統計用語辞典

芝　祐順・渡部　洋・石塚智一　編

新曜社

新曜社

統計用語辞典

初版第1刷発行　1984年5月15日　©
初版第4刷発行　1989年3月20日

編　者　芝　祐順
　　渡部　洋
　　石塚　智一
発行者　堀江　洪
発行所　株式会社 新曜社
〒101 東京都千代田区神田神保町2-10 多田ビル
電話(03)264-4973(代)・振替東京2-108464

印　刷　真珠社　　Printed in Japan
製　本　イマキ製本所
書籍コード　3541-4133-3329

凡　例

1. 見出し語
   (1) 見出し語の配列は，五十音順とした。
      (a) 濁音・半濁音は，清音とみなした。
      (b) 長音は，無視して配列した。
      (c) 促音・拗音は，それぞれ独立の一字とみなして配列した。
   (2) 見出し語には英語を付した。英語が複数該当する場合には，使用頻度が高いと思われるものを先に配列した。
   (3) ＳＳＡなど，通常略記法で呼ばれる項目については，そのまま見出し語に取り上げた。したがって，ＳＳＡは，エスエスエーで検索していただきたい。
2. 参照項目の表示
   (1) 本文中の↑印を付した語と（　）内の→印につづく語は，その語が見出し語となっていることを示す。適宜参照されたい。
   (2) 項目末尾の→は，関連項目を示す。
   (3) 項目末尾の⇒は，同義語を示し，⇒項目の中で解説されていることを示す。
3. 外国人名等のカナ表記
   (1) 外国人名等のカナ表記は，原則として，その国の発音に沿うようにしたが，わが国で一般的に用いられている読み方がある場合には，それに従った。
   (2) 二人以上の人名が連らなる場合は，・で区切った。
4. 文　献
   (1) 文献は，重複と煩雑さを避けるため巻末の文献表に一括した。したがって，各項目末尾には，著者名と年号のみ記してある。文献表，各項目末尾とも，配列はＡＢＣ順である。
   (2) 外国文献で邦訳のあるものについては，文献表の原著の後に（　）を付して明記した。項目末尾の表示においては，原著のみをあげるにとどめ，＊印を付して邦訳のあることを示した。
5. 索　引
   (1) 和文索引の配列は五十音順とし，すべての項目に英語を付した。したがって，統計和英辞典としての利用が可能である。
   (2) 英文索引の配列はＡＢＣ順とし，すべての項目に日本語を付した。したがって，統計英和辞典としての利用が可能である。
   (3) ページを示す数字のうち太字のものは，項目見出しの所在を示し，数字の大小に関係なく最初に配した。
   (4) 数字の後のＬはそのページの左列を，Ｒは右列を示す。

あ

赤池情報量基準〔Akaike's information criterion, AIC〕統計的モデルを評価する基準として提案された次のような統計量.

$$\mathrm{AIC}=-2\log M+2k$$

ここで，M は与えられたデータによるそのモデルの最大尤度（→尤度関数）を表し，k はそのモデルの中で自由に変化させることのできるパラメタの数を表す．この AIC の値が低いほど良いモデルであると評価する．ふつう，最大尤度が大きいほど，モデルは良いとされるが，AIC はパラメタ数の項を含むため，パラメタ数の少ないモデルをより良いものとするという原理が，評価の中に加わってくる．

複数の可能なモデルの中から AIC 最小のモデルを採用するというやり方で選ばれたモデルのもとでのパラメタの推定値を最小 AIC 推定値（MAICE, minimum AIC estimate）と呼ぶ.（中川徹・小柳義夫, 1982；奥野忠一他, 1976；坂元慶行他, 1983；佐和隆光, 1979）〔芝〕

アセンブラ言語〔assembler language〕アセンブリ言語ともいう．†機械語の1つひとつについて，たとえば加算を表す命令を add の頭文字をとって A，減算を表す命令を subtract の頭文字をとって S というように，人間が理解しやすいように記号化した†プログラム言語．アセンブラによって機械語にほぼ1対1の対応（1対多数の場合もある）で翻訳されてから実行される．機械語に近いのでプログラムは難解であるが，使用するコンピュータに対してより柔軟で多様なプログラムが可能である.（大駒誠一, 1979）〔塗師〕

アノヴァ〔ANOVA〕analysis of variance の略．⇒分散分析

アーラン分布〔Erlang distribution〕確率変数 X が，確率密度関数

$$f(x)=\frac{\lambda^{\alpha}e^{-\lambda x}x^{\alpha-1}}{\Gamma(\alpha)},\qquad x>0$$

をもつとき，X はアーラン分布に従うという．ただし，λ と α はパラメタで，$\lambda>0$ であり，α は正の整数である．したがって，†ガンマ分布との違いは，α が整数であるか否かによる．もし，$\alpha=1$ ならば，この分布は†指数分布に一致する．アーラン分布の平均値は α/λ であり，分散は α/λ^2 である．†積率母関数は $t<\lambda$ に対して存在し，$M(t)=\lambda^{\alpha}/(\lambda-t)^{\alpha}$ で与えられる.

確率変数 $X_1, X_2, \cdots$ が独立に上記のアーラン分布に従うものとし，N をある特定の $x(x>0)$ に対して

$$\begin{cases} N=0, & X_1>x \\ N=n, & X_1+\cdots+X_n\leqq x<X_1+\cdots+X_{n+1} \end{cases}$$

で定義される確率変数とすると，N は一般†ポアソン分布に従い，

$$P(N\leqq n)=e^{-\lambda x}\sum_{i=0}^{(n+1)\alpha-1}[(\lambda x)^i/i!]$$

となる．また，確率変数 $X_1, \cdots, X_n$ が $\alpha=1$ のアーラン分布（すなわち指数分布）に従うならば，その線型形式

$$Y=\sum_{i=1}^{n}\lambda_i X_i,\qquad \lambda_i\neq\lambda_j$$

は，密度関数

$$f(y)=\sum_{i=1}^{n}\frac{\lambda_i^{n-2}\,e^{-y/\lambda i}}{\prod_{j\neq i}(\lambda_i-\lambda_j)},\qquad y>0$$

をもち，一般アーラン分布（または一般ガンマ分布）と呼ばれ，システムの†信頼性理論等で利用される.（Johnson, N. L. & Kotz, S., 1970a）〔渡部〕

アルゴル〔ALGOL〕Algorithmic Language の略．主として科学技術計算に用いられる†プログラム言語．文字通りアルゴリズムの記述を目的として考案された言語で，文法記述が厳密であり，歴史的に他のプログラム言語に与えた影響は大きい．実用的には，†フォートランなどに比べると初心者には理解しにくいため，普及度は低い．日本工業規格で規格が5つ

の水準に分けて定められている．(米田信夫・野下浩平，1979)　〔塗師〕

アール推定量〔*R*-estimator〕　†位置母数に関する †頑健性の高い推定量を分類したものの1つ．順位を用いた，位置母数に関する推定の考えに関連づけられるもので，そのような推定において最も「棄却されにくい仮説」となる母数の値を，その推定量とするものである．この中には，†ホッジス・レーマン推定量や，†正規スコア検定から導かれる *R* 推定量，†順位和検定から導かれる *R* 推定量などがある．(竹内啓・大橋靖雄，1981)　〔芝〕

