統計データ解析および学会の組織・活動について

杉山　高一*

The Importance of Statistical Data Analysis and the Organization and Activities of the Japan Statistical Society

Takakazu Sugiyama*

統計データ解析が、いくつかの研究テーマを見出すことにつながることを、簡単なデータ解析を通して話したいと思う。統計の理論を研究している人にとって、統計データ解析を経験することは大切であり、統計学を応用している人との共同研究は、重要であることを２１世紀を担う若い方々に幾分でも伝えられれば幸いである。ここでのデータは中学生の成績で、主成分分析における第１主成分の分散（最大固有値）とその主成分の係数（固有ベクトル）にテーマを限定して述べる。

無作為に抽出した東京都の中学２年生１６６人を対象に、５教科、数学、英語、国語、理科、社会の試験を行い、各生徒の得点から、次のような平均と分散共分散行列が得られたとする (杉山 (1983))。

$$\bar{\boldsymbol{x}} = \begin{pmatrix} 52.4 \\ 45.4 \\ 39.0 \\ 50.1 \\ 39.4 \end{pmatrix} \begin{matrix} \text{国語} \\ \text{数学} \\ \text{英語} \\ \text{理科} \\ \text{社会} \end{matrix} \qquad S = \begin{pmatrix} 470.6 & 384.9 & 493.6 & 333.5 & 363.8 \\ 384.9 & 583.6 & 576.5 & 426.9 & 407.8 \\ 493.6 & 576.5 & 874.7 & 485.3 & 501.6 \\ 333.5 & 426.9 & 485.3 & 462.1 & 383.8 \\ 363.8 & 407.8 & 501.6 & 383.8 & 459.4 \end{pmatrix} \tag{1}$$

この分散共分散行列から主成分分析を行うと次のようになる。

固有ベクトル	1	2	3	4	5
国語 x_1	.388	.120	.808	.366	.222
数学 x_2	.455	−.315	−.410	.679	−.255
英語 x_3	.569	.731	−.309	−.203	.069
理科 x_4	.397	−.524	−.137	−.345	.656
社会 x_5	.401	−.278	.257	−.495	−.671
固有値	2363.9	168.8	146.0	104.5	67.5
寄与率	.829	.059	.051	.037	.024

第１主成分 y_1 は

$$y_1 \quad = \quad 0.388x_1 + 0.569x_2 + 0.455x_3 + 0.397x_4 + 0.401x_5 \tag{2}$$

であり、y_1 の分散 ℓ_1^*（最大固有値）は 2363.9 である。これは５教科の情報の 83% を持った主成分である。

いま、５教科の得点 $\boldsymbol{X}$ が正規分布に従うとする。その母集団からの無作為標本を $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_N$ とし、分散共分散行列を $\boldsymbol{S}$ とすると、上記の分散共分散行列 (1) は、$\boldsymbol{S}$ の一つの

*中央大学理工学部

実現値である。行列 $(N-1)\boldsymbol{S}$ は Wishart 分布に従い、最大固有値 $(N-1)\ell_1^*$（これを以下 ℓ_1 とかく）の分布を求めることができる (杉山 (1966,1967))。

$$\begin{aligned}\Pr\{\ell_1 < x\} &= \Gamma_p\left(\frac{p+1}{2}\right) \Big/ \left[|\boldsymbol{\Sigma}|^{\frac{1}{2}n} 2^{\frac{1}{2}np} \Gamma_p\left(\frac{n+p+1}{2}\right)\right] \\ &\quad \cdot (nx)^{\frac{1}{2}pn} \exp\left(-\frac{nx}{2}\mathrm{tr}\ \boldsymbol{\Sigma}^{-1}\right)\ {}_1F_1\left(\frac{p+1}{2}; \frac{n+p+1}{2}; \frac{nx}{2}\boldsymbol{\Sigma}^{-1}\right) \quad (3)\end{aligned}$$

