(\bar{x}_2-\bar{x}_1)}{s_y\sqrt{1-r^2}}(b-\beta) \tag{74}$$

を構成すると，これは近似的に規準正規分布 $N(0, 1)$ をすることが導かれる。

しかし，もし標本の大きさ n が小さいときは，上の統計量をそのまま使うことは危険であろう。本節は，このような問題を解明することに主眼がおかれる。

さて簡単化のために，確定変数と考えられる x に関して $\Sigma x=0$ と仮定しても議論の本質には変りがないであろう。これは x の原点を移動することによって可能である。したがって $\bar{x}=0$ となるので，0 を基準にして x の級分けを考えればよいことになる。

$y_1, \cdots, y_n$ は，それぞれ $N(\alpha+\beta x_1, \sigma_y^2(1-\rho^2)), \cdots, N(\alpha+\beta x_n, \sigma_y^2(1-\rho^2))$ にしたがって独立に分布している。ところで

$$z_i=y_i-\alpha-\beta x_i \tag{75}$$

と変換すると，これは $N(0, \sigma_y^2(1-\rho^2))$ となることは容易に分るであろう。

前節までの結果を考慮しながら，つぎのような直交変換を用いることにしよう。[5]

(5) 直交変換に関する議論については，たとえは
H. W. Alexander, *Elements of Mathematical Statistics, 1961*, *section* 39, 41, 51, 52, 53, 54, 55 参照。

いま (x, y) が n 対無作為に抽出され，$\bar{x}$（ここでの場合 $\bar{x}=0$）を基準として x は，すでにみたように X_1 と X_2 の2組に分類されたものとしよう。ところで，X_1 には n_1 個，X_2 には n_2 個が入るものとする。n_i は0から n までの値をとりうるであろうから，あらゆる組合せを考慮すればよいであろう。ここではそれぞれ n_1 と n_2 として議論を展開することにする。

$$u_1=\frac{1}{\sqrt{nn_1n_2}}(n_1z_{21}+\cdots+n_1z_{2n2}-n_2z_{11}-\cdots-n_2z_{1n1})$$

$$=\frac{1}{\sqrt{nn_1n_2}}\left\{n_1\sum_{i=1}^{n_2}(y_{2i}-\alpha-\beta x_{2i})-n_2\sum_{j=1}^{n_1}(y_{1j}-\alpha-\beta x_{1j})\right\}$$

$$=\frac{1}{\sqrt{nn_1n_2}}\left\{n_1n_2(\bar{y}_2-\bar{y}_1)-n_1n_2\beta(\bar{x}_2-\bar{x}_1)\right\}$$

$$=\sqrt{\frac{n_1n_2}{n}}\left\{(\bar{y}_2-\bar{y}_1)\beta(\bar{x}_2-\bar{x}_1)\right\}$$

$$=\sqrt{\frac{n_1n_2}{n}}(\bar{x}_2-\bar{x}_1)(b-\beta) \tag{76}$$

$$u_2=\frac{1}{\sqrt{n}}(z_{21}+z_{22}+\cdots+z_{2n1}+z_{11}+\cdots+z_{1n1})$$

$$=\sqrt{n}(a-\alpha) \tag{77}$$

ただし，$a=\bar{y}$

$$u_3=\frac{1}{\sqrt{\Sigma(x_2-\bar{x}_2)^2+\Sigma(x_1-\bar{x}_1)^2}}\left\{\Sigma(x_2-\bar{x}_2)z_2+\Sigma(x_1-\bar{x}_1)z_1\right\}$$

$$=\frac{1}{\sqrt{\Sigma(x_2-\bar{x}^2)^2+\Sigma(x_1-\bar{x}_1)^2}}\Big[\Sigma(x_2-\bar{x}_2)(y_2-\bar{y}_2)+\Sigma(x_1-\bar{x}_1)$$

$$(y_1-\bar{y}_1)-\beta\left\{\Sigma(x_2-\bar{x}_2)^2+\Sigma(x_1-\bar{x}_1)^2\right\}\Big]$$

$$=\frac{\sqrt{n}}{s'_x}(s'_{xy}-\beta(s'_x)^2) \tag{78}$$

ただし[6]