アルファ因子分析〔alpha factor analysis〕†因子分析における解の1つで，因子分析の基本モデル $\boldsymbol{Z=FA'+UD}$ の中の共通因子成分 $\boldsymbol{Y=FA'}$ の線型結合 $\boldsymbol{Yw}$ を因子とするが，そのときの重みベクトル $\boldsymbol{w}$ には線型結合の †アルファ係数を最大とするものを選ぶ．アルファ係数とは線型結合として合成された下位変数の等質性の測度で，心理テストなどの特性を表すのに用いられている係数である．アルファ因子分析の解を求めるには，共通性を要素とする対角行列 $\boldsymbol{H^2}$ を用いて，対称行列 $\boldsymbol{H^{-1}(R-D^2)H^{-1}}$ の固有値の平方根を成分とする対角行列 $\boldsymbol{\theta}$ と，規準化された固有ベクトルを列にもつ行列 $\boldsymbol{Q}$ とを求め，それらから因子負荷行列 $\boldsymbol{A=HQ\theta}$ を求めればよい．(Harman, H. H., 1976；芝祐順，1979)　〔芝〕

アルファ過誤〔alpha error〕 ⇒第一種の誤り

アルファ係数〔coefficient alpha〕テストスコア X が n 個の下位テストスコア $Y_1,\cdots,Y_n$ から構成されているとき，アルファ係数は

$$\alpha=\frac{n}{n-1}\left[1-\frac{\sum_j\sigma^2(Y_j)}{\sigma^2(X)}\right] \qquad (1)$$

によって定義される．テスト X の †信頼性を ρ^2_{XT} とするとき，アルファ係数と信頼性の間には

$$\rho^2_{XT}\geqq\alpha \qquad (2)$$

が成立する．等号は n 個の下位テストが本質的タウ等価（下位テスト同士の真の得点の差がどの個人においても一定であることをいう）であるときに成立する．すなわち，アルファ係数は信頼性の下限を与える．

テストスコアの平方和を個人間平方和，テスト間平方和および個人とテストの交互作用の平方和に分割したとき，アルファ係数の推定値は

$$\hat{\alpha}=1-\frac{\text{交互作用の平均平方}}{\text{個人間の平均平方}}$$

と表すこともできる．

下位テストの得点が2値データとして与えられているとき，下位テストに正答する確率を p_j とすれば Y_j の分散は $\sigma^2(Y_j)=p_j(1-p_j)$ となるので

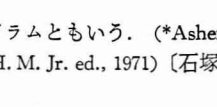
$$\alpha_{20}=\frac{n}{n-1}\left[1-\frac{\sum_j p_j(1-p_j)}{\sigma^2(X)}\right] \qquad (3)$$

を得る．(3)式を †キューダー・リチャードソンの公式20（KR-20）という．アルファ係数の定義は，テストの信頼性の下限を与える KR-20 を一般の下位テストに拡張したものと考えることができる．また，n 個の下位テストの難易度が均等（$p_j=\bar{p}$；$j=1,\cdots,n$）のとき，(3)は

$$\alpha_{21}=\frac{n}{n-1}\left[1-\frac{n\bar{p}(1-\bar{p})}{\sigma^2(X)}\right] \qquad (4)$$

と書くことができる．(4)を KR-21 と呼ぶ．(池田央，1973)　〔石塚〕

アロー ダイアグラム〔arrow diagram〕†因果分析において，変数間や概念間の因果関係を示すために用いられる矢印線図．たとえば右図は，概念 X_1, X_2, X_3, X_4 の間において矢印の向きに因果関係のあることを示している．パスダイヤグラムともいう．(*Asher, H. B., 1976；Blalock, H. M. Jr. ed., 1971)　〔石塚〕



アンコヴァ〔ANCOVA〕 analysis of covariance の略．⇒共分散分析

アンサリ・ブラドレイ の検定〔Ansari-Bradley test〕順位にもとづく †ノンパラメトリック検定の1つで，独立な2群のデータが与えられたとき，それらに対応する母集団分布の

同一性の仮説を検定するもの．対立仮説は，分布のひろがりが異なるというもので，2つの母集団の中央値の等しいことが仮定されている．検定には次のような統計量が用いられる．まず，両群の$N(=n_1+n_2)$個のデータ$x_1, x_2, \cdots, x_{n_1}; y_1, y_2, \cdots, y_{n_2}$をこみにして，小さいものから順に並べる．そして，最小のものと最大のものとに，ともに順位1をつける．続いて最小および最大からそれぞれ2番目のものに順位2をつける．以下順に中央のものまで順位をつけていく．最終の順位は$N/2$となる〔ただし，n_1+n_2が奇数の時は$(N+1)/2$となる〕．こうしてつけられた順位の和Sをxの群について求める．これが検定に用いられる統計量である．データの数の少ない場合の検定のために数表が用意されている．データ数が多いときには，この統計量が，上の仮説のもとで近似的に

$$\text{平均}=\frac{1}{4}n_1(N+2),\quad \text{分散}=\frac{n_1n_2(N^2-4)}{48(N-1)}$$

の正規分布に従うことを利用して検定をする．なお，Nが奇数のときは，

$$\text{平均}=\frac{n_1(N+1)^2}{4N},$$

$$\text{分散}=\frac{n_1n_2(N+1)(N^2+3)}{48N^2}$$

となる．(Hollander, M. & Wolfe, D. A., 1973; 岩原信九郎, 1964)　〔芝〕

アンダーソンのU〔Anderson's U〕　一般†線型仮説の検定に用いられる統計量で，$U=|N\hat{\boldsymbol{\Sigma}}_\Omega|/|N\hat{\boldsymbol{\Sigma}}_\omega|$ によって定義される．ここで，Nは標本数，$\hat{\boldsymbol{\Sigma}}_\omega$は帰無仮説のもとでの分散共分散行列$\boldsymbol{\Sigma}$の最尤推定量，$\hat{\boldsymbol{\Sigma}}_\Omega$は仮説を設けない場合の$\boldsymbol{\Sigma}$の最尤推定量である．複数個の平均ベクトルの間に差が無いというのが帰無仮説であるとき，Uはウィルクスのラムダと呼ばれることが多い．より一般的な場合として，線型モデル$\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{E}$を考えよう．ただし，$\boldsymbol{Y}$は$N\times p$のデータ行列，$\boldsymbol{X}$は$N\times q$の†計画行列，$\boldsymbol{\beta}$は$q\times p$の未知母数行列，$\boldsymbol{E}$は$N\times p$の誤差行列とし，$N\geqq p+q$とする．また，$\boldsymbol{E}$の各行は相互に独立に多変量正規分布$N_p(\boldsymbol{0}', \boldsymbol{\Sigma})$に従うものとする．$\boldsymbol{\Sigma}$は正値の母集団分散共分散行列である．このモデルのもとで，線型仮説$H_0: \boldsymbol{A}\boldsymbol{\beta}=\boldsymbol{C}$を検定したいものとする．ただし，$\boldsymbol{A}$は$r\times q$の階数$r$の任意の指定された行列であり，$\boldsymbol{C}$は$r\times p$のやはりあらかじめ指定された行列とする．このとき，†尤度比検定をおこなうための検定統計量として

$$U=\frac{|N\hat{\boldsymbol{\Sigma}}_\Omega|}{|N\hat{\boldsymbol{\Sigma}}_\omega|}$$

