新統計学シリーズ=5

多変量解析の理論

南山大学教授 理学博士
伊藤孝一 著

培風館

新統計学シリーズ＝5

多変量解析の理論

南山大学教授 理学博士
伊藤孝一 著

培風館

伊藤孝一（いとう こういち）略歴

1946年　東京帝国大学第一工学部応用数学科卒
1954年　Ph. D.（米国セントルイス大学）
1966年　理学博士（九州大学）
現　在　南山大学経済学部教授
専攻分野　数理統計学および統計学，The Annals of Mathematical Statistics, Biometrika その他に論文を発表．
主要著書　現代統計学入門（培風館）

昭和44年9月20日　初版発行
昭和53年10月30日　初版第5刷発行

新統計学シリーズ 5
多変量解析の理論

著者　伊藤孝一
発行者　山本健二
発行所　株式会社 培風館
東京都千代田区九段南 4-3-12・郵便番号 102
電話(03) 262-5256(代表)・振替東京 4-44725
定価 ¥2400.

日東紙工印刷・三水舎製本

3333-0855-6955

まえがき

統計的多変量解析の理論については，すでに数種の専門書が内外において出版されているが（たとえば，[3]†，[11]，[37]，[39]，[43]，[59] 等），本書の目的とするところは，1変量の統計理論を一通り学習した読者に対して多変量の統計理論を紹介することであって，1変量と多変量の関係を重視しながら，分布論，推定論，検定論その他についてできる限り平易に，しかし理論的に厳密に解説を試みたのである．

本書を読むための予備知識としては，基礎数学として線型代数と初等解析を，また，1変量の統計理論における基本的概念と手法に関する知識を必要とする．前者については，たとえば，村上正康・掛下伸一共著：統計のための数学1および2（培風館，新統計学シリーズ3および4，1969），また，後者については，吉村功著：数理統計学（培風館，新統計学シリーズ 1，1969）の知識で十分である．ただ，本書を読むために直接必要な基礎数学の知識については，巻末の付録Aに解説し，また，普通，1変量の統計理論において取り扱われている非心カイ2乗分布と非心F分布についても付録Bにおいて簡単に説明しておいた．

各章末には簡単な演習問題を与えて読者の学習の便宜をはかり，また記述の都合で本文中で取り扱うことのできなかった事項についても，これを演習問題として取り上げるようにし，解答またはヒントはできる限り詳細に与えることにした．演習問題の数値計算において，必要な場合には電子計算機を用いて計算したが，そのプログラムは紙数の関係上割愛しなければならなかった．推定および検定のために必要な一般の統計数値表は他書に譲ったが，付録Cとして多変量解析のために特に必要な統計数値図表を二，三それぞれ原著者および出版社の許可を得て掲載しておいた．

本書では，普通，多変量解析の一分野として取り扱われている因子分析をまったく除外しているが，これについては本シリーズの他書で取り扱われることになっている．また，多変量解析の推測理論を展開するに当たって，すべて母集団の正規性を仮定し，母集団の分布の型を仮定しない，いわゆるノン・パラメトリックの統計理論には触れていないことも断わっておかなければならな

† 巻末参考文献を示す．以下同じ．

い．最近，A. T. James [35] 等によって，多変量解析における標本分布論が直交群上の不変測度の導入によって数学的に美しい形で展開されているが，これは本書の程度を超えるのでまったく触れていない．さらに，多変量解析の種種の概念や手法がいかなる分野において，いかなる条件の下において適用されるかという具体的な応用問題も取り扱っていないが，これについては本シリーズの他書において計画されているのでこれに期待したい．

著者はこれまで，数理統計学および統計学の研究において内外の諸先生から種々のご指導を賜わったのであるが，ここに記して深く感謝申し上げたい．本書の作成に際して，貴重なご助言を頂いた本シリーズの編集委員の諸先生，特に，校正刷を読んで有益なご示唆を下さった慶応大学教授　浦昭二氏には改めて厚く謝意を表したい．終りに，本書の出版に当たって色々とお世話になった培風館の森平勇三氏，小関清氏，牧野末喜氏その他の皆様に心からお礼を申し上げる次第である．