(6) $\sum_{i=1}^{n_2}(x_{2i}-\bar{x}_2)^2$ を簡単化のため，$\Sigma(x_2-\bar{x}_2)^2$ と書く。他も同様。

$$(s'_x)^2=\frac{\Sigma(x_1-\bar{x}_1)^2+\Sigma(x_2-\bar{x}_2)^2}{n} \tag{79}$$

$$s'_{xy}=\frac{\Sigma(x_1-\bar{x}_1)(y_1-\bar{y}_1)+\Sigma(x_2-\bar{x}_2)(y_2-\bar{y}_2)}{n} \tag{80}$$

$$\begin{aligned} u_4&=\Sigma\, a_{42i}\, z_{2i}+\Sigma\, a_{41i}\, z_{1j} \\ &\;\vdots \\ u_n&=\Sigma\, a_{n2i}\, z_{ni}+\Sigma\, a_{n1i}\, z_{1j} \end{aligned} \tag{81}$$

u_4 から u_n までの右辺の係数 a は，直交変換を与えるように選ばれたものとする。

z_i は $N(0,\ \sigma_y{}^2(1-\rho^2))$ にしたがって独立に分布するので，u_i も同様に $N(0,\ \sigma_y{}^2(1-\rho^2))$ にしたがって独立に分布することが容易に導かれる。

以上の結果から

$$b=\sqrt{\frac{n}{n_1n_2}}\frac{u_1}{(\bar{x}_2-\bar{x}_1)}+\beta \tag{82}$$

$$a=\frac{n_2}{\sqrt{n}}+\alpha \tag{83}$$

は，それぞれ $N(\beta,\ n\sigma_y{}^2(1-\rho^2)/n_1n_2(\bar{x}_2-\bar{x}_1)^2)$ および $N(\alpha,\ \sigma_y{}^2(1-\rho^2)/n)$ にしたがって独立に分布することが導かれる。またこれらは $(u_4{}^2+\cdots+u_n{}^2)/\sigma_y{}^2(1-\rho^2)$ とはそれぞれ独立に分布する。　ところで $(u_4{}^2+\cdots+u_n{}^2)/\sigma_y{}^2(1-\rho^2)$ は自由度 $(n-3)$ の χ^2 分布をすることは容易に分るであろう。ところで

$$\begin{aligned} w=\sum_{i=1}^{4}u_i{}^2&=\sum_{j=1}^{n}z_j-u_1{}^2-u_2{}^2-u_3{}^2 \\ &=\Sigma(y-\alpha-\beta x)^2-n(a-\alpha)^2-\frac{n_1n_2}{n}(\bar{x}_2-\bar{x}_1)^2(b-\beta)^2 \\ &\qquad-\frac{n}{(s'_x)^2}\left\{s'_{xy}-\beta(s'_x)^2\right\}^2 \end{aligned} \tag{84}$$

ところで，この結果はつぎのような簡単な形のものになる。すなわち

$$w=n(s'_y)^2\left\{1-\frac{(s'_{xy})^2}{(s'_x)^2(s'_y)^2}\right\} \tag{85}$$

この形は，通常の回帰直線における誤差の変動に類似のものとなる。すな

わち，

$$\sum_{i=1}^{n}(y_i-a-bx_i)^2=ns_y^2\left(1-\frac{s_{xy}^2}{s_x^2 s_y^2}\right)$$

$$=ns_y^2(1-r^2) \tag{86}$$

この自由度は（$n-2$）であるが，われわれの上の結果は自由度が（$n-3$）となっていることに注意。

さて w は a および b とは独立に自由度が（$n-3$）の χ^2 分布をし，$\sqrt{n}(a-\alpha)/\sigma_y\sqrt{1-\rho^2}$ および $\sqrt{n_1n_2}(\bar{x}_2-\bar{x}_1)(b-\beta)/n\sigma_y\sqrt{1-\rho^2}$ はそれぞれ $N(0,1)$ であるから，

$$t_1=\frac{\sqrt{n}(a-\alpha)}{\sigma_y\sqrt{1-\rho^2}}\Bigg/\sqrt{\frac{w}{\sigma_y^2(1-\rho^2)}\Big/(n-3)}$$

$$=\sqrt{n}(a-\alpha)\Bigg/\sqrt{\frac{w}{(n-3)}} \tag{87}$$

$$t_2=\sqrt{\frac{n_1n_2}{n}}\,\frac{(\bar{x}_2-\bar{x}_1)(b-\beta)}{\sigma_y\sqrt{1-\rho^2}}\Bigg/\sqrt{\frac{w}{\sigma_y^2(1-\rho^2)}\Big/(n-3)}$$

$$=\sqrt{\frac{n_1n_2}{n}}(\bar{x}_2-\bar{x}_1)(b-\beta)\Bigg/\sqrt{\frac{w}{(n-3)}} \tag{88}$$

はそれぞれ，自由度が（$n-3$）の t 分布をすることになる。したがって t_1 および t_2 なる統計量を用いて，それぞれ α および β の検定，ならびに信頼限界の設定が可能となる。ただし，推定量そのものに比べて計算は若干複雑なものになることは否めない。

（1964年2月9日）

（後記：印刷の便宜を考え，今回は前回とは記号の使い方が若干異なっている。）