

【調 査】

多変量正規性の検定法*

刈 屋 武 昭

1 序

多変量解析の多くの方法、特に仮説検定法では、データの確率分布として多変量正規分布と仮定している。また一変量時系列分析でも、最初は定常性の仮定から始めても、統計の話や漸近理論になると正規分布を仮定することが多い。加えて不確実性のもとの経済理論でも、しばしば効用関数との関係から多変量正規分布を仮定する。例えば、マーコ・ヴィツ型ポートフォリオ理論を効用関数との関係で議論する場合は、その例である(桐谷(1986))。他方、確率積分を用いるオプション理論の中では、多くの場合資産価格のプロセスとして(多次元)幾何的ブラウン運動に従うことを仮定する。このことは、収益率(対数価格差)が互いに独立に同じ(iid)(多変量)正規分布に従うことを仮定することを意味している。

このように、統計理論や経済理論、金融理論等では、その解析の容易性や自ら設定した枠組との無矛盾性から正規分布を仮定することがきわめて多いが、問題は対象となる変数がその仮定を満たしているかどうかである。従って、観測値を用いて実際の確率的現象が正規分布に従っているかどうかを検定することが重要となろう。本稿は、多変量正規性の検定問題を主として著者のこれまでの研究結果を中心にサーベイ的に解説する。なお、多変量正規性の検定問題は、一変量正規性の検定問題を含むことはもちろん、後にみるように問題の扱い方によっては、一変量時系列のガウス性の検定問題を含ませることができる。

最初フレーム・ワークとして、 p 次元の T 個の観測値は互いに独立に同じ分布に従う、すなわち

$$(1.1) \quad x_1, \dots, x_T \quad \text{iid}$$

とする。そして、その共通な分布が p 次元正規分布に従うという仮説

$$(1.2) \quad H: \mathcal{L}(x_i) = N_p(\mu, \Sigma)$$

μ, Σ は未知

の検定問題を考察する。ここで $\mathcal{L}(x_i)$ は、 x_i の分布観測を示し、(1.2)は、それが平均 μ 、分散共分散行列 Σ の p 次元正規分布に等しい、という仮説を示す。この帰無仮説をたてることはきわめて容易であるが、問題は対立仮説である。以下では対立仮説のたて方に関わるアプローチとして次の3つを区別する。

- (1) 包括的対立仮説(オムニバス仮説)アプローチ(ノンパラメトリック・アプローチ)
- (2) 曲分布族対立仮説アプローチ(パラメトリック・アプローチ)
- (3) インプライド対立仮説アプローチ(正規分布の特性を対立仮説とするアプローチ)

第一のアプローチを述べる。

対立仮説として最も包括的な場合は

$$(1.3) \quad K: \mathcal{L}(x_i) \in \mathcal{P} - \mathcal{N}$$

である。この対立仮説は

$$(1.4) \quad \mathcal{P} = \mathcal{P}^p = p \text{ 次元確率分布全体}$$

から帰無仮説である正規分布

$$(1.5) \quad \mathcal{N} = \{N(\mu, \Sigma) : \mu \in R^p, \Sigma \in \mathcal{S}(p)\}$$

を除いた集合であるから、きわめて大きく(実数より大きい濃度をもち)(ノンパラメトリック的)、非正規性の検証ができて(H_0 が棄却されても)、その乖離の方向性についての情報が提供されないため、正規性のもとで導出された方法や理論がいかなる意味で、非有効的であるかわからない。実際、これらの方法や理論は一定方向の分布の乖離に対してはロバスト性を保有していることもあろう。

そこで第2のタイプの対立仮説として、 K と対極にある対立仮説

$$(1.6) \quad K_0: \mathcal{L}(x_i) \in \mathcal{C}_0$$

を考えよう。ここで $\mathcal{C}_0$ は、正規分布からの乖離が1次元パラメータ θ で表現される、パラメトリック曲分布の1つのクラスである。以下では、正規分布に対応したロケーション・パラメータ μ 、スケール・パラメータ Σ をもつ曲分布のクラスで、密度関数として

* 本研究は文部省科学研究費一般C(04630014)の補助を受けた。記して感謝したい。

$$(1.7) \quad p_{\theta}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = |\boldsymbol{\Sigma}|^{-1/2} f_{\theta}((\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}))$$

の形をもつものを考える。ここで関数 $f_{\theta}(\cdot)$ は $[0, \infty)$ で定義され、 $\theta=0$ のとき

$$(1.8) \quad f_0(z) = (2\pi)^{-1/2} \exp(-z/2)$$

を満たす既知関数とする。すなわち(1.7)の $p_{\theta}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ は $\theta=0$ のとき p 次元正規分布 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ の密度関数になる密度関数のクラスである。関数形 f_{θ} は、知りたい乖離の方向に対応して選択する。実際、 $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ は正規分布と共通なロケーション・スケール・パラメータであるから、もう1つの1次元未知パラメータ θ が0から乖離すると分布族 $\mathcal{C}_{\theta}$ は正規分布から乖離し、確率分布全体 $\mathcal{P}$ の中で1つの曲線を描くことになる。従って、 $\mathcal{C}_{\theta}$ を曲分布族 (curved family of distributions) とよぶ。 f_{θ} の選択が1つの曲分布族の選択に対応し、正規性の検定は、選択された1つの曲分布族毎に行う。すなわち、選択された曲分布族に対して仮説検定問題

$$(1.9) \quad H_0: \theta = 0 \quad \text{vs} \quad K_0: \theta > 0$$

を検定する。ここで対立仮説として片側対立仮説をとるのは、最も知りたい乖離の方法を検証するためである。もちろん両側対立仮説とすることもできる。以下では片側検定問題のみを扱う。この検定問題の設定の仕方では、各曲分布族 $\mathcal{C}_{\theta}$ に対して局所最適な検定方式が導出される (Kariya and George (1993) 等)。

3番目のアプローチとして、正規分布から意味される仮説、すなわちインプライド対立仮説を検定し、それを棄却することで非正規性を主張することである。多変量正規分布の1つの特性に注目して、その特性をデータで検定する。多くの場合、正規分布のもとのモーメントに関する仮説である。その代表的なものは、Mardia による多変量歪度、多変量尖度についての検定である。例として多変量尖度についての検定問題を述べておく。Mardia (1970) は、4次のモーメントをもつ $\mathbf{x}_1$ の分布の多変量尖度を

$$(1.10) \quad \beta_2 = E\{(\mathbf{x}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu})\}^2$$

で定義する。この値は、 $\mathcal{L}(\mathbf{x}_1) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ のとき $\beta_2 = p(p+2)$ となるのでそれをインプライド対立仮説として仮説検定問題

$$(1.11) \quad H: \mathcal{L}(\mathbf{x}_1) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad \text{vs} \quad K_1: \beta_2 \neq p_2(p+2)$$

を考察する。もちろん H が棄却できなくても対立仮説が K_1 であるため H が棄却できないことを直接的に意味しないが、 H が棄却された場合、非正規

性を主張する。Mardia (1970) は、(1.10) に対する標本多変量尖度として、

$$(1.12) \quad b_2 = \frac{1}{T} \sum_{t=1}^T \{(\mathbf{x}_t - \bar{\mathbf{x}})' \mathbf{S}^{-1}(\mathbf{x}_t - \bar{\mathbf{x}})\}^2$$