5 教科で考えているから、上式では $p=5$ となる。例えば上式を用いて確率

$$\Pr(2200 < \ell_1 < 2500)$$

を数値計算で求めることができる。この数値計算ができるか否かは、超幾何関数 ${}_1F_1(a;c;\boldsymbol{Y})$ の数値計算が可能か否かによる。このことはゾーナル多項式が計算できるか否かと同値である。$p=2$ の場合は杉山 (1979) により、ゾーナル多項式の係数が求められている。$p \geq 3$ の場合、橋口 (2000) が係数の一般的な計算法を与えている。$p=7$ までの場合は、超幾何関数 ${}_1F_1(a;c;\boldsymbol{Y})$ を級数展開した場合の係数を求める漸化式が杉山、福田、竹田 (1999) により与えられている。現在の高速計算機を用いても、上記の確率の数値計算は $p=4$ の場合が限界で、$p=5$ の場合は不可能とまでは言わないが、たいへん厳しい。そのような状況を見通して、ℓ_1 の漸近分布の導出を杉浦 (1976) が行った。

帰無仮説 $H_0 : \lambda_1 = \lambda_0$、対立仮説 $H_1 : \lambda_1 \neq \lambda_0$ という仮説検定を統計量 ℓ_1 で検定を行うとする。帰無仮説 H_0 が真のとき、すなわち $\lambda_1 = \lambda_0$ とおいても (3) 式は定まらない。母分散共分散行列 $\boldsymbol{\Sigma}$ の最大固有値が λ_0 である行列 $\boldsymbol{\Sigma}$ は無数に存在する。最大固有値 ℓ_1 を用いての仮説検定では、有意水準 5%点をどう決めるかという問題が残る。

現実のデータ解析では、上記の確率の値は 2 桁の精度があれば十分である。データ解析で知りたい大きい方の固有値（主成分の分散）は、異なる場合がほとんどである。そのような状況のもとで、標本数が 20 あるいは 30 と少ない場合でも、2 桁の精度が保証される簡便な良い近似式が得られないかとの観点から研究が行われ、小西、杉山 (1981) により

$$\ell_1^{\frac{1}{3}} \text{は } N\left(\lambda_1^{\frac{1}{3}}, \sqrt{\frac{2}{9n}}\lambda_1^{\frac{2}{3}}\right) \text{で近似できる}$$

という結果を得ている。これは近似式としては、たいへん優れたものである。

第 1 主成分 y_1 がどのような意味の主成分であるかを推測するには、x_j の係数の大きさと符号が決め手になる。係数ベクトル

$$\boldsymbol{h}_1' = (0.388 \quad 0.569 \quad 0.455 \quad 0.379 \quad 0.401)'$$

は新たな大きさ N の標本を抽出すると、上記の $\boldsymbol{h}_1$ とは異なった値をとる。明らかに係数ベクトル $\boldsymbol{h}_1$ は統計量であり、ある確率分布に従う。$p=2$ の場合には、$\boldsymbol{h}_1 = (\cos\theta \quad \sin\theta)$ と書けるが、この分布を求めたのが杉山の修士論文であり、Ann. Math. Statist. (1965) に 2 頁の論文として掲載されている。一般の場合の分布は杉山 (1967) の中に

$$\begin{aligned}f(\boldsymbol{h}_1)d\boldsymbol{h}_1 &= \left[|\boldsymbol{\Sigma}|^{\frac{n}{2}}\Gamma(\frac{n}{2})\right]^{-1} \Gamma_p(\frac{pn}{2})\ B_{p-1}\left(\frac{p+2}{2}, \frac{n-1}{2}\right) \\ &\quad \sum_{k=0}^{\infty} \frac{(\frac{pn}{2})_k}{k!} \sum_{\kappa} \frac{\left(\frac{n-1}{2}\right)_\kappa}{\left(\frac{n+p+1}{2}\right)_\kappa} C_\kappa\left(-\frac{\boldsymbol{\Sigma}_{p-1}}{\boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1}\right)(\boldsymbol{h}_1'\boldsymbol{\Sigma}^{-1}\boldsymbol{h}_1)^{\frac{1}{2}pn} d\boldsymbol{h}_1 \quad (4)\end{aligned}$$