を考えることができる．これがアンダーソンのUである．ただし，

$$N\hat{\boldsymbol{\Sigma}}_\Omega=(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})$$

$$N\hat{\boldsymbol{\Sigma}}_\omega=(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}^*)'(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}^*)$$

であり，$\hat{\boldsymbol{\beta}}$は制約条件がないときの$\boldsymbol{\beta}$の最尤推定量で$\hat{\boldsymbol{\beta}}^*$は仮説$H_0$のもとでの$\boldsymbol{\beta}$の最尤推定量である．→平均ベクトルに関する尤度比検定，ウィルクスのラムダ (Anderson, T. W., 1958; *Rao, C. R., 1973)　〔渡部〕

安定分布〔stable distribution〕　相互に独立に同じ分布に従う確率変数の線型関数の和が，その分布と同じ†分布族に属するとき，その族は安定分布の族であるといわれる．たとえば，確率変数XとYが相互に独立に正規分布$N(\theta, \sigma^2)$に従うならば，$b_1>0$，$b_2>0$および$b_3>0$として，

$$Z=\frac{b_1X+b_2Y+c}{b_3}$$

はやはり正規分布に従うので，正規分布は安定しているといわれる．より厳密には，操作$*$が†たたみこみを表すとすると，もし，$b_1>0$，$b_2>0$および実数c_1, c_2に対して

$$F\left(\frac{x-c_1}{b_1}\right)*F\left(\frac{x-c_2}{b_2}\right)=F\left(\frac{x-c}{b}\right)$$

が成り立つような正の数bと実数cが存在するならば，†分布関数$F(x)$は安定しているという．

安定分布の応用上の重要性は，たとえば，正規分布よりも裾の重い分布を考えるときや，相互に誤差をもつ変数間の線型回帰を考えるときなどに示される．(Press, S. J., 1972)　〔渡部〕

鞍点(ゲームの)〔saddle point〕　2人ゼロ

和↑決定ゲームにおいて，行プレーヤーのとる最適純粋方略としてのマクシミンの点と，列プレーヤーのとる最適純粋方略としてのミニマックスの点が，一致して鞍点となる．たとえば，3×3 の↑利得行列において，$(C_2 : R_1)=4$ が鞍点となる．

	R_1	R_2	R_3
C_1	2	1	8
C_2	4	5	6
C_3	0	7	3

→マクシミン方略，ミニマックス方略

(*Coombs, C. H. *et al.*, 1970；*Gass, S. I., 1975)

〔岡本〕

い

イーヴィーエム〔EVM〕 errors in the variables model の略で，回帰モデルにおける変数内誤差モデルのこと．⇒変数内誤差モデル

イェーツの連続修正〔Yates' continuity correction〕 度数で表されるような離散型分布をカイ二乗分布や正規分布などの連続型分布で近似することによって統計的仮説検定をおこなおうとするときに用いられる修正手続きのこと．たとえば，次のような 2×2 分割表

特性A	特性B 有	特性B 無	計
有	n_{11}	n_{12}	$n_{1\cdot}$
無	n_{21}	n_{22}	$n_{2\cdot}$
計	$n_{\cdot 1}$	$n_{\cdot 2}$	n

によって示されたデータにおいて↑カイ二乗検定をおこなおうとすれば，検定統計量は

$$\chi^2=\frac{(n_{11}n_{22}-n_{12}n_{21})^2 n}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}$$

で与えられるが，これにイェーツの連続修正をおこなえば

$$\chi^2=\frac{\left(|n_{11}n_{22}-n_{12}n_{21}|-\frac{n}{2}\right)^2 n}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}$$

となる．また，2つの独立した標本における比率が次のようであるときに，

	標本の大きさ	標本比率
標本1	n_1	p_1
標本2	n_2	p_2
全体	n	p

2つの標本比率 p_1 および p_2 の差に関する検定を標準正規分布を利用しておこなおうとするならば，検定統計量は

$$z=\frac{|p_2-p_1|}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

となるが，これに連続修正を加えれば

$$z=\frac{|p_2-p_1|-\frac{1}{2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

となる．修正をおこなうと，おこなわない場合に比べて検定力が低下するが，修正をおこなった方がより正確な確率値をもった χ^2 や z が得られるので連続修正は常に用いた方がよいとする人が多い．(*Everitt, B. S., 1977；*Fleiss, J. L., 1973)

〔渡部〕

閾値誤差損失関数〔threshold loss function〕 推定値の誤差の絶対値が一定の値以下の場合はその誤差は無視し得るが，その値以上になると↑損失が生じるタイプの損失関数をさす．誤差が一定の値以上のときの損失が，誤差の値によらず定数のとき，これは，↑1-0 損失関数と一致する．(Novick, M. R. & Jackson, P. E., 1974)

〔繁桝〕

異常値〔abnormal value, outlier〕 ⇒外れ値

一意性の定理〔uniqueness theorem〕 確率変数の分布 $f(x)$ が与えられたとき，↑積率

が存在すれば，それを計算することが可能である．これとは逆に，一連の積率が与えられたときに，それらに対応する確率変数の分布 $f(x)$ を求める問題を，一般に積率問題（moment problem）という．この問題を解くために，確率変数 X の†積率母関数が定まれば，その分布 $f(x)$ も一意的に定まることが望ましいが，それは一意性の定理によって保証される．すなわち一意性の定理は，0を含む区間 $-h^2<t<h^2$（h は任意の実数）の t のすべての値に対して積率母関数 $M(t)$ が存在するならば，この積率母関数をもつ確率変数はただ1つだけ存在することを述べる．この一意性の定理により，たとえば，2つの確率変数の積率母関数が等しいならば，これら2つの確率変数の分布は等しいことがわかる．（森田優三, 1968）

〔渡部〕

一元配置〔one-way layout〕 ⇒完全無作為配置

一次独立〔linearly independent〕 n 個のベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ からなるベクトルの組に対して，定数 $c_1, c_2, \cdots, c_n$ を考える．これらのベクトルの線型結合 $c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n$ が

$$c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n=\boldsymbol{0}$$

となるのが，$c_1=c_2=\cdots=c_n=0$ のときに限るという場合，ベクトルの組 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ は一次独立であるという．ベクトルの組 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ が一次独立でないとき，すなわち，すべてが0ではない定数 $c_1, c_2, \cdots, c_n$ を用いて

$$c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n=\boldsymbol{0}$$

となる場合に，このベクトルの組は一次従属（linearly dependent）であるという．ベクトルの組 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ は一次独立でなければ，一次従属である．また，一次従属でなければ，一次独立である．一次独立および一次従属という用語の代りに，線型独立および線型従属という用語を用いることもある．また，ある個数のベクトルの組から選ぶことのできる一次独立なベクトルの最大の個数を，そのベクトルの組の†階数という．（岡太彬訓, 1977）

〔岡太〕

イチゼロ（1-0）損失関数〔1-0 loss function〕 母数 θ の推定問題において，θ とその推定値 $\hat{\theta}$ とのずれが，定数 b よりも大きい場合にのみ一定の損失 l が生じると考えられる場合，その損失関数は，

$$\begin{cases} l(\hat{\theta}, \theta)=l, & |\hat{\theta}-\theta|\geq b \\ l(\hat{\theta}, \theta)=0, & |\hat{\theta}-\theta|<b \end{cases}$$