ただし、

$$(1.13a) \quad \bar{\mathbf{x}} = \frac{1}{T} \sum_{t=1}^T \mathbf{x}_t$$

$$(1.13b) \quad \mathbf{S} = \frac{1}{T} \sum_{t=1}^T (\mathbf{x}_t - \bar{\mathbf{x}})(\mathbf{x}_t - \bar{\mathbf{x}})'$$

を提唱し、Mardia and Kanazawa (1973) では、 H のもとでは

$$(1.14) \quad W = \frac{\sqrt{T} [b_2 - p(p+2)]}{\sqrt{8 p(p+2)}} \rightarrow N(0, 1) \quad (T \rightarrow \infty)$$

を示し、 W が棄却域

$$|W| > c \quad \text{ただし} \quad P(|W| > c | H) = \alpha$$

に落ちたとき H_0 を棄却する。この正規性の検定方式は、インプライド帰無仮説のとり方に大きく依存するが、曲分布族の中での正規性の検定方式の場合と同様に、正規分布からの乖離の方法を多変量正規分布の1つの特性と対応を付けることになる。例えば、正規分布の対称性に注目して、その1つのインプライド帰無仮説として歪度が0という仮説をたてることがそれである。1次元正規分布のもとで提案されている多くの検定方式では、この3番目のアプローチをとっている。詳細については例えば、竹内 (1974) 等をみよ。

なお、本稿では扱わないが、Mardia (1970) では多変量歪度を

$$\beta_1 = E\{[(\mathbf{x}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu})]^3\}$$

と定義し、その標本統計量として

$$b_1 = \frac{1}{T^2} \sum_{t=1}^T \sum_{s=1}^T \{(\mathbf{x}_t - \bar{\mathbf{x}})' \mathbf{S}^{-1}(\mathbf{x}_s - \bar{\mathbf{x}})\}^3$$

を提案している。正規分布すなわち帰無仮説のもとでは、漸近的に

$$Tb_1 \rightarrow 6\chi^2_f, \quad f = p(p+1)(p+2)/6$$

であることを証明している。従って

2 包括的対立仮説に対する正規性の検定

1次元の正規性を検定する問題で、最も有名な検定方式は適合度検定によるカイ2乗検定である。このカイ2乗検定は、標本分布のデータの分布が直接正規分布に近いかどうかをみるため、正規性の検定方式として合理的である。また標本分布と標準正規分布の距離に基づく検定方式も検定方式の考え方としては同じものである。もちろんこれらの検定方式

のデータの使い方としては、平均 μ と分散 σ^2 の影響を除去するため基準化統計量

$$(2.1) \quad y_t = (x_t - \bar{x})/s$$

に基づいて構築される。しかしこの1次元の検定方式を多次元に拡張しようとする、いろいろな問題が起る。例えば、カイ2乗検定の場合、各軸の区間分割数を k とすると、 p 次元の場合枠目の数は k^p となり、(例えば $k=6$, $p=10$ の場合でも非常に大きな値となってしまう) データの数よりはるかに大きくなってしまい、多くの枠目はデータを全く含まないことになる。対応して距離に基づく場合でも、その測定の仕方に関して問題が起る。従って、この考え方にに基づく多変量正規性を議論している文献はない。

Henze and Zirkler(1990)では包括的対立仮説に対応する有効な多変量正規性の検定方式を提案している。彼等の検定方式は、経験的特性関数に基づくものである。まず(2.1)に対応する多次元基準化統計量として

$$(2.2) \quad y_t = S^{-1/2}(x_t - \bar{x}) \quad (t = 1, \dots, T)$$

とする。 $S^{-1/2}$ としては、 $S^{-1/2'} S^{-1/2} = S^{-1}$ となる下三角行列とする。そして y_t の経験的特性関数を

$$(2.3) \quad \phi(t) = \frac{1}{T} \sum_{i=1}^T \exp(it'y_i)$$

で定義し、検定統計量

$$(2.4a) \quad D_\beta = \int \left| \phi(t) - \exp\left(-\frac{1}{2} t't\right) \right|^2 S_\beta(t) dt$$

ただし

$$(2.4b) \quad S_\beta(t) = (2\pi\beta^2)^{-p/2} \exp(-t't/2\beta^2)$$

を作り、 D_β が大きいとき仮説を棄却する。 β は1つのウェイトであって $\beta=1$ としてよい。この検定方式はすべての対立仮説に対して一致性をもつことが示されている。その意味でこの検定は1つの重要な検定方式である。しかし、(2.4)はすべてのモーメント構造に関する情報を含んでいるため、きわめて高い次数のモーメントの違いに対して反応する可能性もあり、与えられたデータでは十分識別できない部分を求めている可能性もある。また実際に必要な知識として、低次のモーメント構造から順にチェックしたい部分もあろう。

Kariya, Tsay, Terui and Li(1992)(以下KTTL(1992)と略)では、この点を意識して Q 次までの多変量正規性のモーメント構造を直接的に検定する方法をエルミート多項式を用いて提案している。以下この方法を詳しく述べる。まず、帰無仮説として

$$(2.5) \quad H^*: \mathcal{L}(x_1) \text{ の分布のモーメントは、正規分布 } N_p(\mu, \Sigma) \text{ のモーメントと同じである}$$

を定義する。これは H のインプライド帰無仮説であるが、 $p=1$ のときは H と同等である。多次元の場合、正規分布と同じモーメント構造をもつ分布が正規分布であるが証明されていないが、データの識別能力としては十分な仮説である。 H^* に対する包括的対立仮説は

$$(2.6) \quad K^*: \text{not } H^*$$

である。この仮説検定問題を考察するため、次のエルミート多項式に関する事実を用いる。まず p 次元確率ベクトル

$$x = (x_1, \dots, x_p)'$$

に対して基準化変数

$$(2.7) \quad z_i = (x_i - \mu_i)/\sigma_{ii}^{1/2} \quad (i = 1, \dots, p)$$

を作る。その共分散を

$$(2.8) \quad \phi_{ij} = \text{Cov}(z_i, z_j) = \text{Correl}(x_i, x_j)$$

とおく。また

$$(2.9) \quad w_i^{(m)} = h^{(m)}(z_i) \\ (m \text{ 次のエルミート多項式})$$

で z_i の m 次のエルミート多項式とする。例えば

$$w_i^{(1)} = z_i, \\ w_i^{(2)} = (z_i^2 - 1)/\sqrt{2}, \\ w_i^{(3)} = (z_i^3 - 3z_i)/\sqrt{6}, \\ w_i^{(4)} = (z_i^4 - 6z_i^2 + 3)/\sqrt{24}$$

である。このとき $w_i^{(n)}$ と $w_j^{(m)}$ の共分散

$$(2.10) \quad \phi_{ij}^{(n,m)} \equiv \text{Cov}(w_i^{(n)}, w_j^{(m)})$$

は、もし H^* が正しいとき

$$(2.11) \quad \phi_{ij}^{(n,m)} = \begin{cases} \phi_{ij}^{(n)} & n = m \\ 0 & n \neq m \end{cases}$$