として与えられている。この分布は係数ベクトル $\boldsymbol{h}_1$ の安定性を知るのに有用である。$p=2$ の場合の数値計算は杉山 (1971) によってなされた。p が 3 以上の場合の数値計算は、固有値の場合よりもやっかいである。それ故、杉浦 (1976) は漸近分布を導出している。

第 1 主成分には、5 教科の有している情報を最もよく抽出していると思われる合計点

$$x_1+x_2+x_3+x_4+x_5$$

が出てくることが期待される。主成分の係数の二乗和は 1 であるから、上式は正確には

$$\frac{1}{\sqrt{5}}(x_1+x_2+x_3+x_4+x_5)$$

である。

いま対称行列 Σ を母分散共分散行列として、$\boldsymbol{\eta}_1$ を $\boldsymbol{\Sigma}$ の最大固有値 λ_1 に対応する固有ベクトルとする。ただし、固有値は単根とする。このとき、第 1 主成分が合計点であるという帰無仮説は

$$H_0:\boldsymbol{\eta}_1=\frac{1}{\sqrt{5}}(1,\ 1,\ 1,\ 1,\ 1)'$$

とかける。第 1 主成分に関する仮説検定を例にして述べたが、一般に、α 番目に大きい固有値 λ_α に対応する固有ベクトル $\boldsymbol{\eta}_\alpha$ に関する帰無仮説 H_0 と対立仮説 H_1 は

$$H_0:\boldsymbol{\eta}_\alpha=\boldsymbol{\eta}_0\ ,\quad H_1:\boldsymbol{\eta}_\alpha\neq\boldsymbol{\eta}_0$$

とかける。ただし、$\boldsymbol{\eta}_0$ は既知ベクトルとする。固有ベクトルの検定統計量としては、T.W.Anderson による統計量

$$\Lambda_1 \quad = \quad n\left(\ell_\alpha\boldsymbol{\eta}_0'\boldsymbol{S}^{-1}\boldsymbol{\eta}_0+\frac{1}{\ell_\alpha}\boldsymbol{\eta}_0'\boldsymbol{S}\boldsymbol{\eta}_0-2\right) \tag{5}$$

塚田、杉山 (1997) による統計量

$$\Lambda_2 \quad = \quad n\left(\boldsymbol{\eta}_0'\boldsymbol{S}^2\boldsymbol{\eta}_0-2\ell_\alpha\boldsymbol{\eta}_0'\boldsymbol{S}\boldsymbol{\eta}_0+\ell_\alpha^2\right) \tag{6}$$

$$\Lambda_3 \quad = \quad n\left(\frac{1}{\ell_\alpha}\boldsymbol{\eta}_0'\boldsymbol{S}^2\boldsymbol{\eta}_0-2\boldsymbol{\eta}_0'\boldsymbol{S}\boldsymbol{\eta}_0+\ell_\alpha\right) \tag{7}$$

等が考えられる。検定統計量 Λ_2，Λ_3 は、このタイプの考えられる限りの統計量を書き出し、統計的シミュレーションによって検出力を調べて、良いと確認された 2 つである。統計的シミュレーションを、このように研究の対象を絞るのに用いることもある。Λ_1 の漸近分布は早川 (1978) による。Λ_2，Λ_3 の漸近分布は塚田、杉山の上記の論文の中で与えている。漸近分布による検出力の数値比較から多くの場合、Λ_1 より Λ_2 あるいは Λ_3 の方が優れていることを示している。

