数理解析研究所講究録
44巻 1968年 79-86

Multivariate Tests with Restricted Alternatives

九大 理学部 工藤昭夫

§1. 序

p次元正規分布$N(\underset{\sim}{\theta}, \Sigma)$について。帰無仮説 $H_0 : \underset{\sim}{\theta} = \underset{\sim}{0}$ の検定法は $H_1 : \underset{\sim}{\theta} \neq \underset{\sim}{0}$ のときには次の3種の事情に応じての検定法がある。a) Σ 既知のとき、b) $\Sigma = \sigma^2 \Lambda$と書けて Λは既知、σ^2未知であるが χ^2分布（自由度m）をし 標本平均ベクトルと独立な $m\sigma^2$の推定値が存在する場合 c) Σが全く未知であるときの それぞれの場合の検定は a) 自由度pのχ^2分布によるもの、b) 自由度p, n のF 分布によるもの、c) Hotelling T^2 を用いた検定があり それらの検定の性質もよく知られている。$|\Sigma| = 0$のときには a) および b) の場合には pの代りに$rank(\Sigma)$を用いればよいことも知られている。

ここでは $\underset{\sim}{\theta}' = (\theta_1, \cdots, \theta_p)$ としたとき $H_1 : \theta_i \geq 0$ $i=1, \cdots, p$（p個の不等式のうち少くとも1個は$\neq$） 即ち

multivariate analogue of sided test および H_1: $\theta \geq 0$ $(i=1,\cdots,p)$ 或いは $\theta_i \leq 0$ $(i=1,\cdots,p)$, などの形の対立仮説を持つ検定についての研究を既発表のものと未発表のものをまとめて総合報告をする。

§2. Σ が既知の場合の片側検定

大きさnの標本の平均値ベクトルを $\bar{x}$ とする。尤度比検定の統計量は $n\{\bar{x}'\Sigma^{-1}\bar{x} - \underset{\substack{\theta_i \geq 0 \\ i=1\cdots p}}{\mathrm{Min}} (\bar{x}-\theta)'\Sigma^{-1}(\bar{x}-\theta)\} = n\{\bar{x}'\Sigma^{-1}\bar{x} - (\bar{x}-x^0)'\Sigma^{-1}(\bar{x}-x^0)\}$ である。右辺の x^0 は対立仮説のもとでの平均値ベクトルの最尤推定で unique に存在する。

この統計量の分布は次の定理により与えられる。統計量を $\bar{\chi}^2$ と書くと

$$\Pr(\bar{\chi}^2 \geq \bar{\chi}_0^2) = \sum_{\phi \leq M \leq P} \Pr(\chi^2_{n(M)} \geq \bar{\chi}_0^2) P\{(\Sigma_{M'})^{-1}\} P\{\Sigma_{M:M'}\}$$

ここで Σ は $P=(1,\ldots,p)$ の部分集合M（例えば $M_0 = (1,2,\cdots,m)$）についての和で Mが空集合ϕのこともあり得る。$\chi^2_{n(M)}$ はMの要素の数（$n(M_0)=m$）を自由度に持つ χ^2変量 $P\{\Lambda\}$ は $\underset{\sim}{z}$が $N(0,\Lambda)$ に従って分布するとき、$\underset{\sim}{z}$のすべての component が≥ 0である確率 M'はMの補集合（$M_0' = (m+1,\cdots,p)$）$\Sigma_{M'}$は x_i, $i \notin M$ の分散行列，（$\Sigma_{M_0'}$は，$x_{m+1},\cdots,x_p$の分散行列）$\Sigma_{M:M'}$は x_j : $j \in M'$を条件うけたときの x_i $i \in M$ の分散行列（$\Sigma_{M_0:M_0'}$は $x_{m+1},\cdots, x_p$ を止めたときの $(x_1,\ldots,x_m)$ の条件附分散行列。$\chi_0^2 = 0$,

$P(\Sigma_{p'}^{-1}) = P(\Sigma_{p_i p'}) = 1$

この分布を導くのには次の定理が必要である。k次元空間中に独立なk個のベクトル$(a_1, \cdots, a_k)$を取り、$S = \{y : (a_i, y) \geq 0,\ i = 1, \cdots, k\}$とする。$S$は closed convex polyhedral cone になる。Sの外点y_1に最も近いSの点をy_0とし、部分空間$\{y : (a_i y) = 0,\ i = 1, \ldots, m\}$への$y_1$の射影を$\hat{y}_1$と書く。$a^{(i)}$を$(a_1, \ldots, a_m)$ではる空間にはいり$(a^{(i)}, a_j) = \delta_{ij}$ $(i, j = 1, \ldots, m)$であるようなベクトルとする。