が成立する。また逆に(2.11)が成立するとき H^* が成立する。そこで $z_1, \dots, z_p$ のモーメント構造(2.11)を検定する。まず

$$x_1, \dots, x_T \quad \text{iid}$$

は $2N$ 次のモーメントをもつ、とする。(2.7)の z_i に対応する統計量として

$$(2.12a) \quad u_{it} = \frac{x_{it} - \bar{x}_i}{\sqrt{s_{ii}}} \\ (i = 1, \dots, p; n = 1, \dots, N)$$

$$(2.12b) \quad v_{it} = h^{(n)}(u_{it})$$

とおく。このとき帰無仮説および対立仮説 K^* のもとで(2.10)の共分散 $\phi_{ij}^{(nm)}$ は

$$(2.13a) \quad r_{ij}^{(nm)} = \frac{1}{T} \sum_{t=1}^T v_{it}^{(n)} v_{jt}^{(m)}$$

で強一致的に推定される。他方、帰無仮説 H^* のもとで同じ共分散 $\phi_{ij}^{(n,m)}$ は

$$(2.13b) \quad \hat{\phi}_{ij}^{(n,m)} = \begin{cases} \hat{\phi}_{ij}^{(n)} & n = m \\ 0 & n \neq m \end{cases}$$

で強一致的に推定される。ただし $\hat{\phi}_{ij} = r_{ij}^{(1,1)}$ である。明らかに H^* のもとに限って

$$(2.14) \quad \lim_T r_{ij}^{(n,m)} = \lim_T \hat{\phi}_{ij}^{(n,m)} \\ (i, j=1, \dots, p; n, m=1, \dots, N)$$

が成立する。そこで(2.14)に対応する検定統計量を作る。

そのためまず次の記号を導入する。まず

$$(2.14a) \quad A = (A^{(nm)}) : Np \times Np \text{ 対称行列}$$

$$(2.14b) \quad A^{(nm)} = (a_{ij}^{(nm)}) : N \times N$$

とし、行列 A のベクトル化 $\text{Vec}(A)$ を次のように定義する。 $n=m$ に対して

$$(2.14c) \quad a^{(nm)} = \text{Vec}(A^{(nm)}) \\ = (a_{11}^{(nn)}, \dots, a_{1p}^{(nn)} : a_{22}^{(nn)}, \dots, \\ a_{2p}^{(nn)} : \dots : a_{pp}^{(nn)})' : \frac{p(p+1)}{2} \times 1$$

また $n < m$ に対して

$$(2.14d) \quad a^{(nm)} = \text{Vec}(A^{(nm)}) \\ = (a_{11}^{(nm)}, \dots, a_{1p}^{(nm)} : a_{22}^{(nm)}, \dots, \\ a_{2p}^{(nm)} : \dots : a_{pp}^{(nm)})' : p^2 \times 1$$

とおく。そしてさらに

$$(2.14e) \quad a_{N1} = (a^{(11)'}, a^{(22)'}, \dots, a^{(NN)'})' : f_1 \times 1$$

$$(2.14f) \quad a_{N2} = (a^{(12)'}, a^{(13)'}, \dots, a^{(1N)'}, a^{(23)'}, \dots, \\ a^{(2N)'}, \dots : a^{(N-1,N)'})' : f_2 \times 1$$

$$(2.14g) \quad f_1 = Np(p+1)/2, \quad f_2 = N(N-1)p^2/2$$

$$(2.14h) \quad f = f_1 + f_2 = Np(Np+1)/2$$

とおく。さて上の A に対応するのが

$$(2.15) \quad R = (R^{(nm)}), \quad R^{(nm)} = (r_{ij}^{(nm)})$$

$$(2.16) \quad \Phi = (\Phi^{(nm)}), \quad \Phi^{(nm)} = (\phi_{ij}^{(nm)})$$

である。これらの行列のベクトル化は

$$(2.17) \quad r_N = (r_{N1}', r_{N2}')',$$

$$(2.18) \quad \phi_N = (\phi_{N1}', \phi_{N2}')'$$

となる。このとき対立仮説 K^* のもとで

$$(2.19a) \quad \sqrt{T}(r_N - \phi_N) \rightarrow N(0, A)$$

ただし

$$(2.19b) \quad A = (A^{(nmab)})$$

$$(2.19c) \quad A^{(nmab)} = \text{asy cov}(\sqrt{T}(r^{(nm)} - \phi^{(nm)}), \\ \sqrt{T}(r^{(ab)} - \phi^{(ab)}))$$

が成立する。他方、仮説 H^* のもとでは

$$(2.20a) \quad \phi_N = (\phi_{N1}', 0')',$$

$$(2.20b) \quad \phi_{N1} = (\phi^{(11)'}, \dots, \phi^{(NN)'})'$$

$$(2.20c) \quad \phi^{(nn)} = (\phi_{11}^{(nn)}, \dots, \phi_{1p}^{(nn)} : \phi_{22}^{(nn)}, \dots,$$

$$\phi_{2p}^{(nn)} : \dots : \phi_{pp}^{(nn)})$$

$$(2.20d) \quad \phi_{ij}^{(nn)} = \phi_{ij}^{(n)}$$

であり、その一致推定量は、

$$(2.21a) \quad \hat{\phi}_N = (\hat{\phi}_{N1}', 0')',$$

$$(2.21b) \quad \hat{\phi}_{ij}^{(nn)} = \hat{\phi}_{ij}^{(n)} = (r_{ij}^{(11)})^n$$

となる。そこで

$$(2.22) \quad c_{N0} = [0, I] [r_N - \hat{\phi}_N] \\ = (c^{(22)'}, \dots, c^{(NN)'})' [c^{(12)'}, \dots, c^{(1N)'}, \\ \dots : c^{(N-1,N)'})' \equiv (c_{N1}' | c_{N2}')'$$

とおく。 H^* のもとでは

$$(2.23a) \quad \sqrt{T}c_0 \rightarrow N(0, J(\phi_N) A J(\phi_N)')$$

$$(2.23b) \quad J(\phi_N) = \partial c_{N0} / \partial r_N |_{r_N = \phi_N}$$

が成立するそれゆえ KTTL(1992)では、漸近的カイ2乗検定として検定統計量

$$(2.24a) \quad W_0 = T c_{N0}' [\hat{J} \hat{A} \hat{J}']^{-1} c_{N0} \sim \chi_f^2$$

ただし

$$(2.24b) \quad A = (A^{(nm)}), \quad A^{(nm)} = (\lambda_{ijkh}^{(nmab)})$$

$$(2.24c) \quad \hat{\lambda}_{ijkh}^{(nmab)} = \frac{1}{T} \sum (v_{it}^{(n)} v_{jt}^{(m)} - r_{ij}^{(nm)}) \\ \times (v_{kt}^{(a)} v_{ht}^{(b)} - r_{kh}^{(ab)})$$