となる．これを，1-0損失関数または単純損失関数という．この損失関数が設定された場合の θ のベイズ推定値は，b が十分小さいとき，θ の分布の†中央値である．（Novick, M. R. & Jackson, P. E., 1974）

〔繁桝〕

位置母数〔location parameter〕 確率変数 X の母数 θ が，X にある定数 c を加えたとき，θ から $\theta+c$ に変化するならば，この母数 θ を位置母数と呼ぶ．たとえば，正規分布における平均は位置母数である．

また広義には，母数 θ がその推定量 $\hat{\theta}$ の分布の位置母数となっているときにも，θ を X の分布の位置母数と呼ぶこともある．たとえば，†正規線型モデル等では，広義の意味で用いられる．

〔渡部〕

一様最強力検定〔uniformly most powerful test, UMP test〕 †統計的仮説検定で，1個のパラメタ θ の値について検定する場合のうち，検定仮説が $\theta=\theta_0$，対立仮説が $\theta>\theta_0$ である場合（または $\theta=\theta_0$ 対 $\theta<\theta_0$）の検定に用いられる概念．棄却域の設定のしかたに応じて，検定仮説 $\theta=\theta_0$ を誤って棄却する確率 α と，対立仮説 $\theta=\theta_1$ $(\theta_1>\theta_0)$ を誤って棄却する確率 β は変化する．対立仮説としてはいろいろな θ の値が考えられるので，β を θ の関数として表し $\beta(\theta)$ と書く．α をある指定された数値（†有意水準 α という）に選んで，棄却域を設定するさい，いろいろな棄却域の選び方がある．このとき，対立仮説のパラメタ θ の値が $\theta>\theta_0$ のどの値になっていても $\beta(\theta)$ が最も小さく〔†検定力 $1-\beta(\theta)$ が最も大きく〕なるような棄却域の選び方があれば，それを（有意水準 α の）一様最強力検定，または最良検定という．なお，$\theta=\theta_0$ 対 $\theta\neq\theta_0$ の検定問題においては，どのような θ に対しても検定力が最大にな

るような棄却域は通常はつくれないことがわかっている.(石井吾郎, 1968；竹内啓, 1963)〔生澤〕

一様分布〔uniform distribution〕 連続変数の確率分布で，定義域のどの点の密度も等しい分布のこと．その形から長方形分布，あるいは矩形分布ともいわれる．その区間を a から b までとすれば，次のような密度関数となる．

$$f(x)=\frac{1}{b-a}, \qquad a<x<b$$
$$=0 \quad , \qquad x\leqq a \text{ および } x\geqq b$$

この分布は平均 $E(X)=(a+b)/2$，分散 $V(X)=(b-a)^2/12$ となる．

また，離散変数の場合には，離散一様分布と呼ばれ，すべての値が等確率となるものをさす．たとえば，サイコロの目は $X=1, 2, \cdots, 6$ であり，どの目も 1/6 の確率で出現すると考えればこれは離散一様分布である．これを密度関数として示せば，次のようになる．

$$f(x)=1/6, \qquad x=1, 2, 3, 4, 5, 6$$

(森田優三, 1968)　〔岡本〕

一致係数〔coefficient of concordance〕 対応する k 組の観測値につけられた順位について，その一致の程度を表す係数．→ケンドールの一致係数 (池田央, 1976)

なお，†因子分析では，異なる因子解の間の因子負荷の一致の度合を表す係数を一致係数 (coefficient of congruence) という．(Harman, H. H., 1976)　〔芝〕

一致推定量〔consistent estimator〕 標本の大きさを大きくすればするほど，ますます正確に推定すべき母数に近づくような †推定量を一致推定量と呼ぶ．より厳密には θ を推定すべき母数とし，$\hat{\theta}_n$ を大きさ n の標本にもとづく推定量とするとき，任意の正の値 ε に対して

$$\lim_{n\to\infty} P(|\hat{\theta}_n-\theta|\geqq\varepsilon)=0$$

となるならば，$\hat{\theta}_n$ を θ の一致推定量という．いいかえれば，$n\to\infty$ のとき $\hat{\theta}_n$ が θ に確率収束すればよい．→確率収束 (竹内啓, 1963)　〔渡部〕

一対比較法(いつついひかくほう)〔method of paired comparison〕 官能検査や尺度構成 (→尺度構成法) のために刺激を対にして被験者に提示し，刺激間の選好や大小に関する判断を求める方法．n 個の刺激 $A_1, A_2, \cdots, A_n$ があるとき，$\frac{1}{2}n(n-1)$ 個の刺激対 $(A_1, A_2), (A_1, A_3), \cdots, (A_1, A_n), (A_2, A_3), \cdots, (A_{n-1}, A_n)$ を被験者に提示して判断を求める．提示順序が問題となるときには，さらに，順序を逆にした対を加えて，$n(n-1)$ 個の刺激対に対する判断が必要である．

一対比較法は，被験者に求める判断が，2者の中から1つを選ぶ，という比較的簡単なものであるため広く用いられるが，刺激の数が増えると，必要な判断の総数は急速に増大する．

一対比較データを処理する方法には，†加算モデルを当てはめるシェフェの方法，判定比と呼ばれるパラメタを導入するブラドレーの方法，比較判断の法則と呼ばれる判断モデルを用いるサーストンの方法などがある．

シェフェの方法では，たんに A_i と A_j の大小判断だけでなく，その違いの程度に対する判断も必要とする．すなわち A_i が A_j に対して非常に大きいとき +2，いくぶん大きいとき +1，等しいとき 0，いくぶん小さいとき −1，非常に小さいとき −2 をそれぞれ評点として割り当てる．A_i と A_j に対する l 番目の被験者の評点を x_{ijl} とするとき，モデルは

$$x_{ijl}=(\alpha_i-\alpha_j)+\gamma_{ij}+\delta_{ij}+e_{ijl}$$

で与えられる．ここに，α_i は刺激の平均強度，$\gamma_{ij}=-\gamma_{ji}$ は組合せの効果，$\delta_{ij}=\delta_{ji}$ は提示順序の効果である．

ブラドレーのモデルでは A_i が A_j よりも大きいと判断する確率 p_{ij} を

$$p_{ij}=\frac{\alpha_i}{\alpha_i+\alpha_j}$$

で与える．ここに，α_i は判定比と呼ばれるパラメタで，α_i の大きいほど，刺激の大きいことを示している．

サーストンの比較判断の法則では，刺激 A_i の心理的強度 a_i の心理学的連続体上の分布が $N(\alpha_i, \sigma_i^2)$ に従うことが仮定されている．したがって，A_i が A_j より大きいと判断される確率は

$$p_{ij}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{\alpha_i-\alpha_j}{\sigma_{ij}}} e^{-\frac{1}{2}t^2}dt$$

$$\sigma_{ij}=(\sigma_i^2+\sigma_j^2-2\rho_{ij}\sigma_i\sigma_j)^{1/2}$$

で与えられる．ここに ρ_{ij} は a_i と a_j の相関係数である．サーストンの比較判断の法則は，ρ_{ij} や σ_i に対する仮定の強さによって，5つのケースに分類される．一般には最も強い仮定

$$\rho_{ij}=0,\qquad \sigma_i=\sigma_j=\sigma$$

をおくケースVがよく用いられている．(武藤真介，1982；日科技連官能検査委員会編，1973；西里静彦，1975)　〔石塚〕

一般化逆行列〔generalized inverse (matrix)〕　正方行列で，しかも正則な行列（→正則行列）にのみ定義できる†逆行列の概念を，正則でない正方行列，さらに正方型でない一般の行列に拡張したもの．(n,m) 型行列 $\boldsymbol{A}$を，m 次元ベクトル $\boldsymbol{x}$ の n 次元ベクトル $\boldsymbol{y}$ への線型写像と想定する場合，$\boldsymbol{y}$ を $\boldsymbol{x}$ に移す$\boldsymbol{A}$ の逆写像を $\boldsymbol{A}^-$ と表し，これを一般化逆行列と定義する．一般化逆行列の最も一般的な定義は，次式である．