1969 年 7 月

著　者

目　次

1 序論

1·1 多変量解析とは····1
1·2 記号····4

2 多変量分布

2·1 多変量分布····6
2·2 回帰と相関····12
2·3 多変量正規分布····18
演習問題····26

3 多変量正規母集団からの無作為標本

3·1 1個の母集団の場合····28
3·2 k個の母集団の場合····31
3·3 直線回帰の場合····34
演習問題····37

4 多変量解析における点推定

4·1 同時点推定····39
4·2 正規母集団の母数の同時最尤推定量····42
4·3 正規母集団の母数の同時不偏推定量と同時充足推定量····44
演習問題····45

5 標本分布（I）

5·1 Wishart 分布····47
5·2 $C(S)$ の同時分布····53
5·3 $C(S_1S_2^{-1})$ の同時分布····55
5·4 $C(S_{11}^{-1}S_{12}S_{22}^{-1}S_{12}')$ の同時分布····58
5·5 $C(S^*S^{-1})$ の同時分布····63
5·6 固有根の同時分布の基準型····69
演習問題····70

6 標本相関係数

6·1 標本相関係数····71
6·2 標本偏相関係数····75
6·3 標本重相関係数····80

6·4 標本正準相関係数····85
演習問題····87

7 多変量解析における仮説検定と区間推定
7·1 仮説検定····89
7·2 区間推定····94

8 分散共分散行列に関する仮説検定と区間推定
8·1 $H_0: \Sigma=\Sigma_0$ の場合····97
8·2 $H_0: \Sigma_1=\Sigma_2$ の場合····104
8·3 $H_0: \Sigma_1=\Sigma_2=\cdots=\Sigma_k$ の場合····108
8·4 $H_0: \Sigma_{12}=0$ の場合····109
演習問題····114

9 平均ベクトルに関する仮説検定と区間推定
9·1 $H_0: \mu=\mu_0$ の場合····116
9·2 $H_0: \mu_1=\mu_2$ の場合····122
9·3 $H_0: \mu_1=\mu_2=\cdots=\mu_k$ の場合····124
9·4 線型仮説の検定····131
9·5 $H_0: \mu_1=\mu_2=\cdots=\mu_k, \Sigma_1=\Sigma_2=\cdots=\Sigma_k$ の場合····135
9·6 等分散共分散行列の仮定····137
演習問題····140

10 標本分布 (II)
10·1 尤度比検定統計量の分布····142
10·2 最大固有根と最小固有根の分布····149
10·3 T_0^2 統計量の分布····150
演習問題····154

11 主成分分析
11·1 確率ベクトルの主成分····156
11·2 標本ベクトルの主成分····162
11·3 主成分の推定と検定····163
演習問題····167

12 判別分析
12·1 良好な判別法····169
12·2 2個の正規母集団の場合····174

12・3 k 個の正規母集団の場合‥‥178
演習問題‥‥179

付録
A. 基礎数学‥‥181
B. 非心分布‥‥214
C. 統計数値図表‥‥217

参考文献‥‥235
演習問題解答‥‥240
索引‥‥259

1 序 論

1·1 多変量解析とは

F. Galton（ガルトン）[19], K. Pearson（ピアソン）[49], R.A. Fisher（フィッシャー）[12], [14], [15] 等現代統計学の先覚者たちによって，すでに1920 年代までに種々開拓されてきた多変量解析論は，1928 年の J. Wishart（ウィッシャート）[69] および 1931 年の H. Hotelling（ホテリング）[24] の貢献が契機となって，多くの研究者たちによって数々の新しい概念と手法が導入されて，その後 30 年間にいちじるしい発展を遂げ，今日統計理論における重要な分野の一つとしての地位を確保するに至った．特に，1957 年 M.G. Kendall（ケンドール）[37] および S.N. Roy（ロイ）[59] が，また 1958 年 T.W. Anderson（アンダーソン）[3] がそれぞれ多変量解析を単独に取り扱う統計学の専門書を著わしたことはこのことを明らかにするものということができよう．