を提案している。もちろん棄却域は

$$(2.25) \quad W_0 > c$$

である。また $N(\mu, \Sigma)$ の偶数次のモーメント構造を検定する検定統計量

$$(2.26) \quad W_1 = T c_{N1}' [\hat{J}_1 \hat{A}_{11} \hat{J}_1']^{-1} c_{N1}$$

と $N(\mu, \Sigma)$ の奇数次のモーメント構造を検定する検定統計量

$$(2.27) \quad W_2 = T c_{N2}' \hat{A}_{22} c_{N2}$$

を提案している。これらの統計量のもつ意味は上の作り方から明らかである。

例1 1変量正規性の検定

1変量の場合の正規性の検定は、 $p=1$ とおけばよい。4次までの正規分布のモーメント構造を検定するため $N=2$ とおく。このとき H のもとでエルミート多項式によって変換された変数のモーメント構造は、

$$(2.27a) \quad \phi_{11} = \text{Correl}(v_{1t}^{(1)}, v_{1t}^{(1)}) \\ = \text{Correl}(x_{1t}, x_{1t}) = 1$$

$$(2.27b) \quad \text{Cov}(v_{1t}^{(1)}, v_{1t}^{(2)}) = 0$$

$$(2.27c) \quad \text{Cov}(v_{1t}^{(2)}, v_{1t}^{(2)}) = \phi_{11}^2 = 1$$

となる。(2.27b)は(2.7)の基準化変数に対して $E(z_{1t}) = E(z_{1t}^3) = 0$ と同等である。また(2.27c)は $E(z_{1t}^4) = 3$ と同等である。従って、仮説 H^* は

$$(2.28) \quad \text{歪度} = 0, \text{尖度} = 3$$

の結合仮説に対応する。(2.17)の r_N は

$$(2.29) \quad r_N = (r_{11}^{(1)} r_{11}^{(22)} r_{11}^{(12)})'$$

(2.21)の c_{N0} , (2.23a)の $\hat{J}$ は

$$(2.30) \quad c_{N0} = (r_{11}^{(22)} - (r_{11}^{(1)})^2, r_{11}^{(12)})'$$

$$(2.31) \quad \hat{J} = \begin{pmatrix} -2r_{11}^{(11)} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

となる。従って検定統計量 W_0 は

$$(2.32) \quad W_0 = T c_{N0}' [\hat{J} \hat{\Lambda} \hat{J}]^{-1} c_{N0} \sim \chi^2_2$$

となる。尖度のみの検定は

$$W_1 = T (r_{11}^{(22)} - (r_{11}^{(1)})^2) / \hat{J}_1 \hat{\Lambda}_{11} \hat{J}_1' \sim \chi^2_1$$

$$\hat{J}_1 = (-2r_{11}^{(11)}, 1)$$

となる。また対称性としての歪度に関する検定は

$$W_2 = T (r_{11}^{(12)})^2 / \hat{\lambda}_{1111}^{(1212)} \sim \chi^2_1$$

となる。

3 曲分布族対立仮説

本節では Kariya and George (1993a, b), Kuwana and Kariya (1991) に従って、第1節で述べた曲分布族とのクラスの中で正規性の検定問題を考察する。多変量正規分布をとり得る曲分布族 $\mathcal{C}_\theta$ として、密度関数

$$(3.1) \quad p_\theta(x|\mu, \Sigma) = |\Sigma|^{-1/2} f_\theta((x-\mu)'\Sigma^{-1}(x-\mu))$$

ただし

$$(3.2) \quad 1) \quad f_0(z) = (2\pi)^{-p/2} \exp(-z/2)$$

$$2) \quad \partial f_\theta(z)/\partial \theta \text{ は } (z, \theta) \in [0, \infty) \times [0, a] \text{ 上連続}$$

をもつ分布族を考察する。従って、正規性の検定問題は1次元パラメータ θ を用いて

$$(3.3) \quad H: \theta = 0 \quad \text{vs} \quad K: \theta > 0$$

と表現される。(3.1)の曲分布族の特別な場合として指数 d -分布族を次の形の密度関数をもつ分布族として定義する。

$$(3.4) \quad p_\theta(x|\mu, \Sigma) = c(\theta) |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} d_\theta((x-\mu)'\Sigma^{-1}(x-\mu)) \right\}$$

$d_\theta(z)$ は z の多項式

この分布族の特別な場合として、Kuwana and Kariya (1991) が扱った指数ベキ分布族がある。そこでは

$$(3.5) \quad d_\theta(z) = z^{1-\theta}$$

である。 $\theta=0$ の場合もちろん正規分布であるが、 $\theta > 0$ のとき正規分布より裾の厚い分布となる。 $\theta = 1/2$ の分布を多変量両側指数分布という。(3.1)の定式化で、正規分布と同様に (μ, Σ) をロケーション

ン・スケール・パラメータとしてもつ曲分布族を考察するのは、この分布族のクラスの中では、分布形の検定で不要なパラメータ (nuisance parameter) (μ, Σ) を不変性原理を用いて自然に消去できるからである。すなわち、検定問題(3.3)を検定する場合、群不変性原理によって統計量

$$(3.6a) \quad W = W(X)$$

$$= \frac{1}{T} (X - 1\bar{x}') S^{-1} (X - 1\bar{x}')'$$

$$= Y Y': T \times T$$

ただし

$$(3.6b) \quad Y = \frac{1}{\sqrt{T}} (X - 1\bar{x}') S^{-1/2}$$

$$= [y_1, \dots, y_T]': T \times p$$

に基づいて検定することができることになる。この統計量の分布は未知パラメータ (μ, Σ) に依存しないから、一般性を失うことなく

$$(3.7) \quad \mu = 0, \quad \Sigma = I$$

と仮定できる。 θ のもとの W の確率分布を P_θ^W とすると、 W に基づく検定関数は

$$(3.8) \quad \pi(\phi, \theta) = \int \phi(W) dP_\theta^W$$

となる。このとき Kariya and George (1993) は次の結果を証明している。

定理 3.1

$$(3.9a) \quad \frac{\partial}{\partial \theta} \pi(\phi, \theta)|_{\theta=0} = T E_0[\phi(W) K(W)]$$

ただし $E_0(\cdot)$ は、 $X \sim N(0, I_T \otimes I_p)$ のもとでの期待値、

$$(3.9b) \quad K(W) = \frac{1}{T} \sum_{i=1}^T E_{b,a}[\Psi(b\|y_i\|^2 + 2\sqrt{b} a_1 \|y_i\| + a' a)]$$

$$(3.9c) \quad \Psi(z) = \frac{\partial}{\partial \theta} \log f_\theta(z)|_{\theta=0} = f'_0(z)/f_0(z)$$

$$(3.9d) \quad b \sim \chi^2_r, \quad a \sim N_p(0, I)$$

(3.10) $K(W) > c$ なる棄却域をもつ検定は局所最良不変検定である。

系 3.1

指数 d -分布族の場合、 $K(W)$ は

$$(3.11) \quad K(W) = -\frac{1}{2T} \sum_{i=1}^T E_{b,a}[d'_\theta(b\|y_i\|^2 + 2\sqrt{b} a_1 \|y_i\| + a' a)] + C$$