定理 1. $(a_i, y_0) = 0,\ i = 1, \cdots, m,\ (a_i, y_0) > 0\ i = m+1 \cdots k$ であるための必要十分条件は $(a^{(i)}, y_1) \leq 0,\ i = 1, \cdots, m,$ $(a_i, \hat{y}_1) > 0\quad i = m+1, \cdots, k$

$\bar{X}$の分散行列$\frac{1}{n}\Sigma$に対して$\frac{1}{n}A\Sigma A' = I$となる行列$A$を取り$Y = A\bar{X}$と変数変換をすれば、密度関数は$C \exp\left[-\frac{1}{2}(y - A\theta)'(y - A\theta)\right]$と書けるので対立仮説のもとでの$\theta$、最尤推定の問題は、$y$が与えられたとき

$\underset{A^{-1}m \geq 0}{\mathrm{Min}}\ (y - m)'(y - m) = (y - m_0)'(y - m_0)$ を満足するm_0を求める問題に帰着する。θの最尤推定量$\hat{\theta}$のコンポーネントは幾つかは$= 0$（例えば$\theta_i = 0,\ i \notin M$）幾つかは$> 0$ $(\theta_i > 0,\ i \in M)$。MはPの部分集合のどれかであり2^pどうりの可能性がある。標本平均値ベクトルの空間もそれに対応して2^p通りにわけられる。定理1は yが標本平均値の空間のどこには

いるかを特徴づける定理である。

この§の内容は Kudô 1963, Nüesch 1966 に拠っている

§3 Σ が既知の場合の両側検定

ここで云う両側検定とは $H_1; \theta_i \geqq 0$ $(i=1,\cdots,p)$ 或いは $\theta_i \leqq 0$ $(i=1,\cdots,p)$ (どちらの場合でも少くとも一つの i について $\neq 0$ が成立する) を対立仮説とする検定である。対立仮説の空間は 片側検定の場合の対立仮説 $H_1^{(+)}$ の空間 $\Omega^{(+)}$ と $\geqq 0$ を $\leqq 0$ で入れ替えて出来た対立仮定 $H^{(-)}$ の空間 $\Omega^{(-)}$ との和集合 $\Omega^{(+)} \cup \Omega^{(-)}$ となっている。検定問題 $(H_0, H_1^{(+)})$, と $(H_0, H_1^{(-)})$ の検定統計量を $\bar{\chi}^{(+)^2}$, $\bar{\chi}^{(-)^2}$ と書くと (H_0, H_1) のそれは $\bar{\bar{\chi}}^2 = \mathrm{Max}(\bar{\chi}^{(+)^2}, \bar{\chi}^{(-)^2})$ となるが その分布および検定力関数は Kudô and Fujisawa 1964, 1966 2次元の場合のみ求められている。次元数が2より大なる場合には 非常に複雑になるが $\Sigma = I$ の場合には

$$
\begin{aligned}
\Pr(\bar{\bar{\chi}}^2 \geqq C^2) = \frac{1}{2^{k-1}} \sum_{m=1}^{k} {}_kC_m \Pr(\chi_m^2 \geqq C^2) \\
- \frac{1}{2^k} \sum_{m=1}^{k} {}_kC_m \Pr(\chi_m^2 \geqq C^2) \Pr(\chi_{k-m}^2 \geqq C^2)
\end{aligned}
$$

が 九大の棄能哲氏により求められており、数表の作製を計画中である。

§3 Σ が既知でない場合

§1で述べたように b) c) 2つの場合があるが b)

の場合には　片側検定の検定統計量は　$n(\bar{x}-x^0)'\Lambda^{-1}(\bar{x}-x_0)/\hat{\sigma}$ となる。但し　$\hat{\sigma}^2$は　σ^2，不偏推定量である。

この統計量　$\bar{F}$の分布は

$$P_r(\bar{F} \geq \bar{F}_0)=\sum_{\phi \leq M \leq P} P_r(F_{n(M)}, m \geq n(M)\bar{F}_0) P(\Sigma_{M'}^{-1}) P(\Sigma_{M;M'})$$

となる。両側検定も同様である。c) の場合には．解くことは非常に困難で　Nüesch 1966　で提案している検定は 1．計算間違いがあるので尤度比検定と云えない。2．その結果統計量自身が未知母数　Σ　に依存する　3．H_0のもとでの分布が　Σ　に依存する欠点がある。Σを標本分散行列から推定し．それを　Σであるかの如く統計量を計算すれば　2 の欠点は克服出来るが　3．の欠点　すなはち分布が　Σ に無関係に出来るかどうかは分らない。