さて，多変量解析とは，（1）一般に個体 α に関して $p(>1)$ 個の特性を考慮するとき，その各々についての測定値が p 次元確率分布に従う確率ベクトル

$$\boldsymbol{x}_\alpha=\begin{bmatrix} x_{1\alpha} \\ x_{2\alpha} \\ \vdots \\ x_{p\alpha} \end{bmatrix}$$

の実現値として表わされ，（2）N 個の個体（$\alpha=1, 2, \cdots, N$）に対する独立な測定結果はその各列が独立に p 次元確率分布に従う確率行列

$$\boldsymbol{X}=[\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N]$$

$$= \begin{bmatrix} x_{11} & x_{12} & & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & & \vdots \\ x_{p1} & x_{p2} & & x_{pN} \end{bmatrix}$$

の実現値によって与えられる場合の統計理論を取り扱うものであって，$\boldsymbol{x}_\alpha$ と $\boldsymbol{x}_\beta$ $(\alpha, \beta=1, 2, \cdots, N, \alpha \neq \beta)$ は独立であるが，$\boldsymbol{x}_\alpha$ の要素 $x_{i\alpha}$ と $x_{j\alpha}$ $(i, j=1, 2, \cdots, p, i \neq j)$ は一般に独立でないことを特徴とするのである．すなわち，多変量解析は互いに独立でない確率変数 $x_1, x_2, \cdots, x_p$ を要素とする確率ベクトル $\boldsymbol{x}$ の要素間の関係を取り扱うのであって，Kendall [37] によれば，$\boldsymbol{x}$ の特定の1個の要素 x_i（または数個の要素 $x_i, x_j, \cdots$）のその他の要素に対する従属関係（dependence）を取り扱う問題と $\boldsymbol{x}$ の要素の間の相互依存関係（interdependence）を取り扱う問題に二大別される．現在その理論を構成する重要な項目は，1変量の場合の統計的概念と手法を多変量の場合へ直接拡張したもの，各種の相関概念とそれに関する統計的推論，回帰論，主成分分析，判別分析，因子分析など，さらにそれに関連する種々の統計量の標本分布の問題，特に確率行列の固有根の標本分布などである．

統計調査あるいは実験の見地からすれば，各個体について測定する特性の個数，すなわち，確率ベクトルの次元数が増加すればするほどそれだけ情報量が増加するのであるから，自然科学，社会科学，各種産業における統計的現象の多変量的取扱いが盛んになるのは自然の傾向のように思われる．特に，過去30年にわたって多変量解析の理論が整備されると共に，近年，電子計算機の急速な進歩によって多変量解析の応用に伴う数値計算が容易になったため，この傾向は益々助長されるようになってきた．しかしながら，Kendall [37] が指摘するごとく，理論統計家と応用統計家の間には多変量解析に対する態度に根本的な相違があるようである．すなわち，前者は統計理論における1変量の場合の諸概念と手法を2変量の場合へ，また2変量の場合の結果を p 変量の場合へと拡張，一般化することを指向するに対して，後者は応用の場において与えられる多変量のデータについてその次元数（p の値）をできるだけ減少しようと努めるのである．従来，多変量解析の理論は主として生物測定学，計量心理学，計量経済学，人体測定学，品質管理等における応用問題の解決に関連して発展してきたのであるが，これ以外の分野においても，1変量の統計理論が応用される限りは多変量の統計理論の応用の可能性が存在するものということが

できよう．C.R. Rao（ラオ）[56] によれば，多変量解析の応用に関する統計家の見解は多岐にわたっているのであって，その 2，3 をあげれば，

（1） M.S. Bartlett（バートレット）のごとく，きわめて楽観的であって，多変量解析はすでに理論的に安定した地位を確保しているゆえに，今後は各分野における応用を一層盛んにするための研究が望ましいとするもの，

（2） F. Yates（イェーツ）および M.J.R. Healy（ヒーリィ）のごとく，多変量解析の方法に疑問を持ち，統計解析の第一歩はまず個々の測定結果を簡単な 1 変量の手法で解析することが本質的であって，ただそれが絶対必要な場合においてのみ複雑な多変量解析の手法に訴えるべきであるとするもの，