となる。

また $K(W)$ の分布を考えるため、Kuwana and

Kariya(1991)と同じ方法で次の漸近的結果を示している。まず $K(W)$ を含む統計量として

$$(3.12) \quad U(W) = \frac{1}{T} \sum_{i=1}^T E_{\theta}[H(b\|y_i\|^2)]$$

を定義し、関数 $H(\cdot)$ に対して

$$(3.12a) \quad H \text{ は 2 回連続微分可能}$$

$$(3.12b) \quad E_{\theta}[H(\|x_1\|^2)] < \infty,$$

$$(3.12c) \quad E_{\theta}[\|x_1\|^2 H(\|x_1\|^2)] < \infty$$

$$(3.12d) \quad \xi \equiv E_{\theta}[\|x_1\|^2 H'(\|x_1\|^2)] < \infty$$

$$(3.12e) \quad G(u) = \sum_{k=k_1}^{N_2} a_k u^k \text{ なる関数が存在し}$$

て

$$|H'(u)| \leq G(u),$$

$$E_{\theta}[\|x_1\|^4 G(\|x_1\|^2)] < \infty$$

を仮定する。このとき次の定理が成立する。

定理 3.2 (3.12a)~(3.12d) のもとで、 $T \rightarrow \infty$ のとき

$$(3.13a) \quad \sqrt{T}[U(W) - \mu(\theta)] \rightarrow N(0, \nu(\theta))$$

が成立する。ただし $\rightarrow$ は分布収束を示し、

$$(3.13b) \quad \mu(\theta) = E_{\theta}[H(\|x_1\|^2)]$$

$$(3.13c) \quad \nu(\theta) = \text{Var}_{\theta}\left\{H(\|x_1\|^2) - \left(\frac{1}{p}\|x_1\|^2 - 1\right)\xi\right\}$$

系 3.2 $T \rightarrow \infty$ のとき局所最良検定統計量 $K(W)$ の帰無分布は

$$(3.14) \quad \sqrt{T}K(W) \rightarrow N(0, \nu(0))$$

となる。 $\nu(0)$ は関数形

$$(3.15a) \quad H(b\|y_i\|^2) = E_a[\Psi(b\|y_i\|^2 + 2\sqrt{b}a_1\|y_i\| + a'a)] \\ = \exp\left\{-\frac{b}{2}\|y_i\|^2\right\} \sum_{k=1}^{\infty} \frac{\lambda_k(b\|y_i\|^2)^k}{(2k)!}$$

$$(3.15b) \quad \lambda_k = E_a[\Psi(a'a) a_1^{2k}]$$

に依存する。

ϕ_{LBI} で密度関数形(3.1)をもつ各曲分布族 $\mathcal{C}_{\theta}$ に対する局所最良不変検定を表すと、(3.9a)より正規分布($\theta=0$)の近傍で検出力を最大にするのが ϕ_{LBI} であるから

$$(3.16) \quad \frac{1}{\sqrt{T}} \frac{\partial}{\partial \theta} \pi(\phi_{LBI}, \theta)|_{\theta=0}$$

は、曲分布族 $\mathcal{C}_{\theta}$ が正規分布から乖離しよう($\theta > 0$)とする局所的な大きさを表す。従って(3.16)の値は正規分布からの局所的乖離の測度として利用できる。もちろんそれは T に依存するので、Kariya

and George(1993a)はその極限値を局所的乖離の測度として提案している。その極限値は次の通りである。

定理 3.3

$$(3.17) \quad \delta_0 = \lim_{T \rightarrow \infty} \frac{1}{T} \frac{\partial}{\partial \theta} \pi(\phi_{LBI}, \theta)|_{\theta=0} \\ = \nu(0)^{1/2} (2\pi)^{-1/2} \exp(-c^2/2)$$

ただし、 c は標準正規分布関数 $\Phi(c)$ を用いて

$$1 - \Phi(c) = \alpha$$

から決まる棄却点で、有意水準 α が共通であるかぎり、同じである。

従って、局所的乖離の測度 δ_0 は局所最良不変検定の統計量 $K(W)$ の漸近的標準偏差 $\nu(0)^{1/2}$ に比例的である。すなわち $K(W)$ の漸近的標準偏差は正規分布からの局所的乖離を測る測度である。この値について、密度関数(3.1)で示される分布族を順序づけることができる。

例 3.1 (3.5)の $d_{\theta}(z)$ をもつ指数ベキ分布族の場合 Kuwana and Kariya(1991)の結果から、局所最良不変検定統計量は、かなり複雑で

$$(3.18a) \quad K(W) = -1 - \frac{p \log T}{2} + \frac{1}{T^2} \sum_{i=1}^T \sum_{k=0}^{\infty} \frac{k}{2r+p} \times \frac{\Gamma(k + (T-1)/2)}{\Gamma((T-1)/2) \Gamma(k+1)} \\ (1 - z_i' z_i)^{(T-1)/2} (z_i' z_i)^k$$

ただし

$$(3.18b) \quad z_i = \sqrt{T} y_i (1 + T y_i' y_i)^{-1/2}$$

で与えられる。この $K(W)$ の漸近分布は、設定した条件(3.12a, b, c)をチェックすることが困難だけでなく、漸近的平均、分数も計算が難しく、未知のままである(例 3.3)をみよ。

μ が既知の場合の局所最良不変検定

μ が既知の場合の局所最良不変検定は未知の場合のそれと大きく異なる。例 3.1 でみたように μ が未知な場合の局所最良不変検定は、複雑なことが多くその漸近分布の導出も困難なことが多い。注意すべき点は、 μ が既知な場合の局所最良不変検定統計量に、 μ の代わりに $\bar{x}$ を代入しても、 μ が未知の場合の局所最良不変検定統計量にならないことである。例外は次節でみる Mardia の尖度検定である。 μ が既知の場合を扱う積極的理由は、多変量線形回帰モデルの誤差項の多変量正規性の検定に応用可能であ

るからである。具体的には

$$(3.19) \quad U = Z\gamma + E$$

を多変量線形モデルとする。ただし $U = (u_1, \dots, u_n)'$: $T \times p$, $Z = (z_1, \dots, z_T)'$: $T \times k$ は確定行列, γ : $k \times p$ は未知係数行列, $E = (e_1, \dots, e_T)'$: $T \times p$ は誤差行列である。ここでの問題は e_t は iid か $E(e_t) = 0$ を満たす e_t の p 次元正規性を検定することである。しかし、このモデルでは、 $\gamma'z_t$ が各 $u_t = \gamma'z_t + e_t$ においてロケーション・パラメータであるが、モデル(3.19)のもとで不要なパラメータ γ を消去できる群不変構造はない。また、 $(\hat{\gamma}, \hat{E}'\hat{E})$ は、 $\theta=0$ の場合を除いてモデル(3.19)に対して十分統計量でもない。ただし

$$\hat{\gamma} = (Z'Z)^{-1}Z'U, \quad \hat{E} = U - Z\hat{\gamma}$$

である。従って、局所最良不変検定を知ることができない。この場合、ここで提案する方法は、残局 $\hat{e}_t = u_t - \hat{\gamma}'z_t$ を平均 0 の iid データとみなして、 e_t に基づいた局所最良不変検定の中に代入することである。

以下では指数 d 分布族モデル(3.4)において一般性を失うことなく $\mu=0$ とし、 $d_\theta(z)$ は(3.5)で与えられるものとする。このとき定理 3.1 に対応する結果として次の結果をえる。

