評価関数の成分分析の適用例

中　島　達　夫

事象またはものの集合のなかから，その構成要素についていくつかの特性を総合評価して，ある部分集合を選抜する場合を考える．具体例としては，数科目の試験結果の総合評価によって，応募者のなかから一定数の合格者を決める選抜試験があげられよう．

いま評価すべき特性すなわち評価成分をA，Bのふたつとし，それぞれの得点のウエイトつきの総和により選抜するものとする．選抜率を Po とする．

A，B の得点を x_1，x_2とし，重み係数 C_1，C_2をつけ

$$E_0 = C_1 x_1 + C_2 x_2 \tag{1}$$

の値により判定する．全体のA，Bの得点の平均を μ_1，μ_2，標準偏差を σ_1，σ_2とする．

$$u = (x_1 - \mu_1)/\sigma_1,\quad v = (x_2 - \mu_2)/\sigma_2$$

$$C = C_2\sigma_2 / C_1\sigma_1 \tag{2}$$

と正規化すると，評価関数として

$$E = u + Cv \tag{3}$$

が得られる．この係数CがAに対するBの評価比重を示すもので，C の値が小さいほどAに対してBのウエイトが小さくなるものと考えてよい．

なお u，v はともに標準正規分布 $N(0, 1)$に したがい，両者の相関係数は ρ であるとする．したがって u，v の確率密度は次式で示される．

$$f(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}}\, exp \left\{ -\frac{1}{2(1-\rho^2)}(u^2 - 2\rho uv + v^2) \right\} \tag{4}$$

図1に示すように Lu，Lv をそれぞれ u，v 軸に直交し，その外側確率が P_0である直線とする．すなわち Lu，Lv の座標軸切片長を a とすると，

$$\Phi(a) = \frac{1}{\sqrt{2\pi}}\int_0^a e^{-t^2/2}dt = 1/2 - Po \tag{5}$$

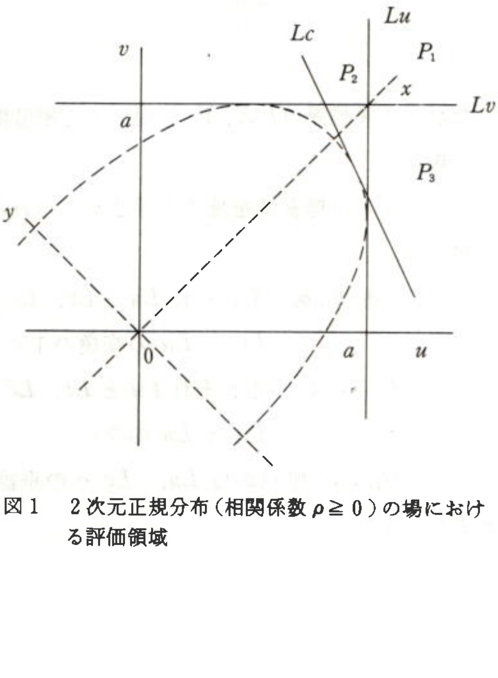

図1　2次元正規分布（相関係数 $\rho \geqq 0$）の場における評価領域

評価を A，B いずれか単独で行うとき Lu，Lv がそれぞれの合格限界線を示し，この線の右側または上側に入るものが合格となる．

つぎに評価関数

$$E=u+Cv$$

で判定する場合にはその合格限界線 Lc は，u，v 軸との切片長比が C：1 であり，その外側確率が Po の直線となる．3 線 Lu，Lv，Lc はともに等確率密度線

$$u^2-2\rho uv+v^2=a^2$$

の接線となる．

Lu と Lv の外側の共通領域内の確率を P_1，Lc の外側と Lu の内側の共通領域内の確率を P_2，Lc の外側と Lv の内側の共通領域内の確率を P_3 とすると，