（3） D.J. Finney（フィネー）のごとく，多変量解析のある種の応用に批判的であって，それは科学的研究に何ら貢献するところなく，ただ，研究者個人の数学的興味を満足するにすぎないのであって，一般に多変量解析の手法を実験または調査の結果，得られたデータの解釈という見地から再検討する必要があるとするもの，

（4） J. Tukey（トゥキイ）のごとく，統計家はその問題の所在が不明のとき，とかく多変量解析に逃避しがちであるという痛烈な批判をくだすもの，

などがある．これらの見解，なかんずく，多変量解析の応用についての批判的見解は多変量解析論の健全な発展のために必要であり，特に多変量解析を無用とする批判に対して，Rao は次のごとく答えているが，これは傾聴に値するものである．すなわち，

「統計資料はそれが収集された特定の目的のためのみに吟味されなければならない，あるいは収集された資料は眼前の興味ある問題に解答を与えることができれば十分である，とするのは統計家の正しい態度とはいえない．この態度は確かに科学の進歩のために好ましいものではない．統計的検定の結果の解釈にありがちの陥穽におちいらないようにするためには，精巧な解析法と包括的資料が不可欠であって，科学史をひもとくとき，観察結果に基づく証拠によって得られた新発見のあるものは，資料収集の段階においては，研究者がまったく予期しなかったものであることを想起しなければならない」と．

統計的品質管理に多変量解析の手法を導入した Hotelling [27] は品質管理または抜取検査において多変量解析の手法が 1 変量解析のそれに比して有利であると思われるのは，数多くの部品を組み立てて作った製品の良否を判定する場

合であるとしている．すなわち，簡単な部品に対しては，個別的に1個の特性について計量するか，あるいは単に良品，不良品の判定をくだすことによって抜取検査を実施することもできるが，それぞれ種類を異にする欠点を有する可能性のある部品を数多く組み合わせた複雑な製品については，その性能の如何に基づいて良否の判定をくださなければならない．このためには一特性のみの計量では不十分であって，少なくとも数個の計量を総合的に判断しなければならないのである．

以上において，多変量解析に対する種々の見解を紹介したのであるが，この理論の限界を認めつつ，その範囲内において多変量解析の手法が一層多くの分野に適用されることを期待するものである．

1·2 記 号

本書においては，特に断わらない限り，次の記号を用いることにする．母集団パラメータをギリシャ文字によって，定数をローマ文字アルファベットの前半によって，また，確率変数，標本変量をその後半によって表わす．行列は肉太の大文字で，ベクトルは肉太の小文字で表わす．転置行列はプライム記号をつけて表わす．たとえば，$\boldsymbol{A}'$ は $\boldsymbol{A}$ の転置行列である．ベクトル $\boldsymbol{x}$ は列ベクトルを表わすことにし，したがって，行ベクトルはその転置，$\boldsymbol{x}'$ によって表わす．三角行列は，主対角線の上側の要素がすべて0であるものとし，これをたとえば，$\tilde{\boldsymbol{T}}$ と表わす．主対角線の下側の要素がすべて 0 である三角行列は $\tilde{\boldsymbol{T}}'$ となる．p 行 q 列の行列 $\boldsymbol{A}$ の行数と列数を明示する必要のあるときには，$\boldsymbol{A}(p\times q)$，$\underset{p\times q}{\boldsymbol{A}}$，$\underset{q}{\boldsymbol{A}_p}$ などと書くことにする．したがって，$\boldsymbol{a}(p\times 1)$ は列ベクトル $\boldsymbol{a}$ が p 個の要素をもつことを表わし，これを $\underset{p\times 1}{\boldsymbol{a}}$ とも書くことにする．行列 $\boldsymbol{M}$ の第 i 行，第 j 列要素（これを (i,j) 要素ともいう）が m_{ij} であるとき，$\boldsymbol{M}$ を $[m_{ij}]$ で表わし，また，m_{ij} を $(\boldsymbol{M})_{ij}$ で表わすこともある．正方行列 $\boldsymbol{A}$ の行列式は $|\boldsymbol{A}|$ で，$\boldsymbol{A}$ が正則の正方行列であるとき，その逆行列を $\boldsymbol{A}^{-1}$ で表わす．記号 $\mathrm{tr}\,\boldsymbol{M}$ は正方行列 $\boldsymbol{M}$ の跡（trace）であって，主対角線上の要素の和である．$a_1, a_2, \cdots, a_p$ を対角要素とする対角行列を $\boldsymbol{D}(a_i)$ と書く．特に，単位行列には記号 $\boldsymbol{I}$ を用い，$p\times p$ の単位行列を $\boldsymbol{I}(p)$ によって表わす．正方行列 $\boldsymbol{M}(p\times p)$ が $\boldsymbol{M}\boldsymbol{M}'=\boldsymbol{I}(p)$ であれば，$\boldsymbol{M}$ は直交行列であるといい，また，$\boldsymbol{L}(p\times q)$ $(p<q)$ が $\boldsymbol{L}\boldsymbol{L}'=\boldsymbol{I}(p)$ ならば，$\boldsymbol{L}$ は半直交行列であるという．

