6. 統計量の独立性と自由度の問題

坂元平八

(19年8月14日受付)

I. 正規標本相関係数の分布法則

正規標本相関係数の分布法則を求める問題はR.A.Fisher[1]が幾何学的方法に依て解決し，この意味をD.Jackson[2]が解析的に解釈して説明している。亦S.Kullback[3]等は特性函数を用いて矢張同じ分布法則を導き出している。以下本稿では統計量の独立性を利用して，この分布法則を求める一方法を示そう。

(1) R.A.Fisher; Frequency Distribution of the Values of the Carrelation Coefficient in Samples from an infinitely Large pakulation, Biometrika, Val X. (1914-1915) 517-521.

(2) Dumhum Jackson; mathematical principle's in the Theory of Small Samples. The American mathematical monthly, Vol XIII 1935. Number 6, June-July. 344-364.

(3) Salomon kullback; An application of characteristic functions to the distrukation problem of statistics. The Annals of mathematical statistics. Vol. V. 1934
Wishart. J; The generalized product moment distribution in samples from a normal multivariate nokulation. Biometrikas. Vol XXA (1928) P.P. 32-52

今(x_1, y_1) $(x_2, y_2), \cdots\cdots (x_n, y_n)$を分布法則が

$$\frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} e^{-\frac{1}{1-\rho^2}\left(\frac{x^2}{2\sigma_1^2}-\frac{2\rho xy}{2\sigma_1\sigma_2}+\frac{y^2}{2\sigma_2^2}\right)}dxdy \qquad \text{--------} (1)$$

に従ふ様な二変量正規相関母集団から独立に取れるn組の標本値からなる一つの標本とする。然る時，$(x_i, y_i)(i=1, 2, \text{------}, n)$の同時分布法則は

$$\frac{1}{(2\pi\sigma_1\sigma_2)^n(1-\rho^2)^{\frac{n}{2}}} e^{-\frac{1}{1-\rho^2}\left(\frac{(x\cdot x)}{2\sigma_2^2}-\frac{2\rho(x\cdot y)}{2\sigma_1\sigma_2}+\frac{(y\cdot y)}{2\sigma_2^2}\right)}dxdy \qquad \text{---------} (2)$$

で與へられる(4) 但しここに,4)

$$\mathfrak{x} = (x_1, x_2 \cdots\cdots, x_n), \quad \mathfrak{y} = (y_1, y_2 \cdots\cdots, y_n)$$

とする。

(4) 坂元平八「統計量の独立性に就て」、數学物理学会誌(1944)

の記号用例を参照の事.

(2) に於て, $\mathfrak{x}$, $\mathfrak{y}$ なる變數を $\mathfrak{x}$, $\mathfrak{z} = \frac{1}{(1-\rho^2)^{\frac{1}{2}}} \frac{\sigma_1}{\sigma_2} \mathfrak{y} - \frac{\rho}{(1-\rho^2)^{\frac{1}{2}}} \mathfrak{x}$ なる變數に變換すれば そのヤコービ行列式の値が $\left\{ (1-\rho^2)^{\frac{1}{2}} \frac{\sigma_2}{\sigma_1} \right\}^n$ であるから (2) は

$$\frac{1}{(2\pi\sigma_1^2)^n} e^{-\frac{1}{2\sigma_1^2}\{(\mathfrak{x},\mathfrak{x})+\mathfrak{z}\mathfrak{z})\}} d\mathfrak{x}, d\mathfrak{z} \cdots\cdots (3)$$

なる $\mathfrak{x}$ と $\mathfrak{z}$ の同時分布法則を示す式に變換される。

今こゝに n 次の対稱行列 A があり $AA' = A$ なる性質があり, 且

A の階數は m であるとせよ(5)

斯かる A に対して

$$R = \frac{(A\mathfrak{x},\ A\mathfrak{y})}{|A\mathfrak{x}||A\mathfrak{y}|}$$

の分布法則を求めるのが本稿の目的である。そのために先づ

$$r = \frac{(A\mathfrak{x};\ A\mathfrak{y})}{|A\mathfrak{x}||A\mathfrak{y}|}$$

4) 坂元平八「統計量の独立性に就て」數学物理学会誌(1944)の記号同例を参照の事.
5) 坂元平八「 〃 」の[定理 II]参照の事.