§4．応用と拡張

線型回帰モデル，$E(y)=\lambda_1 x_1+\cdots+\lambda_k x_k$，$V(y)=\sigma^2$を考え　正規性を仮定すれば　$\lambda_1,\ldots,\lambda_k$の線型最良推定$(\hat{\lambda}_1,\ldots,\hat{\lambda}_k)$　の分散は　部分的に既知　即ち b)　の場合になる。$\lambda_1,\ldots,\lambda_k$ の　m個の線型関数　$L_1,\ldots,L_m$　について．$H_0: L_i=0\ (i=1,\cdots,m)$．$H_1: L_i \geq 0\ (i=1,\cdots,m)$　とする検定問題は　L_i　が独立で　$m \leq k$ の場合には　L_i の最良線型推定量　に片側検定の手法を適用すればよい。Bartholomew 1959, 1961　に述べられている問題は上に述べた一般の formulation のうちにはいる。問題は $m \geq k$ の場合である。

変数変換をすれば $(\hat{\lambda}, \cdots, \hat{\lambda}_k)$ は 分散行列 $\sigma^2 I$, を持つ変量 $y_1, \cdots, y_k$ に変換出来るので、そこから問題を考えることにする。

$y_1, \cdots, y_p$ が独立で平均 $\theta_1, \cdots, \theta_p$ 共通の分散1を持つとする。$\underset{\sim}{a}_1, \cdots, \underset{\sim}{a}_k$ $(k > p)$ を p 次元 vector とし $\underset{\sim}{\theta}' = (\theta_1, \cdots, \theta_p)$ とする。$(\underset{\sim}{a}_1, \cdots, \underset{\sim}{a}_k)$ の張る空間 $\dim\{\underset{\sim}{a}_1, \cdots, \underset{\sim}{a}_k\} = p$ とする。(この仮定は必ずしも必要ではない。$H_0 : (\underset{\sim}{a}_i, \underset{\sim}{\theta}) = 0$, $H_1 : (\underset{\sim}{a}_i, \underset{\sim}{\theta}) \geq 0$ なる検定問題を考える. このとき $(\underset{\sim}{a}_i, \underset{\sim}{\theta}) \geq 0$ を満足する θ の集合は空ではないとする. 尤度比検定は簡単に求められる 即ち $\bar{\chi}^2 = \underset{\sim}{y}'\underset{\sim}{y} - \underset{(\underset{\sim}{a}_i, \underset{\sim}{\theta}) \geq 0}{\mathrm{Min}} (\underset{\sim}{y} - \underset{\sim}{\theta})'(\underset{\sim}{y} - \underset{\sim}{\theta})$ である。

この統計量の分布を求めるためには定理1の拡張が必要である。S, $\underset{\sim}{y}_1$, $\underset{\sim}{y}_0$, $\hat{\underset{\sim}{y}}_1$ を定理1の場合と同様に定義する。$\dim\{\underset{\sim}{a}_1, \cdots, \underset{\sim}{a}_m\} = l\ (\leq m)$ とする。$a_1, \cdots, a_m$ から $l-1$ 個の独立なベクトル, 一般性を失うことなく $\underset{\sim}{a}_1, \cdots, a_{l-1}$ を取る。$a^{(l)}$ としては、$a^{(l)} \in \{a_1, \cdots, a_m\}$, $(a^{(l)} a_i) = 0$, $i = 1, \cdots, l-1$, $(a^{(l)}, a_i) > 0$, $i = l, l+1, \cdots, m$ となるような $a^{(l)}$ を考える。.

定理 2. $(a_i, y_0) = 0$ $(i = 1, \cdots, m)$, $(a_i, y_0) > 0$ $(i = m+1, \cdots k)$ であるための必要十分条件は, 上により定めたすべての $a^{(l)}$ に対して $(a^{(l)}, y_1) \leq 0$, および $(a_i, \hat{y}_1) > 0$, $i = m+1$,

$\cdots, k$.