正方行列 $\boldsymbol{B}$ の固有根の集合には $C(\boldsymbol{B})$ なる記号を用い，最大根を $C_{\max}(\boldsymbol{B})$，最小根を $C_{\min}(\boldsymbol{B})$，第 i 根を $C_i(\boldsymbol{B})$ で表わす．すべての要素が 0 であるベクトルおよび行列をそれぞれ $\mathbf{0}, \boldsymbol{O}$ で表わす．対称行列 $\boldsymbol{U}$ が正の定符号であること，非負であることをそれぞれ $\boldsymbol{U}>0$, $\boldsymbol{U}\geq 0$ で表わし，また，$\boldsymbol{U}>\boldsymbol{V}$, $\boldsymbol{U}\geq\boldsymbol{V}$ によって，$\boldsymbol{U}-\boldsymbol{V}$ がそれぞれ，正の定符号，非負であることを意味することにする．また，行列 $\boldsymbol{A}$ の階数を $\operatorname{rank}\boldsymbol{A}$ によって表わす．

$\boldsymbol{A}(p\times q)=[a_{ij}]$ のとき，$\prod_{i=1}^{p}\prod_{j=1}^{q} da_{ij}$ を $d\boldsymbol{A}$ によって，また $\boldsymbol{a}'(1\times p)=[a_1, a_2, \cdots, a_p]$ のとき，$\prod_{i=1}^{p} da_i$ を $d\boldsymbol{a}$ によって表わす．変換 $\boldsymbol{X}\to\boldsymbol{Y}$ のヤコビアンを $J(\boldsymbol{X}:\boldsymbol{Y})$ で表わす．すなわち，$\boldsymbol{X}, \boldsymbol{Y}$ がそれぞれ m 個の独立の要素からなるとき，

$$J(\boldsymbol{X}:\boldsymbol{Y})=\frac{\partial(x_1, \cdots, x_m)}{\partial(y_1, \cdots, y_m)}$$

また，記号 $Pr\{\ \}$, $Pr(\ )$ によって，$\{\ \}$，$(\ )$ の中の関係が成立する確率を表わすことにし，確率変数 x の期待値，分散をそれぞれ $E(x)$, $V(x)\equiv E(x-E(x))^2$ によって表わし，確率変数 x と y の共分散を $\operatorname{Cov}(x, y)\equiv E(x-E(x))(y-E(y))$ によって表わす．自由度 f のカイ 2 乗分布に従う確率変数を $\chi^2(f)$，また，その上側確率 α の点を $\chi^2(f, \alpha)$ で表わし，自由度 f_1, f_2 の F 分布に従う確率変数を $F(f_1, f_2)$ によって，その上側確率 α の点を $F(f_1, f_2; \alpha)$ によって表わすことにする．

2 多変量分布

確率変数を要素とする確率ベクトルの従う多次元確率法則は，一般には，離散型あるいは連続型の種々のものを考えなければならないが，従来展開されてきた多変量解析の理論はほとんどすべて多変量正規分布に基づくものである．その理由としては，一方，理論的には一般化された中心極限定理によって多変量正規分布が導入され†，他方，実用的にはこの数学モデルが非常に多くの応用の場に適用可能であること，などをあげることができるであろう．この章では，まず確率ベクトルの従う一般の多変量分布について述べた後，多変量正規分布についてその重要な性質を考察することにする．