(a)　$\boldsymbol{AA^-A=A}$

また，この式から次式が成立する．

(i)　$\mathrm{rank}(\boldsymbol{A^-})\geqq\mathrm{rank}(\boldsymbol{A})$

(ii)　$\mathrm{rank}(\boldsymbol{A})=\mathrm{rank}(\boldsymbol{A^-A})=\mathrm{rank}(\boldsymbol{AA^-})$

(iii)　$\boldsymbol{Ax=0}\Rightarrow\boldsymbol{x=(I-A^-A)z}$ ($\boldsymbol{z}$は任意のn次元ベクトル)

(iv)　$\boldsymbol{Ax=b}\Rightarrow\boldsymbol{x=A^-b+(I-A^-A)z}$

(v)　$\boldsymbol{A(A'A)^-A'A=A}$

(vi)　$\boldsymbol{A(A'A)^-A'}$　は対称行列

一般に与えられた行列 $\boldsymbol{A}$ が正則な正方行列でないと(a)を満たす一般化逆行列は1通りに定まらないので，(a)の他に下記の3つの条件をつけることがある．

(b)　$\boldsymbol{A^-AA^-=A^-}$

(c)　$(\boldsymbol{AA^-})'=\boldsymbol{AA^-}$

(d)　$(\boldsymbol{A^-A})'=(\boldsymbol{A^-A})$

(a)と(b)を満たす一般化逆行列は，反射型一般化逆行列と呼ばれ，これを $\boldsymbol{A_r^-}$ とすると次式が成立する．

(vii)　$\mathrm{rank}(\boldsymbol{A_r^-})=\mathrm{rank}(\boldsymbol{A})$

また(a)と(c)を満たす一般化逆行列は最小二乗型一般化逆行列と呼ばれ，これを $\boldsymbol{A_l^-}$ で表すと

(viii)　$\boldsymbol{A'AA_l^-=A'}$，$\boldsymbol{AA_l^-=A(A'A)^-A'}$

が成立する．

次に(a)と(d)を満たす一般化逆行列はノルム最小型一般化逆行列と呼ばれ，これを $\boldsymbol{A_m^-}$ で表すと

(ix)　$\boldsymbol{A_m^-AA'=A'}$，$\boldsymbol{A_m^-A=A'(AA')^-A}$

が成立する．そして，(a)，(b)，(c)，(d)をみたす一般化逆行列はムーア・ペンローズ逆行列と呼ばれるもので，これを $\boldsymbol{A^+}$ と書くと

$$\boldsymbol{A^+=A'(AA')^-A(A'A)^-A'}$$

$\mathrm{rank}\boldsymbol{A}=m$ ならば，

$$\boldsymbol{A^+=(A'A)^{-1}A'}$$

$\mathrm{rank}\boldsymbol{A}=n$ ならば，

$$\boldsymbol{A^+=A'(AA')^{-1}}$$

となり，与えられた行列 $\boldsymbol{A}$ に対して，一意的に逆行列が定められる．たとえば，

$$\boldsymbol{A}=\begin{bmatrix}0 & 1 & 2\\ 1 & 1 & -1\end{bmatrix}\text{ とおくと，}$$

$$\boldsymbol{A^-}=\begin{bmatrix} a & 1+3b\\ (1-2a)/3 & -2b\\ (1+a)/3 & b\end{bmatrix}$$

(a, b は任意の定数)

さらに，ムーア・ペンローズ逆行列は上式で，$a=\frac{1}{14}$，$b=-\frac{3}{14}$ とおいて得られる

$$\boldsymbol{A^+}=\begin{bmatrix}\frac{1}{14} & \frac{5}{14}\\ \frac{2}{7} & \frac{3}{7}\\ \frac{5}{14} & -\frac{3}{14}\end{bmatrix}$$

となる．(伊理正夫・韓太舜，1977；中谷和夫，1977；*Rao, C. R. & Mitra, S. K., 1971；柳井晴夫・竹内啓，1983)　〔柳井〕

一般化決定係数〔generalized coefficient of determination〕　同一個体について測定された2組の変数群 $(x_1,x_2,\cdots,x_p)$，$(y_1,y_2,\cdots,y_q)$ の全体としての関連の程度を示す指標．上記の2つの変数群を行列 $\boldsymbol{X}$, $\boldsymbol{Y}$ で表し，$\boldsymbol{X}$, $\boldsymbol{Y}$ によって構成される正射影行列を $\boldsymbol{P_X=X(X'X)^-X'}$，$\boldsymbol{P_Y=Y(Y'Y)^-Y'}$ とおくと，一般化決

定係数は,

$$d^2_{X\cdot Y}=\mathrm{tr}(\boldsymbol{P}_X\boldsymbol{P}_Y)/m$$

によって表される．ただし，m は rank($\boldsymbol{X}$) と rank($\boldsymbol{Y}$) のうち小さい方の数を表し，$(\boldsymbol{X}'\boldsymbol{X})^-$ は $\boldsymbol{X}'\boldsymbol{X}$ の†一般化逆行列を表す．$d^2_{X\cdot Y}$ は，$\boldsymbol{X}$ と $\boldsymbol{Y}$ に含まれる変数が相互に独立な場合には，$d^2_{X\cdot Y}=0$，$\boldsymbol{X}$ または $\boldsymbol{Y}$ の一方の変数の組が他方の変数の組によって完全に説明される場合には，$d^2_{X\cdot Y}=1$ となる．

なお，$\boldsymbol{X},\boldsymbol{Y}$ に含まれるそれぞれの変数の平均がゼロの場合，$d^2_{X\cdot Y}$ は†正準相関係数の平方の平均に一致する．さらに $\boldsymbol{Y}$ に含まれる変数が1つの場合は†重相関係数 $R_{y\cdot X}$ の平方に，$\boldsymbol{X},\boldsymbol{Y}$ に含まれる変数がともに1つの場合は†相関係数の平方 r^2_{xy} となる．このように $d^2_{X\cdot Y}$ は†決定係数（または多重決定係数）の最も一般化された形といってよい．さらに，$\boldsymbol{X}$, $\boldsymbol{Y}$ が名義尺度の†ダミー変数行列の場合，$d^2_{X\cdot Y}$ は†クラメールの連関係数に一致し，一方の変数群がダミー変数，他方の変数群が1つの連続変数だけからなる場合には，†相関比の平方に一致する．（竹内啓・柳井晴夫, 1972；柳井晴夫・高根芳雄, 1977）〔柳井〕