この定理を用いれば $\bar{\chi}^2$ の分布を求めることが出来るがその具体的な形は $(a_1, \cdots a_k)$ で非常に複雑な形で定まるので一般の形は省略する。

この部分は 葉, 工藤の最近の結果で. 未発表のものである。

実例として X_1, X_2, X_3 が独立な正規分布 $N(\theta_i, 1)$ $(i=1,2,3)$ をし, $H_0: \theta_1=\theta_2=\theta_3=0$, $H_1: \theta_1\geq 0, \theta_2\geq 0, \theta_3\geq 0, \theta_1+\theta_2-\theta_3\geq 0$ を持つ仮説検定問題を考える。(sample size が 1 と云う仮定, 分散が 1 と云う仮定は. 前述の理由からここでは一般性を失わない)、尤度比検定は $\bar{\chi}^2=\{x'x-\underset{(a_i',\theta)\geq 0}{\mathrm{Min}}(x-\theta)'(x-\theta)\}$ と検定統計量として持つ、但し $a_1'=(1,0,0)$, $a_2'=(0,1,0)$, $a_3'=(0,0,1)$, $a_4'=(1,1,-1)$。

定理2. を用いて $\bar{\chi}^2$の分布を計算すると

$$
\begin{aligned}
P_r(\bar{\chi}^2\geq c^2) &= P_r(\chi^2_{(0)}\geq c^2)\,P_r(\bar{x}_1\leq 0, \bar{x}_2\leq 0, \bar{x}_1+\bar{x}_3\leq 0, \bar{x}_2+\bar{x}_3\leq 0)\\
&+P_r(\chi^2_{(1)}\geq c^2)[P_r(\bar{x}_1>0, \bar{x}_2\leq 0, \bar{x}_3\leq 0)+P_r(\bar{x}_2>0, \bar{x}_3\leq 0, \bar{x}_1\leq 0)\\
&\quad+P_r(\bar{x}_2+\bar{x}_3>0, \bar{x}_2-\bar{x}_3\leq 0, 2\bar{x}_1-\bar{x}_2+\bar{x}_3\leq 0)\\
&\quad+P_r(\bar{x}_1+\bar{x}_3)>0, \bar{x}_1-\bar{x}_3\leq 0, 2\bar{x}_2-\bar{x}_1+\bar{x}_3\leq 0]\\
&+P_r(\chi^2_{(2)}\geq c^2)[P_r(\bar{x}_1>0, \bar{x}_2>0, \bar{x}_3\leq 0)+P_r(\bar{x}_3>0, \bar{x}_2-\bar{x}_3>0,\\
&\quad\bar{x}_1\leq 0)+P_r(\bar{x}_1-\bar{x}_3>0, \bar{x}_3>0, \bar{x}_2\leq 0)+P_r(2\bar{x}_1-\bar{x}_2+\bar{x}_3>0,\\
&\quad 2\bar{x}_2-\bar{x}_1+\bar{x}_3>0,\ \bar{x}_1+\bar{x}_2-x_3\leq 0)]
\end{aligned}
$$

$$+ P_r(\bar{\chi}^2_{(3)} \geq c^2) P_r(\bar{x}_1 > 0, \bar{x}_2 > 0, \bar{x}_3 > 0, \bar{x}_1 + \bar{x}_2 - \bar{x}_3 > 0)$$

$$= \frac{1}{6} P_r(\bar{\chi}^2_{(0)} \geq c^2) + \left[\frac{3}{4} - \frac{1}{4\pi} \cos^{-1}\left(-\frac{1}{\sqrt{3}}\right)\right] P_r\left(\bar{\chi}^2_{(1)} \geq \frac{c^2}{n}\right)$$

$$+ \frac{1}{3} P_r(\bar{\chi}^2_{(2)} \geq c^2) + \left[\frac{3}{8} - \frac{1}{4\pi}\left(\cos^{-1}\left(\frac{-2}{\sqrt{6}}\right) + \cos^{-1}\left(\frac{1}{\sqrt{3}}\right)\right)\right] P_r\left(\bar{\chi}^2_{(3)} \geq \frac{c^2}{n}\right)$$

となる。

References

Bartholomew, D. J., A test of homogeneity for ordered alternatives. Biometrika, 46, 35-48, (1959a)

Bartholomew, D. J., A test of homogeneity for ordered alternatives. II. Biometrika, 46, 328-35, (1959b)

Bartholomew, D. J., A test of homogeneity of means under restricted alternatives. J. R. Statisti. Soc. B, 23, 239-81, (1961a)

Bartholomew, D. J., Ordered tests in the analysis of variance. Biometrika, 48, 325-32, (1961b)

Kudô, A., A multivariate analogue of the one-sided test. Biometrika, 50, 3 and 4, 403-18, (1963)

Kudô, A. and Fujisawa, H., Some multivariate tests with restricted alternative hypotheses. Multivariate analysis, Proceeding of an international symposium held in Dayton, Ohio, June 14-19, (1965)

Kudô, A. and Fujisawa, H., A bivariate normal test with two sided alternative. Mem. Fac. Sci., Kyushu Univ., Ser. A, Math., 18, 1, 104-08, (1964)

Nüesch, P. E., On the problem of testing location in multivariate populations for restricted alternatives. Annals Math. Stat., 37, 1, 113-19, (1966)