この Λ_1，Λ_2，Λ_3 による帰無仮説 (1) の検定を有意水準 5% で検定すると次のようになる。

	検定統計量の値	近似による棄却点
Λ_1	49.30	9.95
Λ_2	18.44×10^6	3.14×10^6
Λ_3	78.01×10^2	13.4×10^2

この場合、第 1 主成分が合計点であるという帰無仮説は、どの統計量でも有意水準 5% で棄却される。次に、第 1 固有ベクトルが各変数の標準偏差に比例したベクトル、つまり

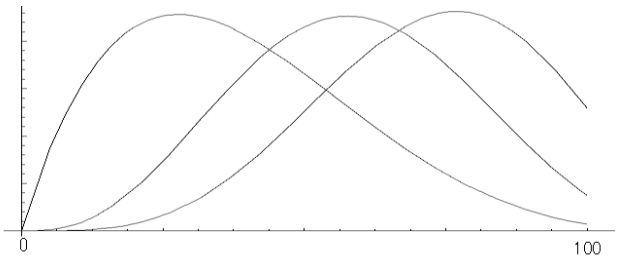
$$\boldsymbol{\eta}_1 = (.406, .452, .553, .402, .401)'$$

という帰無仮説を有意水準 5%で検定すると、

	検定統計量の値	近似による棄却点
Λ_1	1.627	9.95
Λ_2	5.032×10^5	31.4×10^5
Λ_3	2.129×10^2	13.4×10^2

となり、帰無仮説はどの統計量でも有意水準 5% で棄却されない。このような場合、第 1 主成分の x_i の係数は、変数 x_i の標準偏差に比例するのは直観的にはわかるが、理論的にはどのように説明できるか興味がある。

ここで一つの疑問が出てくる。各教科の成績の得点分布は正規分布としたが、その前提条件は妥当であろうか？ 各教科の得点分布は下記の図のようにいろいろな場合がある。



正規分布の定義域は $-\infty$ から ∞ までであり、また大きな特徴は平均点を中心に左右対象であることである。得点分布の場合、定義域は 0〜100 であり、また分布も明らかに左右対象ではない。正規分布を仮定して構成した理論を用いての確率計算や検定法は信頼できるであろうか。すなわち、同じ状況のもとで同じ調査をしたときに、安定した、あるいは再現性のある結果が得られるのであろうか。これはロバストネス（頑健性）の問題であり、たいへん重要な研究テーマである。理論的な研究は難しいが最近のワークステーションを用いた統計的シミュレーションによる研究は可能である。この場合の問題は 2 変量間の関連性の尺度をどのように入れて、乱数を発生させるかである。

多変量正規分布に従うことが不明である場合、あるいは明らかに多変量正規分布とは考えられない場合の検定法としては、ノンパラメトリック検定が考えられる。これに関連しては、パーミュテーションテストを考えて、固有値、固有ベクトルに関しての研究が進められている。主成分分析の最大固有値とそれに対応する固有ベクトルを例にして述べてきた。統計学における研究問題は、統計データ解析を行っていると、いろいろと出てくる。その意味で諸分野のデータ解析に関わることは重要である。

統計学が活用されているそれぞれの専門分野との共同研究は、新しい理論的な研究テーマを提供してくれる。統計学が活用されて、その威力を発揮している応用領域は多く、お互いが共同研究を進めることで、それぞれが得るところは大きい。

データ解析の重要性をのべてきたが、そもそもデータは統計調査、実験等によって得られたものである。統計家には、データを得るにさいして守るべき倫理がある。伊藤孝一先生は会報 No. ８７（１９９６年）で、I.S.I.（１９８５年）で採択された倫理綱領のガイドラインについて言及し、その見出しを次のように紹介している。