2·1 多変量分布††

この節に限って，記述を明瞭にするために，一般の記号と異なって次の記号を用いることにする．確率変数をローマ文字の大文字，それを要素とする確率ベクトルを肉太の大文字で，また，確率変数の実現値をローマ文字の小文字，それを要素とするベクトルを肉太の小文字で表わすことにする．確率変数を要素とする行列，すなわち，確率行列はその行数，列数を明示して肉太の大文字で表わす．さて，p 変量分布とは，p 次元標本空間 R_p における p 個の確率変数の組 $(X_1, \cdots, X_p)$ の同時分布，あるいは，これをベクトルで表わした $1\times p$ の確率ベクトル $\boldsymbol{X}'=[X_1, \cdots, X_p]$ の分布である．

1. 分布関数，密度関数

1変量のときと同様に，$1\times p$ の確率ベクトル

† たとえば Cramér [9] p. 316f を参照せよ．

†† 詳細については，たとえば，Cramér [9] p. 260ff, Wilks [68] p. 39ff を参照せよ．

$$\boldsymbol{X}'=[X_1, \cdots, X_p]$$

の分布は，$x_1, \cdots, x_p$，すなわち，$1\times p$ のベクトル $\boldsymbol{x}'=[x_1, \cdots, x_p]$ の実数値1価関数

$$F(\boldsymbol{x})\equiv F(x_1, \cdots, x_p)=Pr\{X_1\leq x_1, \cdots, X_p\leq x_p\} \qquad (1)$$

によって一意的に決定される．ここに，$F(\boldsymbol{x})$ は確率ベクトル $\boldsymbol{X}$ の（累積）分布関数（cumulative distribution function, c.d.f.）と呼ばれ，次の性質をもっている．

1° $0\leq F(\boldsymbol{x})\leq 1$

2° p 次の階差を $\Delta_p F(\boldsymbol{x})$ で表わせば，$\Delta_p F(\boldsymbol{x})\geq 0$

3° $F(-\infty, x_2, \cdots, x_p)=\cdots=F(x_1, \cdots, x_{p-1}, -\infty)=0$

4° $F(+\infty, \cdots, +\infty)=1$

5° $F(x_1, \cdots, x_p)$ は $x_1, \cdots, x_p$ の各々について右側から連続である．

もし，確率ベクトル $\boldsymbol{X}$ の分布が連続型であるならば，ほとんどいたる所において，

$$f(\boldsymbol{x})\equiv f(x_1, \cdots, x_p)=\frac{\partial^p F(x_1, \cdots, x_p)}{\partial x_1 \cdots \partial x_p} \qquad (2)$$

が存在する．この $f(\boldsymbol{x})$ は $\boldsymbol{X}$ の密度関数（probability density function, p.d.f.）と呼ばれ，$f(\boldsymbol{x})\geq 0$ である．p 次元ユークリッド空間の任意の可測集合 E に対して，

$$Pr\{\boldsymbol{x}\in E\}=\int_E \cdots \int f(\boldsymbol{x})d\boldsymbol{x} \qquad (3)$$

である．

2. 周辺分布，統計的独立

$1\times p$ の確率ベクトル

$$\boldsymbol{X}'=[X_1, \cdots, X_p]$$

の分布関数を $F(\boldsymbol{x})$ とするとき，$X_1, \cdots, X_p$ の部分，たとえば，$X_{i_1}, \cdots, X_{i_q}$ $(q<p)$ の分布関数は

$$\begin{aligned} F_{i_1 \cdots i_q}(x_{i_1}, \cdots, x_{i_q}) &= Pr\{X_{i_1}\leq x_{i_1}, \cdots, X_{i_q}\leq x_{i_q}\} \\ &= F(+\infty, \cdots, x_{i_1}, +\infty, \cdots, +\infty, x_{i_q}, +\infty, \cdots, +\infty) \end{aligned} \qquad (4)$$