一般化最小二乗推定量〔generalized least squares estimator〕 †回帰分析等において，誤差項間に相関があるとき，†最小二乗推定量の代りにしばしば誤差間の関係を考慮した一般化最小二乗推定量（略して GLS 推定量）が用いられる．すなわち，線型モデル

$$\boldsymbol{y}=\boldsymbol{X}\beta+\boldsymbol{e},\quad E(\boldsymbol{e})=\boldsymbol{0}$$

において，$\boldsymbol{e}$ の分散共分散行列を

$$V(\boldsymbol{e})=\boldsymbol{G}$$

としよう．ただし，$\boldsymbol{y}$ は $n\times1$ の従属変数ベクトル，$\boldsymbol{X}$ は $n\times q$ の階数 q の定数行列，β は $q\times1$ の未知母数ベクトル，$\boldsymbol{e}$ は $n\times1$ の誤差ベクトル，$\boldsymbol{G}$ は $n\times n$ の既知の正値行列（→正値）とする．このとき，β の最小二乗推定量 $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ は，β の不偏推定量ではあるが，†最良不偏推定量とはならない．そこで，$\boldsymbol{G}$ の逆行列で重みづけた推定量

$$\hat{\beta}=(\boldsymbol{X}'\boldsymbol{G}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{G}^{-1}\boldsymbol{y}$$

を考えるならば，この $\hat{\beta}$ は β の†最良線型不偏推定量（BLUE）となり，最小二乗推定量よりも好ましいということになる．この $\hat{\beta}$ を一般化最小二乗推定量と呼ぶ．（佐和隆光, 1970）〔渡部〕

一般化最小二乗法〔generalized least squares method〕 線型回帰式（→線型回帰分析）において，誤差項が独立でない場合に，回帰パラメタを推定する方法．すなわち，線型回帰式 $\boldsymbol{y}=\boldsymbol{X}\beta+\boldsymbol{e}$ の誤差項 $\boldsymbol{e}$ の分散共分散行列が $V(\boldsymbol{e})=\boldsymbol{G}$〔$\boldsymbol{G}$ は正値行列（→正値）〕の場合 $\boldsymbol{G}=\boldsymbol{T}\boldsymbol{T}'$ として $\boldsymbol{T}\boldsymbol{y}=\boldsymbol{T}\boldsymbol{X}\beta+\boldsymbol{T}\boldsymbol{e}$ に最小二乗法を適用すると，結局，

$$(\boldsymbol{y}-\boldsymbol{X}\beta)'\boldsymbol{G}^{-1}(\boldsymbol{y}-\boldsymbol{X}\beta)$$

を最小にする β を求めればよいことになる．上記の方式によって β の推定値を求める方式を一般化最小二乗法または重み付最小二乗法という．

なお，一般化最小二乗法は，ベクトル幾何学的には，基準変数ベクトル $\boldsymbol{y}$ を，説明変数ベクトルの組 $\boldsymbol{X}$ で張られる空間に斜交射影することに相当するもので，この考え方を用いると，誤差項の共分散行列 $\boldsymbol{G}$ が正値でない場合にも β の推定量を求めることが可能となる．（*Rao, C. R., 1973；佐和隆光, 1979；Takeuch, K. *et al.*, 1982）〔柳井〕

一般化自然共役事前分布〔generalized natural conjugate prior distribution〕 †ベイズ統計学において事前情報を分布として表す際に，†自然共役事前分布では制約が強すぎる場合がある．このとき自然共役事前分布の族を拡張して数学的な便利さはやや低下するが事前情報を表すのにより柔軟な事前分布の族を構成し，それを事前分布の分布族として指定することがある．そのような事前分布を一般化自然共役事前分布と呼ぶ．たとえば，†モデル分布が†行列正規分布のときに用いられる．（Press, S. J., 1972）〔渡部〕

一般化分散〔generalized variance〕 2つ以上の変数の測定値に関するバラツキを示す指標．n 組の2次元データを示す点 $P_1(x_1,y_1)$, $P_2(x_2,y_2),\cdots,P_n(x_n,y_n)$ が与えられているとき，これらの n 個の点の全体の重心を $\bar{P}(\bar{x},\bar{y})$ と定義する．そして3点 P_i と P_j $(i\neq j)$ と $\bar{P}$ を結んで定義されるすべての三角形 $P_iP_j\bar{P}$ の

面積の2乗の平均値を求めると，2つの変数 x, y の一般化分散は

$$\begin{vmatrix} s_x^2 & s_{xy} \\ s_{yx} & s_y^2 \end{vmatrix} = s_x^2 s_y^2 (1-r_{xy}^2)$$

となる．ただし s_x^2, s_y^2 はそれぞれの分散，s_{xy} は共分散，r_{xy} は相関係数を表す．これを変数が $\boldsymbol{X}=(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p)$ の p 個ある場合に一般化すると，一般化分散は，p 個の変数ベクトル $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p$ の張る p 次元の平行体(平行四辺形の概念を p 次元に拡張したもの)の体積の2乗，すなわち，共分散行列 $\boldsymbol{C}_{XX}$ の行列式の値 $|\boldsymbol{C}_{XX}|$ に一致する．したがって，p 個のベクトルが一次従属である場合には一般化分散は0，p 個の変数がすべて相関のない場合には，一般化分散の値は p 個の変数の分散の積になる．(Anderson, T.W., 1958；林知己夫他, 1970; Tatsuoka, M.M., 1971)　〔柳井〕

移動平均〔moving average〕 時系列データから偶然誤差による変動を除き，時間 $t_j (j=1, \cdots, n)$ に関する滑らかな†トレンドを得るための方法．トレンドが，t_{i-k} から t_{i+k} の区間で線型であると仮定できるとき，移動平均 m_i は

$$m_i = \frac{1}{2k+1} \sum_{j=i-k}^{i+k} x_j$$

によって与えられる．ただし，x_j は時間 t_j における観測値である．また，トレンドが t_{i-k} から t_{i+k} の区間で非線型のときは加重移動平均

$$m_i = \sum_{j=-k}^{k} a_{j+k+1} x_{i+j}$$

が用いられる．a_j は $\sum_{j=1}^{2k+1} a_j = 1$ の制約に従う重みで，多項式模型などによってあらかじめ決定される．なお，移動平均を計算するためには，$2k+1$ 個の観測が必要なので，最初の k 個の観測と最後の k 個の観測を除いた $k+1 \leqq i \leqq n-k$ についてのみ計算できる．なお，多項式模型による a_j の決定や移動平均 m_i の分散や信頼限界についても研究されている．(*Brandt, S., 1976)　〔石塚〕

イメージ分析〔image analysis〕 †因子分析モデルに準じ，変数のもつ共通変動分と独自な変動分とを分析するモデル．データをとった個体を，ある母集団からの標本とみなすことはふつうおこなわれるが，変数についても，そこで取り上げられている変数は，もっと多くの変数の総体からの標本であるとみなす．そして，ある1つの変数の変動のうち，他の変数の総体を用いた多重線型回帰(→回帰分析)によって予測される部分が，その変数のもつ共通部分であり，これをその変数のイメージと呼ぶ．そして，予測されない残差の部分を反イメージと呼ぶ．現実のデータでは，変数の総体が用いられることはなく，そのうちの一部として n 個の変数が用いられる．したがって，このイメージを求めることはできない．そのかわりに，ある変数を残りの $n-1$ 個の変数による多重線型回帰で予測する．こうして求められるイメージ，反イメージを，変数の総体を用いたものと区別して，部分イメージ，部分反イメージなどと呼ぶことがある．イメージ分析の基本的な分解は

$$\boldsymbol{R} = \boldsymbol{G} - \boldsymbol{S}^2 \boldsymbol{R}^{-1} \boldsymbol{S}^2 + 2\boldsymbol{S}^2$$