１．社会に対する責任
- 1.1 対立する利害関係の考慮
- 1.2 統計の範囲の拡大
- 1.3 客観性の追求

２．調査主体に対する責任
- 2.1 責任と役割の明確化
- 2.2 複数の代替的方法の公平な評価
- 2.3 結果先取りの回避
- 2.4 守秘義務

３．調査実施協力者・統計学界に対する責任
- 3.1 統計に対する信頼性の維持
- 3.2 統計的方法と調査結果の開示と再吟味
- 3.3 倫理原則の伝達

４．調査対象に対する責任
- 4.1 プライバシーの尊重
- 4.2 インフォームド・コンセントの取得
- 4.3 インフォームド・コンセントの代替案
- 4.4 調査対象の権利の保護
- 4.5 調査原票の秘密保持

伊藤孝一先生は「この倫理綱領は官庁統計のみならず、広く一般の社会現象、自然現象を対象にした統計調査、実験に適用されるものである。… 」と述べ、日本における調査環境の悪化をふまえて、重要な問題提起をされている。私達統計家は倫理綱領を常に意識して調査する側とされる側の相互信頼とより良い協力関係のもとに、質の良いデータを得るように努めることが大切である。

以後は日本統計学会の組織と活動に関連して述べたいと思います。よく御存知のように、日本統計学会会則には「本会は、統計学の研究および普及を促進し、その発達に貢献することを目的とする」とあります。この目的を遂行していくために、現時点においてはどのような組織と活動が適切かは、それぞれにいろいろなお考えがあることと思います。昨年の評議員会で報告された９８・９９年度学会活動特別委委員会（主査・藤越康祝先生）の報告書は、示唆に富んでいます。また、私が理事長のとき会長でした三浦由己先生のお考えも入りますが、考えていることを少し述べさせていただこうと思います。

日本統計学会（以後は統計学会と省略）は分科会方式を真剣に検討してよい時期に来ているように思います。分科会という言葉が適当でなければ、研究部会、研究分科会等の別の名称で結構ですが、ここでは一応「分科会」という言葉を使わせていただきます。すべての学会員がいずれかの分科会に所属する数学会のような形ではなく、いくつかの分科会があって、その分科会の活動に関心ある方が所属する形でよいと思います。わかりやすい例として、仮に統計教育委員会が分科会になったとします。分科会には小・中・高等学校の統計教育に関心がある人、大学での統計教育に関心のある人、企業での統計教育に関心のある人等これまで同様にいろいろな方々が所属することになります。分科会主催のシンポジウムや分科会としての電子ジャーナルなども考えられます。また、統計学会員以外の方で分科会の活動に関心のある方に、具体的には小・中・高等学校の教員、大学や世の中で統計教育に携わっている方々に分科会会員になっていただくことを考えてもよいのではないでしょうか。このような準会員のような制度ができたとして、ここではそれを分科会会員と書くことにいたしますが、例えば分科会会員の年会費は千円程度で、日本統計学会会報と分科会の企画等の案内を送付することが考えられます。ある年の岡山大会のときに、統計教育の研究集会に協力し参加した小・中・高の先生方が、その後も継続的に統計学会

とつながりをもっているようには思えません。統計教育には強い関心をもっているが統計学会に入会するのは躊躇してしまう人もおられることでしょう。分科会および分科会会員制度はその方々とつながりを持ちつづけて、統計教育の研究と普及にこれまで以上の広がりを持たせることができるのではないでしょうか。官庁統計でも同様のことが考えられます。この場合分科会会員には各県、各市で統計実務に携わっておられる方々になっていただくことによって、同様の広がりを期待できます。分科会会員から正会員になる方々も出てくるかもしれません。あまり厳しい条件を設定しないで、分科会がつくれれば、統計学会の活性化につながります。また、５年、１０年後に分科会というものの性格が固まっていけばよいというような長い目で捉えて試みてみることが大切かと思います。