によって与えられ，これを $X_{i_1}, \cdots, X_{i_q}$ の周辺分布関数（marginal distribution function）という．ここに，(4) の右辺は $\boldsymbol{X}$ の分布関数 $F(x_1, \cdots, x_p)$ の中の $x_{i_1}, \cdots, x_{i_q}$ 以外の $(p-q)$ 個の変数を $+\infty$ としたものであって，p 次

元空間における確率分布を q 次元部分空間に射影したものである．分布が連続型であるならば，$\boldsymbol{X}$ の密度関数 $f(x_1, \cdots, x_p)$ を $x_{i_1}, \cdots, x_{i_q}$ 以外の $(p-q)$ 個の変数についてその全変域において積分すれば，$X_{i_1}, \cdots, X_{i_q}$ の周辺密度関数 $f_{i_1 \cdots i_q}(x_{i_1}, \cdots, x_{i_q})$ が得られる．

確率変数 $X_1, \cdots, X_p$ が条件

$$F(x_1, \cdots, x_p) = F_1(x_1) \cdots F_p(x_p) \qquad (5)$$

を満足するとき，$X_1, \cdots, X_p$ は統計的に独立であるという（以下，単に独立ということもある）．ここに，$F_i(x_i)$ は X_i の周辺分布関数である．もし分布が連続型であるならば，（5）の条件は

$$f(x_1, \cdots, x_p) = f_1(x_1) \cdots f_p(x_p) \qquad (6)$$

と同等である．ここに，$f_i(x_i)$ は X_i の周辺密度関数である．また，$(p+q) \times 1$ の確率ベクトル

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix}_q^p$$

の分布関数 $F(\boldsymbol{x})$ が

$$F(\boldsymbol{x}) \equiv F(\boldsymbol{x}_1, \boldsymbol{x}_2) = F_1(\boldsymbol{x}_1) F_2(\boldsymbol{x}_2) \qquad (7)$$

と書くことができるならば，$p \times 1$ の確率ベクトル $\boldsymbol{X}_1$ と $q \times 1$ の確率ベクトル $\boldsymbol{X}_2$ は統計的に独立である．ここに，$F_1(\boldsymbol{x}_1), F_2(\boldsymbol{x}_2)$ はそれぞれ $\boldsymbol{X}_1, \boldsymbol{X}_2$ の周辺分布関数である．

3. 条件付分布

$$\boldsymbol{X}_1 = \begin{bmatrix} X_1^{(1)} \\ \vdots \\ X_p^{(1)} \end{bmatrix}, \qquad \boldsymbol{X}_2 = \begin{bmatrix} X_1^{(2)} \\ \vdots \\ X_q^{(2)} \end{bmatrix}$$

をそれぞれ $p \times 1, q \times 1$ の確率ベクトルとし，E_1, E_2 をそれぞれの空間における可測集合とする．このとき，事象 $\boldsymbol{X}_2 \in E_2$ の下における事象 $\boldsymbol{X}_1 \in E_1$ の条件付確率は

$$Pr\{\boldsymbol{X}_1 \in E_1 | \boldsymbol{X}_2 \in E_2\} = \frac{Pr\{\boldsymbol{X}_1 \in E_1, \boldsymbol{X}_2 \in E_2\}}{Pr\{\boldsymbol{X}_2 \in E_2\}} \qquad (8)$$

によって与えられる．ただし，$Pr\{\boldsymbol{X}_2 \in E_2\} > 0$ とする．ここにおいて，

$$E_1 = \{\boldsymbol{X}_1 : X_1^{(1)} \leq x_1^{(1)}, \cdots, X_p^{(1)} \leq x_p^{(1)}\},$$
$$E_2 = \{\boldsymbol{X}_2 : X_1^{(2)} \leq x_1^{(2)}, \cdots, X_q^{(2)} \leq x_q^{(2)}\}$$

とすれば，（8）の右辺は $\boldsymbol{X}_1$ と $\boldsymbol{X}_2$ の同時分布関数と $\boldsymbol{X}_2$ の分布関数の比で