の分布法則を求める．但し $|A\mathfrak{x}|$ は $(A\mathfrak{x}, A\mathfrak{x})^{\frac{1}{2}} = (A\mathfrak{x}, \mathfrak{x})^{\frac{1}{2}}$ を示すものとする．扨 $\dfrac{(A\mathfrak{x},\mathfrak{x})}{\sigma_1^2}$, $\dfrac{(A\mathfrak{z},\mathfrak{z})}{\sigma_1^2}$, $\dfrac{(A\mathfrak{x}, A\mathfrak{z})}{|A\mathfrak{x}||A\mathfrak{y}|}$ なる三つの統計量の同時分布の特性函数を計算すると．

$$\varphi(t_1, t_2, t_3) = \frac{1}{(2\pi\sigma_1^2)^n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-\frac{(\mathfrak{x},\mathfrak{x})+(\mathfrak{z},\mathfrak{z})}{2\sigma_1^2} + it_1\frac{(A\mathfrak{x},\mathfrak{x})}{\sigma_1^2} + it_2\frac{(A\mathfrak{z},\mathfrak{z})}{\sigma_1^2} + it_3\frac{(A\mathfrak{x},\mathfrak{z})}{|A\mathfrak{x}||A\mathfrak{z}|}}\, d\mathfrak{x}\, d\mathfrak{z}$$

扨 A は適当な直交行列 T を選ぶことに依って

$$T'A\,T = \begin{pmatrix} \overbrace{1 \cdots 1}^{m} & & 0 \\ & \overbrace{0 \cdots 0}^{n-m} & \\ 0 & & \end{pmatrix}$$ ならしめ得るから

上式に於て $\mathfrak{x}, \mathfrak{z}$ に夫々 T なる一次変換を施せば，そのヤコービ行列式の値は 1 である．従って

$$\varphi(t_1, t_2, t_3) = \frac{1}{(2\pi\sigma_1^2)^n}\overbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}^{2n} e^{-\frac{(\mathfrak{x},\mathfrak{x})+(\mathfrak{z},\mathfrak{z})}{2\sigma_2^2} + it_2\frac{(\mathfrak{x}',\mathfrak{x}')}{\sigma_1^2} + it_2\frac{(\mathfrak{z}',\mathfrak{z}')}{\sigma_1^2} + it_3\frac{(\mathfrak{x}',\mathfrak{z}')}{|\mathfrak{x}'||\mathfrak{z}'|}}\, d\mathfrak{x}\, d\mathfrak{z}$$

$$= \frac{1}{(1-2it_1)^{\frac{m}{2}}} \frac{1}{(1-2it_2)^{\frac{m}{2}}} \frac{1}{(2\pi)^m} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\frac{(\mathfrak{r}',\mathfrak{r}')+(\mathfrak{z}',\mathfrak{z}')}{2} + it_3 \frac{(\mathfrak{r}',\mathfrak{z}')}{|\mathfrak{r}'||\mathfrak{z}'|}} d\mathfrak{r}' d\mathfrak{z}'^{(\text{註})} \cdots\cdots (4)$$

(註) こゝで積分の範囲が矢張実数軸上で$(-\infty, \infty)$なることは

簡単に分る.

但しこゝに $\mathfrak{r}' = (x_1, x_2, \cdots\cdots, x_n)$, $\mathfrak{z}' = (z_1, z_2 \cdots\cdots, z_n)$ とする. 依って$\varphi(t_1, t_2, t_3) = \varphi_1(t_1)\varphi_2(t_2)\varphi_3(t_3)$の形に書け $\varphi_1(t_1)$, $\varphi_2(t_2)$, $\varphi(t_3)$は夫々

$$a^2 = \frac{(A\mathfrak{r}, \mathfrak{r})}{\sigma_1^2}, \quad b^2 = \frac{(A\mathfrak{z}, \mathfrak{z})}{\sigma_1^2}, \quad r = \frac{(A\mathfrak{r}, \mathfrak{z})}{|A\mathfrak{r}||A\mathfrak{z}|}$$

なる統計量の特性函数を表すから a^2, b^2, r は相互に独立であり且 a^2, b^2 の分布法則はその特性函数から明らかに自由度mのχ^2分布をなす。汞rの分布法則もその特性函数が

$$\varphi_3(t_3) = \overbrace{\int_{-\infty}^{\infty} \cdots\cdots \int_{-\infty}^{\infty}}^{2m} \frac{1}{(2\pi)^m} e^{-\frac{(\mathfrak{r}',\mathfrak{r}')+(\mathfrak{z}',\mathfrak{z}')}{2} + it_3 \frac{(\mathfrak{r}',\mathfrak{z}')}{|\mathfrak{r}'||\mathfrak{z}'|}} d\mathfrak{r}' d\mathfrak{z}' \cdots\cdots\cdots (5)$$

なる計算で得られる事から容易に[7]

$$\beta(r)=\frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{m-1}{2}\right)}\,(1-r^2)^{\frac{m-3}{2}}$$

で與へられることが証明出来る.