統計学会は、性能のよいサーバを有し、８０ギガを超えるハードディスクを備えています。分科会の会員への連絡や、電子ジャーナルの発行等に活用できます。実際に統計教育委員会の研究会の案内は郵送ではなく、電子メールで行っています。絶版になった書籍の電子化を会員の有志が精力的に取り組んでいますが、それらを会員や一般の方々にお使いいただくことも可能です（作成された方々の御了承がいただければの話ですが…）。一般の方々にと書きましたが、これは会則にある統計学の普及という目的に合っていて、問題はないと考えます。

統計学会の国際化への対応は美添泰人元理事長が強く提案され、前理事会では国際関係担当理事をおき、その先生の努力で成果をあげました。いまは、竹村彰通理事が引き継ぎ、I.S.I. をはじめ、諸外国の統計学会との連帯と協力を推し進めています。教育に、研究に、さらに大学内外で重要な役割のお仕事をされている先生には、いくら有能であっても限界はあります。一つの考え方は、竹村国際関係担当理事を中心に国際交流委員会を設けて組織的に対応することです。例えばアジア地区担当、米国・中南米担当、欧州担当、…、I.S.I. 担当等の役割でそれぞれが仕事をしていくことも考えられます。この分類は一例で、別の分類で仕事を割り振ることも考えられます。国際関係担当理事を委員長とした５，６人からなる有機的な組織を構成できればと考えます。これまで統計学会は２０年ほどの間、日中統計会議、日韓統計会議を後援してきました。これらの会議を含めて、アジア諸国の統計学会等との協力関係をどのようにすすめていくかを、長期的な視野に立って検討していくことも重要かと思います。

統計学会には常設の研究部会があり、期間は２年以内という制限がついています。約４件が常時活動しています。予算申請の段階ですが、来年度は１件増やした予算額が提案されいてます。この研究部会が、分科会設置の核になりうる例も出てくるのではないかと思います。先の報告書では従来の研究部会に加えて、５年程度更新可の新しい研究部会の提案がされています。限られた予算枠の中で、庶務会計担当理事はたいへん御苦労をされていますが、研究部会新設に限っては、優れたものであれば今後とも柔軟に対応していただければと思います。

統計学会会報 No.93(1997) に美添泰人元理事長が文部省の科学研究費等を利用した研究活動の一層の活性化を提案しています。複合領域に設置された分科「統計科学」について、学会員が積極的に申請されることの必要性と重要性を述べています。それから４年経ったいまでも美添先生が危惧されて書かれた状況は変わっていないと感じています。これに対処していく研究費担当理事のようなものが考えられないでしょうか。その担当理事が委員長になり、科研費を含めた研究費全般についていろいろと対応を検討し、ときどき会報や

ホームページを通して会員にアドバイスしていただければと思います。

日本統計学会・応用統計学会（会長・小西貞則）・日本計量生物学会（会長・柳川堯）は２００２年度に統計関連学会連合大会を開催いたします。本大会連絡委員会の議長は小西貞則先生です。３年前の中央大学での日本統計学会大会には８３０名程の参加者がありました。その内の非会員の参加者は２２０名です。東京は首都でもあり、東京での開催はいつも参加者が多めです。次回は３学会合同ですから、少なくとも１０００名をこえる方々が参加されることと思います。３学会の相乗効果があって、参加者が１１００～１２００名になるかもしれません。統計学会の大会は懇親会を行ったり、皆で写真を撮ったり、以前はバスを借りきって観光までしていました。お祭り的な雰囲気もある明るい大会です。各学会は、それぞれに雰囲気をもっています。各学会の良い所を生かした実りの多い大会になればと願います。大会は長い時間軸の１点です。大会も大切ですが、学会誌のこと、会報のこと、国際対応のこと等の協力関係についても真剣に進めてくださるように、連合大会議長の小西先生にはよろしくお願いしたいと思います。