定理 3.4 $\mu=0$ のとき、仮説検定問題(3.3)に対して

$$(3.20) \quad K(W^0) = \frac{1}{T} \sum_{i=1}^T E_b[\Psi(b\|y_i\|^2)]$$

は局所最良不変検定統計量である。ただし

$$(3.21) \quad Y^0 = [y_1^0, \dots, y_T^0]' = X(X'X)^{-1/2}/\sqrt{T}, \\ W^0 = Y^0 Y^{0'} \\ \Psi(z) = f_0'(z)/f_0(z)$$

である。また 1 つの不変検定 ϕ については

$$(3.22) \quad \frac{\partial}{\partial \theta} \pi(\phi, \theta)|_{\theta=0} = TE_0[\phi(W^0)K(W^0)]$$

が成立する。さらに(3.12)の条件のもとで漸近的に正規性

$$\sqrt{T}[K(W^0) - \mu(\theta)] \rightarrow N(0, \nu(\theta))$$

が成立する。

例 3.2 指数ベキ級数モデル(3.5)の場合、局所最良不変検定統計量は

$$(3.23) \quad K(W^0) = \frac{1}{2T} \sum_{i=1}^T \|y_i\|^2 \log \|y_i\|^2 + c_0 \\ = \frac{1}{2T} \sum_{i=1}^T x_i' S^{0-1} x_i \log x_i' S^{0-1} x_i + c$$

ただし、

$$c_0 = (p/2)[\psi(T/2) - \psi(p/2)] - 1 + (p/T),$$

$$\psi(a) = \int_0^\infty z^{a-1} (\log z) e^{-a} dz / \Gamma(a)$$

で与えられる。 μ が未知な場合と比べてきわめて単純化されている。漸近的分散は

$$\begin{aligned} \nu(0) &= p(p+2)\psi'\left(\frac{p}{2}+1\right) - 2p \\ &= p(p+2) \left[\frac{\pi^2}{2} - 4 \sum_{r=1}^{(p+1)/2} \frac{1}{(2r-1)^2} \right] - 2p \quad (p \text{ 奇数}) \\ &\quad p(p+2) \left[\frac{\pi^2}{6} - \sum_{r=1}^{p/2} \frac{1}{r^2} \right] - 2p \quad (p \text{ 偶数}) \end{aligned}$$

で与えられる。したがって(3.17)より $\nu(0)$ は正規分布からの乖離度に対応する。

例 3.3 $\mu=0$ のときの正規混合分布

(1.7)の密度関数が正規混合分布

$$(3.24) \quad |\Sigma|^{-1/2} f_\theta(x' \Sigma^{-1} x) \\ = (1-\theta) |\Sigma|^{-1/2} f_0(x' \Sigma^{-1} x) \\ + \theta |\Sigma|^{-1/2} g(x' \Sigma^{-1} x)$$

である場合の局所最良不変検定を導出しよう。ただし $g(x'x)$ は固定した密度関数である。定理 3.3 より

$$(3.25) \quad \int_{GL(p)} \prod_{j=1}^n g(y_j' A' A y_j) |\det A|^{n-p} dA \\ < \infty$$

が成立するとき、仮説検定問題 $H: \theta=0$ vs $K: \theta > 0$ の局所最良不変検定統計量は

$$(3.26) \quad U(W) = \frac{1}{T} \sum_{j=1}^T E_b \left[\frac{g(b\|y_j\|^2)}{f_0(b\|y_j\|^2)} \right] \\ = (2\pi)^{p/2} \frac{1}{T} \sum_{j=1}^T (1+\eta_j)^{T/2} \\ E_b[g(b\eta_j)]$$

で与えられる。ここで $b \sim \chi_{r^2}$ および

$$(3.27) \quad \eta_j = \|y_j\|^2 / (1 - \|y_j\|^2)$$

である。 $U(W) + 1 = K(W)$ であることに注意せよ。

$U(W)$ の最初の表現は、 g と f_0 の尤度比の加重平均である。2 番目の表現は、 $g(\cdot)$ が $[0, \infty)$ 上解析的のとき次のように議論できる。まず g を

$$(3.28) \quad g(z) = \sum_{k=0}^{\infty} \alpha_k z^k / k! \quad \text{ただし} \\ \alpha_k = \partial^k g(z) / \partial z^k|_{z=0}$$

と展開する。このとき U はさらに

$$(3.29) \quad U(W) = \frac{1}{T} \sum_{j=1}^T (1 + \eta_j)^{T/2} \left[\sum_{k=0}^{\infty} r_k(\eta_j) \beta_k \right]$$

ただし

$$r_k(\eta_j) = \frac{\Gamma\left(\frac{T}{2} + K\right)}{k! \Gamma\left(\frac{T}{2}\right)} \eta_j^k,$$

$$\beta_k = (2\pi)^{p/2} 2^k \alpha_k$$

と表現される。この LBI 統計量の表現は次の興味ある解釈ができる。

(1) $g(z)$ は $[0, \infty)$ 上解析的であるから $g(z)$ は $z=0$ でのその微係数 α_k で決定される。正規密度の場合 $\alpha_{kN} = (2\pi)^{-p/2} (-2)^{-k}$ であり、(3.29) の $\beta_k = (-1)^k \alpha_k / \alpha_{kN}$ に対する固定した密度 g と $\beta_k = (-1)^k$ に対する正規密度の差を評価する。

(2) $w_{ji} = \|y_j\|^2 = \frac{1}{n} x_j' S^{-1} x_j (\leq 1)$ は第 j 観測値 x_j の原点からの距離を示す。もし(3.24)で θ が 1 に近く $g(z)$ の裾が正規分布に比べて厚い(薄い)ならば、 $\|y_j\|$ 、従って η_j は全体として相対的に大きな(小さな)値となる傾向をもつ。(3.24)の f_θ のパラメータ θ と g は関数形を反映する。もし θ が 1 に近いとき、 $\|y_j\|^2$ 、従って η_j 、 g の関数形をより大きく反映するし、もし θ が 0 に近ければ正規密度 f_0 の特徴を示すことになる。

(3) (3.29) の $r_k(\eta_j)$ は、係数 β_k にウェイトを与える。他方、 β_k は g の正規性からの乖離を評価する。従って η_j が大きいとき $r_k(\eta_j)$ は k が大きいとき β_k により大きなウェイトを与える。従って固定した T に対していくつかの $\|y_j\|^2$ が大きいとき高次の微係数の比率 $\beta_k = \alpha_k / \alpha_{kN}$ が正規性の乖離をより検出することになる。逆に、すべての $\|y_j\|^2$ が小さいとき、正規性の検出のため低次の微係数により大きなウェイトが与えられる。 T が大きくなると、 $(1 + \eta_j)^{T/2}$ によって大きな $\|y_j\|^2$ の効果がよりはっきりし、高次の微係数の違いの情報が、正規性の乖離をみる上で重要な役割を果たす。

$U(W)$ の漸近分布は $H(z) = g(z) / f_0(z)$ とおくと

$$(3.30) \quad \sqrt{n} [T(W) - 1] \rightarrow N(0, v(0))$$

for $\theta = 0$

ただし

$$v(\theta) = \text{Var}_\theta \left[\frac{g(\|x_1\|^2)}{f_0(\|x_1\|^2)} - \frac{1}{p} \|x_1\|^2 \delta(\theta) \right]$$

$$\delta(\theta) = E_\theta \left[\|x_1\|^2 \frac{g'(\|x_1\|^2)}{f_0(\|x_1\|^2)} + g(\|x_1\|^2) \right]$$