依って a, b, r の同時分布法則は

$$f(a,b,r)\,da,db,dr=\frac{a^{m-1}e^{-\frac{a^2}{2}}\,b^{m-1}\,e^{-\frac{b^2}{2}}\quad\Gamma\left(\frac{m}{2}\right)}{2^{\frac{m-2}{2}}\,\Gamma\left(\frac{m}{2}\right)\,2^{\frac{m-2}{2}}\Gamma\left(\frac{m}{2}\right)\sqrt{\pi}\,\Gamma\left(\frac{m-1}{2}\right)}$$

$$(1-r^2)^{\frac{m-3}{2}}\;da\,db\,dr \text{ - - - - - - - - - - - } (6)$$

で與へられる. 故 $\zeta=\dfrac{1}{(1-\rho^2)^{\frac{1}{2}}}\;\dfrac{\sigma_1}{\sigma_2}\,\eta-\dfrac{\rho}{(1-\rho^2)^{\frac{1}{2}}}\,\xi$

であるから

$$A\zeta=\frac{1}{(1-\rho^2)^{\frac{1}{2}}}\;\frac{\sigma_1}{\sigma_2}\,A\eta-\frac{\rho}{(1-\rho^2)^{\frac{1}{2}}}\,A\xi$$

である。此処で $\quad r=\dfrac{(A\xi,A\zeta)}{|A\xi||A\zeta|}=\cos\theta,$

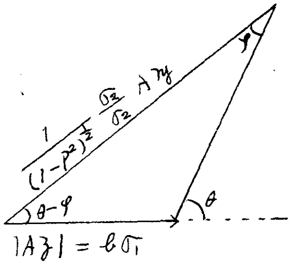
$$R=\frac{(A\xi,A\zeta)}{|A\xi||A\zeta|}=\cos\varphi \text{ と置けば}$$

$(0\leqq\theta\leqq\pi,\ 0\leqq\varphi\leqq\pi)$, (7) は下図の如き関係を示す.

---

(7) 此は r の分布法則が $\dfrac{1}{(\sqrt{2\pi})^m}\,e^{-\frac{(\xi',\xi')+(\zeta',\zeta')}{2}}\,d\xi'd\zeta'$
なる分布法則に従ふ m 次元空間の二つのヴェクトルのなす角の余弦
即ち $\cos\theta=\dfrac{(\xi,\zeta)}{|\xi||\zeta|}$ の分布法則に等しいことを意味している.
これより容易

KとYの分布法則が與へられる。(2)のJacksonの論文を参照されたし。



故に

$$\left|\frac{\rho}{(1-\rho^2)^{\frac{1}{2}}}A\tau\right|=\frac{\rho}{(1-\rho^2)^{\frac{1}{2}}}a\sigma_2$$

$$a=\frac{(1-\rho^2)^{\frac{1}{2}}}{\rho}b\frac{\sin(\theta-\varphi)}{\sin\varphi}\cdots\cdots(8)$$

なる関係を得。従って(6)に於てY, aなる変数をθ,φなる変数に変換しbに就て(0. ∞)で積分すれば

$$-\frac{n-2}{2\pi}(1-\rho^2)^{\frac{m}{2}}\frac{\{\sin\theta\sin\varphi\sin(\theta-\varphi)\}^{m-1}}{\{(1-\rho^2)\sin^2(\theta-\varphi)+\rho^2\sin^2\varphi\}^m}d\theta d\varphi$$

$$\cdots\cdots\cdots\cdots\cdots\cdots(9)$$

此処で$\frac{\sin(\theta-\varphi)}{\rho\sin\theta}=e^{z}$と置いて$\theta$なる変数を$z$なる変数に変換し$\varphi$の分布法則を求めると

$$-\frac{m-1}{\pi}(1-\rho^2)^{\frac{m}{2}}\sin^{m-2}\varphi\int_0^\infty\frac{dz}{(\cosh z-\rho\cos\varphi)^m}d\varphi$$

依って$\cos\varphi=R$と置けばRの分布法則として

$$\frac{(m-1)}{\pi}(1-\rho^2)^{\frac{m}{2}}(1-R^2)^{\frac{m-3}{2}}\int_0^\infty\frac{dz}{(\cosh z-\rho R)^m}dR\cdots\cdots(10)$$

を得(8)

II 応用例

次に本論説の実際の場合に於ける応用に就て述べる.