P_1＝A，B それぞれ単独判定の場合に両方に合格する人の割合

P_2＝A 単独のとき不合格であるが，A，B の複合判定のときに合格する人の割合

P_3＝B 単独のとき不合格であるが，A，B の複合判定のときに合格する人の割合

を意味し，また

P_0-P_1＝単独評価のとき一方で合格，他方で不合格となる人の割合

である．

つぎに確率 P_1，P_2，P_3 を計算してみよう．

座標軸 u,v を45°回転したものを x，y 軸とし，さらに

$$X=x/\sqrt{1+\rho},\quad Y=y/\sqrt{1-\rho} \tag{6}$$

と座標変換すると

$$f(X,\ Y)=\frac{1}{2\pi}exp\left\{-\frac{1}{2}(X^2+Y^2)\right\} \tag{7}$$

したがって座標系 O-X, Y を用いると密度関数は標準正規分布となる．

図 1 にこの座標変換を施すと図 2 がえられる．図 2 において

$\theta_1, \theta_2, \theta_3$：それぞれ Lu と Lv，Lc と Lv，Lc と Lu の交角の 1／2

G_1, G_2, G_3：それぞれ Lu と Lv，Lc と Lv，Lc と Lu の交点

H_1, H_2：原点から Lu，Lc への垂線の足

とする．また

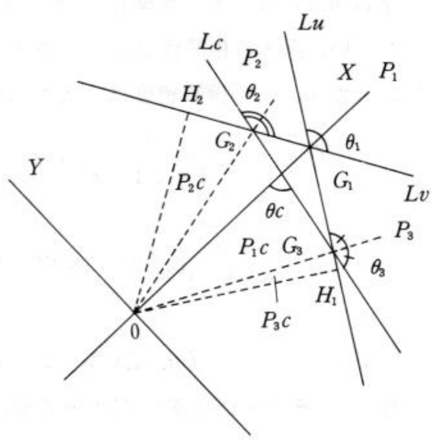

図 2　2 次元正規分布（相関係数 $\rho=0$）の場への変換

P_{1c}, P_{2c}, P_{3c}：それぞれΔOG_1H_1，ΔOG_2H_2，ΔOG_3H_1 の領域内の確率とする．$\theta_1, \theta_2, \theta_3$ は次式で計算される．

$$\left.\begin{array}{l}\theta_1=\tan^{-1}\sqrt{(1+\rho)/(1-\rho)}\\ \theta_2=(\theta_1+\theta_c)/2\\ \theta_3=(\theta_1-\theta_c+\pi)/2\end{array}\right\} \quad (8)$$

ただし

$$\theta c=\tan^{-1}\left(\frac{1+C}{1-C}\sqrt{\frac{1+\rho}{1-\rho}}\right) \quad (9)$$

P_{ic}（$i=1, 2, 3$）を数値積分を用いて計算し，

$$P_i=P_0+2P_{ic}+\theta_i/\pi-1/2$$

によってP_1，P_2，P_3を求めることができる．計算結果例を図３に示す．

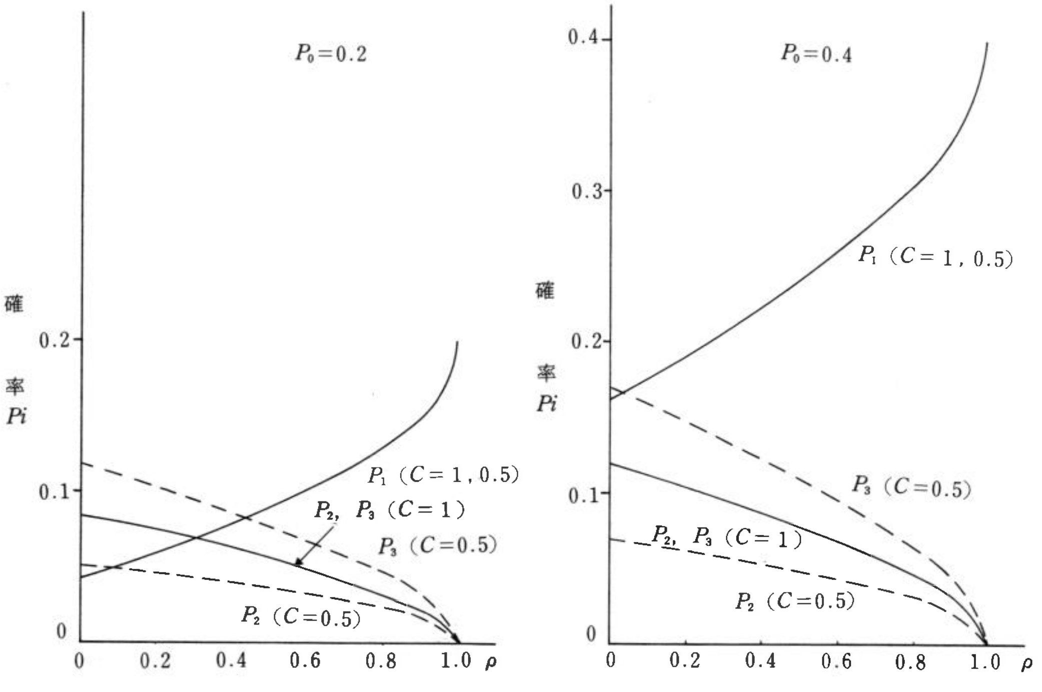

図３　確率 P_1，P_2，P_3 の計算結果

図より明らかなように相関係数 ρ が大きくなるにつれ P_1 が増大し，P_2，P_3が減少する．またウエイト係数 C が小さくなるほど P_2 が小さくなり，逆に P_3 が大きくなる．