となる。例として

$$(3.31) \quad |\Sigma|^{-1/2} g(x' \Sigma^{-1} x)$$

$$= |\Sigma|^{-1/2} \frac{a \Gamma(p/2)}{(2^{1/a} \pi)^{p/2} \Gamma(p/2a)}$$

$$\times \exp \left\{ -\frac{1}{2} [(x' \Sigma^{-1} x)]^a \right\}$$

の場合をみてみよう。ここで $a > 0$ は固定されている。このとき LBI 検定統計量は(3.26)より

$$(3.32) \quad U(W) = \frac{1}{T} \sum_{j=1}^T E_b H(b \|y_j\|^2),$$

ただし $b \sim \chi_r^2$

$$H(z) = \frac{2^{p(a-1)/2} a \Gamma(p/2)}{\Gamma(p/2a)}$$

$$\times \exp \{ (z - z^a) / 2 \}$$

となる。(3.29)の表現は、

$$(3.33) \quad U(W) = \frac{a \Gamma(p/2)}{(2^{1/a} \pi)^{p/2} \Gamma(p/2a) \Gamma(T/2)}$$

$$\times \frac{1}{T} \sum_{j=1}^T (1 + \eta_j)^{T/2}$$

$$\times \left[\sum_k \frac{(-1)^k 2^{(a-1)k}}{k!} \Gamma\left(\frac{T}{2} + ak\right) \eta_j^{ak} \right]$$

$U(W)$ の漸近的結果(3.29)が成立するが、 $a < 1$ のときは定理 3.2 の仮定が満たされないので漸近的正規性は証明されていない。

4 Mardia の尖度検定

1 節で述べたように、Mardia(1970)は正規分布の特性値として多変量尖度を定義し、その標本統計量に基づいて正規性の検定法を提案した。Kariya and George(1993)では、この Mardia の尖度検定が、実は

$$(4.1) \quad \frac{\partial}{\partial \theta} d_\theta(z) |_{\theta=0} = -cz^2 \quad (c > 0)$$

を満たす指数 d 分布族に対して局所最良不変検定となることを示した。実際、(3.11)より(4.1)のもとでは

$$(4.2) \quad K(W) = \frac{c}{T} \sum_{i=1}^T E_{b,a} [b \|y_i\|^2 + 2\sqrt{b} a_1 \|y_i\| + a' a^2] + C$$

$$= \frac{c}{T} \sum T(T+2) \|y_i\|^4$$

$$- T(T+2)p(p+2) + C$$

となり、

$$(4.3) \quad K_M \equiv \frac{1}{T} \sum_{i=1}^T \|y_i\|^2 = \frac{1}{T} \sum_{i=1}^T [(x_i - \bar{x})'$$

$$\times S^{-1} (x_i - \bar{x})]^2$$

であるから、これは Mardia の多変量尖度統計量である。(4.1)を満たす $d_\theta(z)$ として、例えば

$$(4.4) \quad d_\theta(z) = z(1 - \theta z)^2$$

をとることができる。3節で述べた漸近的結果を用いるか、あるいは Mardia and Kanazawa(1973)により

$$\sqrt{T}[K_M - p(p+2)] \rightarrow N(0, 8p(p+2))$$

が成立する。従って正規分布との局所乖離度は

$$\begin{aligned} \delta_M &= \lim_{T \rightarrow \infty} \frac{\partial}{\partial \theta} \pi(\phi_M, \theta) |_{\theta=0} \\ &= (8p(p+2))^{1/2} (2\pi)^{-1/2} \exp(-c^2/2) \end{aligned}$$

となる。 $8p(p+2)^{1/2}$ は、指数べき分布族の場合と比べて小さいので、Mardiaの尖度統計量を局所最良不変検定にする指数 d 分布族は局所的に近い。

5 1 変量時系列でのガウス性の検定

2節で述べた包括的対立仮説に対する多変量正規性の検定法は、1変量時系列のガウス性の検定に応用できる。実際、KTTL(1992)では、次の議論を行っている。まず $\{y_t\}$ を平均、自己共分散

$$(5.1) \quad E(y_t) = \mu, \quad \text{Cov}(y_t, y_{t-k}) = \gamma_k$$

をもつ定常時系列とする。後の漸近的正規性のために

$$(5.2) \quad \sum_{k=-\infty}^{\infty} |k| |\gamma_k| < \infty$$

を仮定する。多変量正規分布の次元 p に対応する変数として

$$(5.3) \quad x_{it} = y_{t-i+1} \quad (i = 1, \dots, p)$$

と定義し、その基準化変量を

$$(5.4) \quad z_{it} = \frac{x_{it} - \mu}{\sqrt{\gamma_0}}$$

とおく。このとき z_{it} と z_{jt} の共分散は

$$(5.5) \quad \text{Cov}(z_{it}, z_{jt}) = \gamma_{i-j}/\gamma_0 \equiv \rho_{i-j} \equiv \phi_{ij}$$

となる。またエルミート多項式による変数変換

$$(5.6) \quad w_{it}^{(n)} = h^{(n)}(z_{it})$$

とし、その共分散を

$$(5.7) \quad \phi_{ij}^{(nm)} = \text{Cov}(w_{it}^{(n)}, w_{jt}^{(m)})$$

とおく。このときもし $\{y_t\}$ がガウス過程ならば

$$\begin{aligned} (5.8) \quad \phi_{ij}^{(nm)} &= \text{Cov}(w_{it}^{(n)}, w_{jt}^{(m)}) \\ &= \phi_{ij}^{(n)} \quad n = m \\ &= 0 \quad n \neq m \end{aligned}$$

が成立する。そこで与えられたデータ $\{y_t: t=1, \dots, T\}$ に対して

$$(5.9) \quad u_{it} = \frac{(x_{it} - \bar{x})}{\sqrt{s_{ii}}}, \quad \bar{x} = \bar{y} = \frac{1}{T} \sum y_t$$

$$(5.10) \quad s_{ij} = \hat{\gamma}_{i-j} = \frac{1}{T} \sum_{t=i-j+1}^T (y_t - \bar{y}) \times (y_{t-i+j} - \bar{y})$$

とおき、 u_{it} のエルミート多項式による変換を

$$(5.11) \quad v_{it}^{(n)} = h^{(n)}(u_{it}) \quad (i = 1, \dots, p; n = 1, \dots, N)$$

とおく。このとき(5.7)の $\phi_{ij}^{(nm)}$ は、定常性と仮定する。このとき(5.2)のもとで

$$(5.12) \quad r_{ij}^{(nm)} = \frac{1}{T} \sum_{t=\max(i,j)}^T v_{it}^{(n)} v_{jt}^{(m)}$$

によって一貫的に推定される。他方、ガウス性が成立すると

$$(5.13) \quad \hat{\phi}_{ij}^{(nm)} = \begin{cases} \hat{\phi}_{i-j}^{(n)} & n = m \\ 0 & n \neq m \end{cases}$$