国際統計協会の２００１年大会の共通テーマは１６４あり、統計学の大きな広がりを感じます。その中から適当に抽出した一部を以下に記します。

3.Stochastic Processes	6.Multivariate Methods	9.Classification and Clustering
11.EM Algorithm	17.Time Series	20.Linear Models
23.Categorical Data	28.Robust Statistics	33.Nonparametric Statistics
36.Partial Least Squares Methods	47.Population Statistics	48.Official Statistics
55.Issues related to Metadata	64.Data Collection Systems	69.City Comparison Statistics
76.Transportation Statistics	77.Employment Statistics	80.Longitudinal and Panel Data
81.Sampling	84.Survey Methodology	91.Internet Survey
96.Biometrics	98.Statistics in Medicine	105.Statistical Modeling in Ecology
110.Bioinformatics	116.Neural Networks	118.Spatial Statistics
122.Computational Statistics	124.Simulation	127.Data Analysis
133.Statistical Education	136.Econometrics	137.Psychometrics
143.Statistics and Human Rights	149.Sports Statistics	156.Financial Statistics
157.Experimental Design	162.Reliability	163.Statistical Quality Control

統計学会も I.S.I. の在り方を参考にして、２１世紀の学会の進む方向を真剣に検討する時期に来ているように思います。

学会賞の充実、法人格の取得等、いろいろと書きたいことはありますが、別の機会にしたいと思います。最後に、会長、理事長、理事をはじめとした役員の任期は２年ですが、２年というのは何かを行うには期間が短すぎます。任期３年を検討するか、あるいは現状の任期のままの場合は I.S.I. が行っているような配慮が必要かと考えています。

参 考 文 献

1. Hashiguchi, H., Nakagawa, S. & Niki, N. (2000) Algebraic Simplification of the Laplace-Beltrami operator. *Math. Comput. Simulation*, **51**, no. 5, 489–496.
2. Hayakawa, T. (1978) The asymptotic expansion of the distribution of Anderson's statistic for testing a latent vector of a covariance matrix. *Ann. Inst. Statist. Math.*, **30**, no. 1, 51–55.
3. Konishi, S & Sugiyama, T. (1981) Improved approximations to distributions of the largest and the smallest latent roots of a Wishart matrix. *Ann. Inst. Statist. Math.*, **33**, no. 1, 27–33
4. Sugiura, N. (1976) Asymptotic expansions of the distributions of the latent roots and the latent vector of the Wishart and multivariate F matrices. *J. Multivariate Anal.*, **6**, no. 4, 500–525.
5. Sugiyama, T. (1965) On the distribution of the latent vectors for principal component analysis. *Ann. Math. Statist.*, **36**, 1875–1876.
6. Sugiyama, T. (1966) On the distribution of the largest latent root and the corresponding latent vector for principal component analysis. *Ann. Math. Statist.*, **37**, 995–1001.
7. Sugiyama, T. (1966) On the distribution of the largest latent root and the corresponding latent vector for principal component analysis. *Ann. Math. Statist.*, **37**, 995–1001.
8. Sugiyama, T. (1967) On the distribution of the largest latent root of the covariance matrix. *Ann. Math. Statist.*, **38**, 1148–1151
9. Sugiyama, T. (1971) Tables of Percentile Points of a Vector in Principal Component Analysis. *J. Japan Statist. Soc.*, **1**, no.2 63–68
10. Sugiyama, T. (1979) Coefficients of zonal polynomials of order two. *Computer Science Monographs*, **12**, Institute of Statistical Mathematics, Tokyo.
11. 杉山　高一 (1983) 多変量データ解析入門　朝倉書店
12. Sugiyama, T., Fukuda, M & Takeda, Y. (1999) Recurrence relations of coefficients of the generalized hypergeometric function in multivariate analysis. *Comm. Statist. Theory Methods*, **28**, no. 3-4, 825–837
13. Tsukada, S. & Sugiyama, T. (1997) On the distributions of likelihood ratio criterion for equality of characteristic vectors in two populations. *Bull. Fac. Sci. Engrg. Chuo Univ. Ser. I Math.*, **40** ,1–11