によって表される．ここで $\boldsymbol{R}$ は n 変数間の相関行列を，$\boldsymbol{G}$ は n 変数のイメージの共分散行列を表す．また，対角行列 $\boldsymbol{S}^2 = (\mathrm{diag}\boldsymbol{R}^{-1})^{-1}$ の各成分 s_j^2 を1から引いた $1-s_j^2$ が，変数 j と他の $n-1$ 個の変数との重相関係数の平方(いわゆる †SMC)に等しい．変数の母集団全体について分析したときには，イメージ共分散行列は，対角成分に共通性を入れた相関行列に等しくなる．

イメージ分析の近似として，部分イメージの共分散行列を用いて，ふつうの因子分析をすることができるが，この場合には共通性の推定を改めておこなう必要がない．イメージ分析に関連したものとして，イメージ因子分析と呼ばれるモデルがある．これは

$$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{A}' + \theta \boldsymbol{S}^2$$

という基本式で表される．ここで，$\boldsymbol{A}$ は因子負荷行列を，θ は $\boldsymbol{R}$ によって定まる定数である．(Mulaik, S. A., 1972；芝祐順, 1979)　〔芝〕

因果分析〔causal analysis〕 事象間に因果関係を想定したモデルの妥当性を検証し，その関係の強弱を明らかにする方法．

2つの変数 X と Y の間に因果関係が存在すると言い得るためには，次の3つの条件が必要であるといわれている．すなわち，

(1) X と Y の間に共変動が存在する．
(2) X と Y の間に時間的順序を特定できる．
(3) X，Y 双方に影響を与える変数の影響を取り除いても，X と Y の共変動は0とならない，すなわち，X, Y 以外のすべての変数の集合を Ω として，$V_{XY.\Omega} \neq 0$.

因果分析では，これら3つの条件を手がかりとして，モデルを設定し，また，検証する．(1)および(2)は，主としてモデルを設定，または，改訂するときの手がかりとして用いられ，(3)は主として，モデルを検証するための手段となる．しかし，実際問題として，X と Y 以外のすべての変数を考慮することは不可能である．そこで，X と Y に関係が深いと考えられるいくつかの変数（Ω の部分集合）を選んでモデルを構成する．

因果分析のモデルは，主として，因果関係を矢印で示した†アロー ダイアグラムの形で与えられる．また，変数が規準化されているとき，因果関係の強さの程度を与えるパラメタをパス係数と呼ぶ．モデルの検証とモデル パラメタの推定は，具体的には，†パス解析の方法や†多重指標法，†共分散構造分析の方法などによっておこなわれる．モデルの検証を通じて，いくつかの変数がモデルから取り除かれたり，他にモデルに加えるべき変数の存在することが明らかになったりする．また，このように，モデルを改訂すべき証拠のないときは，新しい証拠の現れるまで現在のモデルを使って，パラメタを推定することができる．(*Asher, H. B., 1976; Blalock, H. M. Jr. ed., 1971)　〔石塚〕

因子スコア〔factor score〕　因子分析モデルの中では，因子と呼ばれる潜在変数が用いられるが，この変数の値をさして因子スコアという．実際のデータから得られる因子解としては，ふつう因子負荷が数値的に与えられるが，因子スコアの値は他の未知数（独自性のスコア）のため，確定することができない．因子スコアの推定のために，いろいろな推定式が提案されている．まず，n 個の変数への各因子の回帰によって因子スコアを推定する方法では，因子スコアの推定値の行列 $\hat{F}$ は $\hat{F}=ZR^{-1}A$ によって与えられる．ここで，Z は標準化されたデータ行列，R は変数間相関行列，A は因子負荷行列を表す．これは真の因子スコアへの最小二乗近似による推定値でもある．また，独自因子を最小化することによって求めた因子スコアの推定値は $\hat{F}=ZD^{-2}A(A'D^{-2}RD^{-2}A)^{-\frac{1}{2}}$ である．ただし，D^2 は独自性を成分とする対角行列である．さらに，共通因子成分に関する最小二乗的方法による推定値は $\hat{F}=ZR^{-1}(R-D^2)A(A'A)^{-1}$ によって与えられる．このほかにも多数の推定式があるが，それらのいずれが推定式として最も望ましいものであるかは一義的に評価できない．→因子分析　〔芝〕

因子スコアの不定性〔indeterminacy of factor score〕　ふつう†因子分析においては，†因子スコアを直接求めるのではなく，因子と観測された変数との相関係数，すなわち，因子負荷によって因子を間接的に定める．この場合に，因子分析モデルでは，因子スコアの値を確定的に求めることができない．このことを因子スコアの不定性と呼ぶ．

ふつう，因子スコアは，推定値として求められる．ここで，推定値というのは，いわゆる母集団値の推定を意味するものではなく，数学的な意味で，与えられたデータから因子スコアを確定することができないということを意味している．いま，n 個の変数について，m 個の共通因子が得られたとすると，このモデルは m 個の共通因子と n 個の独自因子とを合わせた $m+n$ 次元のベクトル空間によって説明される．そして n 個の変数はこの空間の中での n 次元の部分空間を構成する．したがって，これら n 個の変数の線型結合は，すべてこの部分空間の中にあり，これによって $m+n$ 次元空間の中の共通因子や独自因子を完全に表すことはできない．

因子の不定性の程度を表す指標として，各因子と n 個の変数との重相関係数の平方が用いられる．R を n 変数間相関行列とし A を因子負荷行列（斜交モデルの場合は因子構造行列）とすると，この重相関係数の平方は $A'R^{-1}A$ の対角成分として与えられる．(Mulaik, S. A., 1972)　〔芝〕

因子の不変性〔factorial invariance〕 †因子分析モデルを用いて，実際のデータから求めた因子負荷によって因子を説明するとき，それがどこまで一般性をもったものと考えてよいかが問題となる．これを標本理論の枠組で考えれば，標本分布や，標本抽出の問題となるが，そのほかに，個体の一部が除かれた場合，あるいは，変数の一部が除かれたり，新たに加えられたりした場合に，それが因子解にどのような変化をもたらすかは，結果の一般化の上で重要な問題である．もし，標本の差異や変数の選択にもかかわらず，同じ因子が現れるということであれば，一般化が可能となる．このような傾向を因子の不変性と呼んで，そのための条件や，その程度について検討されてきた．(Gorsuch, R.L., 1974; Mulaik, S.A., 1972)　〔芝〕