によって一貫的に推定される。そこで以下2節と同様な議論をし、Keenan(1983)による漸近的正規性

$$(5.14) \quad \sqrt{T}(r_N - \phi_N) \rightarrow N(0, A)$$

を用いると、検定統計量 W_0, W_1, W_2 を得る。

帰無仮説のもとでの W_0 の漸近的分布が重要となるが、KTTL(1992)では、シミュレーションを行っている。まず1変量時系列として

$$x_t = 0.5x_{t-1} + \varepsilon_t, \quad \varepsilon_0 = 0$$

を設定し、 ε_t の分布として 1)標準正規分布 $N(0, 1)$, 2)指数分布 $\text{Exp}(1)$, 3)コーシー分布 $C(0, 1)$, 4)ガンマ分布 $G(2)$, 5)一様分布 $U(0, 1)$ をとり、それぞれ乱数を発生させる。 $p=2, N=2$ の場合に、 W_0 の分布を帰無分布で近似する妥当性・非ガウスの場合の検出力をみている。

例 Terui and Kariya(1994)では、株価収益率の1変量時系列としてのガウス性を検定している。データは1988年1月5日から1989年10月31日の日次データである(標本数479)。 $p=2, N=2$ の場合の東証株価指数(TOPIX)および日経平均の結果は次のとおりである。

	W_0	p 値
TOPIX	14.83	0.0383
NIKKEI	16.32	0.0224

いずれも5%水準でガウス性は棄却される。 W_1 と W_2 の統計量の方がもっとはっきりした結果を生むであろう。

(一橋大学経済研究所)

参考文献

- Andersson, S. A.(1978), Invariant measures, Technical Report No.129, Stanford University, Department of Statistics.
Baringhaus, L. and Henze, N.(1988), A consistent test

- for multivariate normality based on the empirical characteristic function, *Metrika* **35**, 339-348.
- Baringhaus, L., Danschke, R. and Henze, N.(1989), Recent and classical tests for normality—a comparative study, *Comm. Statist. B* **18**, 363-379.
- Billingsley, P.(1986), *Probability and Measure*, 2nd Edition. John Wiley, New York.
- Cox, D. R. and D. V. Hinkley(1974), *Theoretical Statistics*, Chapman and Hall, London.
- Csorgo, S.(1986), Testing for normality in arbitrary dimension, *Ann. Statist.* **14**, 708-723.
- Csorgo, S.(1989), Consistency of some tests for multivariate normality, *Metrika* **36**, 107-116.
- Davis, A. W.(1979), On the effects of moderate multivariate nonnormality on Wilks's likelihood ratio criterion, *Biometrika* **67**, 419-427.
- Epps, T. W. and Pulley, L. B.(1983), A test for normality based on the empirical function, *Biometrika* **70**, 723-726.
- Fang, K. T. and Y. T. Zhang(1990), *Generalized Multivariate Analysis*, Science Press Beijing, Hong Kong.
- Gastwirth, J. L. and Owens, M. E. B.(1977), On classical tests of normality, *Biometrika* **64**, 135-139.
- Henze, N.(1990), The asymptotic distribution of Mardia's measure of multivariate kurtosis, Technical Report, Institut für Mathematische Stochastik, Hannover University, West Germany.
- Henze, N. and B. Zirkler(1990), A class of invariant consistent test for multivariate normality, Technical Report, Institut für Mathematische Stochastik, Hannover University, West Germany.
- Kariya, T. and B. K. Sinha(1988), *The Robustness of Statistical Tests*, Academic Press.
- Kariya T. and E. George(1992), Locally best invariant tests for multivariate normality in curved families with μ known. To appear from the *Proceedings of the International Symposium on Multivariate Analysis and its Applications*.
- Kariya T. and E. George(1993), Locally best invariant test for multivariate normality in curved families and Mardia's test. To appear from *Sankhya*.
- Kariya, T. and N. Terui(1994), Testing the Gaussianity of returns of Japanese stocks, Mimeo.
- Kariya, T., R. Tsay, N. Terui and H. Li(1992), Tests for multinormality with application to time series, Discussion paper 264, Hitotsubashi University.
- Kelker, D.(1970), Distribution theory of spherical distributions and a location-scale parameter generalization, *Sankhya* **A43**, 419-430.
- Kozioł, J. A.(1982), A class of invariant procedures for assessing multivariate normality, *Biometrika* **69**, 423-427.
- Kozioł, J. A.(1986), A note on the asymptotic distribution of Mardia's measure of multivariate kurtosis, *Commun. Statist. Theor. Meth* **15**, 1507-1513.
- Kuwana, Y. and T. Kariya(1991), LBI tests for multivariate normality in exponential power distributions, *Jour. Multivariate Analysis*, **39**, 117-134.
- Lehmann, E. L.(1959), *Testing Statistical Hypotheses*, John Wiley, New York.
- Machado, S. G.(1983), Two statistics for testing for multivariate normality, *Biometrika* **70**, 713-718.
- Malkovich, J. F. and Afifi, A. A.(1973), On tests for multivariate normality, *Jour. Amer. Statist. Assoc.* **68**, 176-179.
- Mardia, K. V.(1970), Measures of multivariate skewness and kurtosis with applications, *Biometrika* **57**, 519-530.
- Mardia, K. V.(1974), Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies, *Sankhya* **B36**, 115-128.
- Mardia, K. V.(1975), Assessment of multinormality and robustness of Hotelling's T^2 test, *Jour. Roy. Statist. Soc. C24*, 163-171.
- Mardia, K. V.(1980), Tests of univariate and multivariate normality, In *Hand book of Statistics I. Analysis of Variance*(P. R. Krishnaiah, ed.), North Holland.
- Mardia K. V. and M. Kanazawa(1983), The null distribution of multivariate Kurtosis, *Commun. Statist.* **12**(5), 569-576.
- McCulloch, J. H.(1971), Measuring the term structure of interest rates, *Jour. Business*, 19-31.
- Paulson, A. S., P. Roohan and P. Sullo(1987), Some empirical distribution function tests for multivariate normality, *Jour. Statistical Computation and Simulations* **28**, 15-30.
- Royston, J. P.(1982), An extension of Shapiro and Wilk's W test for normality to large samples, *Appl. Statist.* **31**, 115-124.
- Schwager, S. J.(1985), Multivariate skewness and kurtosis, In *Encyclopedia of Statistical Sciences* **6**(S. Kotz, N. L. Johnson and C. B. Read, Eds.), 122-125, John Wiley.
- Shapiro, S. S. and Wilk, M. B.(1965), An analysis of variance test for normality, *Biometrika* **52**, 592-611.
- Terui, N. and T. Kariya(1994), Testing Gaussianity for stock returns. mimeo.
- Wijsman, R. A.(1967), Cross-sections of orbits and their applications to densities of maximal invariants, *Fifth Berk. Symp. Math Statist. Prob 1*, University of California, Berkeley, 389-400.
- 刈屋武昭(1994)『計量経済分析法の新展開』岩波書店.
- 桐谷 維(1986)『資産選択の理論』東洋経済新報社.
- 竹内 啓(1974)「正規性の検定について」『経済学論集』39巻4号.
- 竹内 啓(1974)「多変量正規性の検定」『経済学論集』40巻2号.