(a) 今 $(x_1, y_1), (x_2, y_2) \cdots\cdots (x_n, y_n)$ を (1) なる分布法則に従ふ様な二変量正規母集団から独立に取れる n 組の標本値からなる一つの標本とする. これから計算せる見本相関係数を

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{m}(t_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{(A\mathfrak{x}, \mathfrak{y})}{|A\mathfrak{x}||A\mathfrak{y}|} \cdots\cdots (11)$$

とする。こゝに $\mathfrak{x} = (x_1, x_2, \cdots\cdots x_n)$,

$$\mathfrak{y} = (y_1, y_2, \cdots\cdots y_n).$$

$$A = \begin{pmatrix} 1-\frac{1}{n}, & -\frac{1}{n}, & \cdots\cdots\cdots\cdots, & -\frac{1}{n} \\ -\frac{1}{n}, & 1-\frac{1}{n}, & \cdots\cdots\cdots\cdots, & -\frac{1}{n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ -\frac{1}{n}, & -\frac{1}{n}, & \cdots\cdots\cdots\cdots, & 1-\frac{1}{n} \end{pmatrix}$$

と置く. $AA' = A$ で A の階数は $(n-1)$ であるから R は (10) に於て $m = n-1$ と置いた分布法則で與へられる。

(6) (1) なる分布法則に従う様な二変量正規母集団から次の如き二つの標本

$(x_1, y_1), (x_2, y_2) \cdots\cdots\cdots (x_{n_1}, y_{n_2})$

及び $(x'_1, y'_2), (x'_2, y'_2) \cdots\cdots\cdots (x'_{n_2}, y'_{n_2})$

をとる。その各々について $\varphi = \sum_{i=1}^{x} (x_i - \bar{x})(y_i - \bar{y})$

$\eta^2 = \sum_{i=1}^{n_1} (x_i - \bar{x})^2, \quad \varphi^2 = \sum_{i=1}^{n_1} (y_i - \bar{y})^2$

$\varphi' = \sum_{i=1}^{n_2} (x'_i - \bar{x}')(y_i - \bar{y}') \quad \eta'^2 = \sum_{i=1}^{n_2} (x'_i - \bar{x}')^2$

$\varphi'^2 = \sum_{i=1}^{n_2} (y'_i - \bar{y}')^2$ が計算してあるとする。此の二つより次の如くしてRを計算する即ち

$$R = \frac{\varphi + \varphi'}{\sqrt{\eta^2 + \eta'^2}\sqrt{\varphi^2 + \varphi'^2}} = \frac{(A\mathfrak{x}, \mathfrak{y})}{|A\mathfrak{x}||A\mathfrak{y}|}$$ で計算する

但しこゝに $\mathfrak{x} = (x_1, x_2, \cdots, x_{n_1}, x'_1, x'_2 \cdots x'_{n_2})$

$\mathfrak{y} = (y_1, y_2, \cdots y_{n_1}, y'_1, y'_2 \cdots y'_{n_2})$

$$A = \left(\begin{array}{cccc|cccc} 1-\frac{1}{n}, & -\frac{1}{n}, & \cdots, & -\frac{1}{n}, & & & & \\ -\frac{1}{n}, & 1-\frac{1}{n}, & \cdots, & -\frac{1}{n}, & & 0 & & \\ \cdots & \cdots & \cdots & \cdots & & & & \\ -\frac{1}{n}, & -\frac{1}{n}, & \cdots & 1-\frac{1}{n}, & & & & \\ \hline & & & & 1-\frac{1}{n}, & -\frac{1}{n_2} & \cdots & -\frac{1}{n_2} \\ & 0 & & & -\frac{1}{n_2}, & 1-\frac{1}{n_2} & \cdots & -\frac{1}{n_2} \end{array}\right)$$

$$\left(\quad\quad\quad\quad \begin{array}{|l} \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ -\frac{1}{n_1}, -\frac{1}{n_2} \cdots\cdots, 1-\frac{1}{n_2} \end{array}\right)$$

とする。$AA' = A$ で A の階数は $n_1 + n_2 - 2$ であるから容易に R は (10) に於て $m = n_1 + n_2 - 2$ と置きたる分布法則に従ふことが云へる。