因子分解定理〔factorization theorem〕
⇒ネイマン・フィッシャーの因子分解定理

因子分析〔factor analysis〕 一組の多変数データを，比較的少数の共通な †潜在変数の線型結合として，集約的に表すことを目的としたモデルである．個体 i についての，変数 j の観測値 x_{ij} を，変数ごとに標準化したものを z_{ij} で表すと，因子分析モデルの基本式は

$$z_{ij}=a_{j1}f_{i1}+a_{j2}f_{i2}+\cdots+a_{jm}f_{im}+e_{ij}$$

となる．ここで，$f_{i1}, f_{i2}, \cdots, f_{im}$ などは m 個の潜在変数を表し，共通因子スコア，あるいは，共通因子と呼ばれる．$a_{j1}, a_{j2}, \cdots, a_{jm}$ はその潜在変数にかかる重み係数で，因子負荷と呼ばれる．e_{ij} はこのような共通因子で説明しきれない誤差分を表す．上式を行列の記法によって表すと $\boldsymbol{Z}=\boldsymbol{FA}'+\boldsymbol{UD}$ となる．$\boldsymbol{Z}$ は $N\times n$ のデータ行列を，$\boldsymbol{F}$ は $N\times m$ の共通因子行列を，$\boldsymbol{A}$ は $n\times m$ の因子負荷行列を表す．残りの誤差分は e_{ij} をそのまま行列で表してもよいが，説明の便宜上，$N\times n$ の独自因子スコアの行列 $\boldsymbol{U}$ と，$n\times n$ の対角行列 $\boldsymbol{D}$（この行列の各対角成分を平方したものを独自性と呼ぶ）の積によって表す．因子分析の課題は所与のデータの組に対し，これを説明するための潜在変数として，いくつの共通因子を設けたらよいか，どのような変数を選ぶのがよいか，などである．

因子分析の解を求めるということは，因子負荷を求めることを意味することが多い．それは，因子負荷が因子の特徴を集約的に表しているからである．因子解は，因子が互いに直交するように定められる直交解と，直交の制約を課さない斜交解とがある．多変数にみられる変動を集約的に表す潜在変数として，直交する因子を設けることは，簡明で理解しやすいという利点がある．しかし，ときには直交する因子によるよりも，斜交する因子によって多変数の関係を説明する方が自然なこともある．直交解の場合には因子負荷は基本モデルにおける重みを表すと同時に，因子と変数との相関をも表す．すなわち，重み係数 a_{jp} はまた，因子 p と変数 j との相関係数でもある．しかし，斜交因子解では，重み係数は必ずしも因子と変数との相関係数に一致しないので，因子負荷と呼ぶかわりに，重み係数の方を因子パタンと呼び，相関係数の方を因子構造と呼ぶ．

ふつう因子解を求めるには，変数間相関行列が用いられる．ただし，その対角成分には，†共通性と呼ばれる値が入れられる．これは，基本式から $\boldsymbol{R}=\boldsymbol{AA}'+\boldsymbol{D}^2$ という関係が得られることによる．斜交解の場合には，$\boldsymbol{R}=\boldsymbol{ALA}'+\boldsymbol{D}^2$ となる．ここで $\boldsymbol{L}$ は因子間の相関行列である．

因子解としてもっとも基本的なものは主因子解である．これは，データ行列に含まれる共通変動の中で，もっとも顕著な変動傾向を表す因子をまず定め，以下，これに直交し，しかも残された変動をもっともよく表すような因子を，順次定めていく解法である．この解は，共通性を対角成分にもつ相関行列 $\boldsymbol{R}^+=\boldsymbol{R}-\boldsymbol{D}^2$ の †固有分解によって得られる．すなわち，主因子解の因子負荷は $\boldsymbol{R}^+\boldsymbol{A}=\boldsymbol{A\Lambda}$ によって表される．$\boldsymbol{\Lambda}$ は $\boldsymbol{R}^+$ の固有値を成分とする対角行列である．このように，相関行列から直接求められる解を直接解と呼ぶが，直接解としては，このほか，セントロイド解（→セントロイド法），群因子解，直接バリマックス解（→バリマックス法）などがある．

直接解によって定められた因子を線型変換することによって，より適切な因子を求めることがある．このようにして得られる解を変換解と呼ぶ．より適切な解とは†単純構造を目指すものや，特定の因子構造によく近似されるようにするものを指す．単純構造を目指す方法として提案されているものには，†バリマックス法，†ジェオマックス法，†オーソマックス法などの直交解，†オブリマックス法，†オブリミン法，†プロマックス法，マックスプレイン法，斜交ジェオマックス法などの斜交解がある．特定の因子構造に近似する方法としては，†プロクラステス法（直交あるいは斜交）がある．これは仮定された因子構造（あるいは因子パタン）に最小二乗法的にもっともよく近似するように解を求めるものである．

因子負荷を求めることによって目的を達することもあるが，場合によっては，因子スコア $\boldsymbol{F}$ を推定することもある．因子負荷が定められたとしても，独自性スコアが未知であるため，因子スコアを確定することはできない．因子スコアは，因子負荷とデータ行列とから，なんらかの方法，たとえば回帰などによって推定される．推定式としてはいろいろ提案されているが，その例を挙げると $\hat{\boldsymbol{F}}=\boldsymbol{ZR}^{-1}\boldsymbol{A}$, $\hat{\boldsymbol{F}}=\boldsymbol{ZA}(\boldsymbol{A}'\boldsymbol{A})^{-1}$, $\hat{\boldsymbol{F}}=\boldsymbol{ZR}^{-1}(\boldsymbol{R}-\boldsymbol{D}^2)\boldsymbol{A}(\boldsymbol{A}'\boldsymbol{A})^{-1}$ などである．

因子分析モデルには，このほか異なった特徴をもつものもある．与えられたデータを，母集団からの無作為標本とみなし，因子分析モデルに標本理論の考えを導入することもある．この場合には，因子分析の基本式は $\boldsymbol{x}=\boldsymbol{fA}'+\boldsymbol{uD}$ と表される．ここで $\boldsymbol{x}$ は n 個の確率変数 $x_1, x_2, \cdots, x_n$ を成分とする確率ベクトル（それらの期待値は0とする），$\boldsymbol{f}$ は m 個の確率変数 $f_1, f_2, \cdots, f_m$ を成分とする確率ベクトル，$\boldsymbol{u}$ は n 個の確率変数 $u_1, u_2, \cdots, u_n$ を成分とする確率ベクトルを表す．ただし $\boldsymbol{x}$, $\boldsymbol{f}$, $\boldsymbol{u}$ はいずれも横ベクトルを表す．このようなモデルを設け，また因子の確率分布についての仮定を設けることによって，因子負荷の最尤推定をおこなうことができる．

また，変数の方について，その母集団を想定することによって，†イメージ分析のモデルが論じられている．(*Comrey, A. L., 1973; Gorsuch, R. L., 1974; Harman, H. H., 1976; *Horst, P., 1965; (*)Lawley, D. N. & Maxwell, A. E., 1971; Mulaik, S. A., 1972; 芝祐順, 1979)　〔芝〕

インドスケール〔INDSCAL〕 individual differences scaling の略．†個人差多次元尺度法の1つ．また，そのモデルとアルゴリズムを用いて書かれたコンピュータプログラムを指すこともある．→多次元尺度法　〔石塚〕

う

ウィシャート分布〔Wishart distribution〕 $p\times p$ 確率変数行列 $\boldsymbol{W}$ は，対称でかつ正値とする．このとき，$\boldsymbol{W}$ の要素の同時分布が連続的で密度

$$f(\boldsymbol{W})=\frac{c|\boldsymbol{W}|^{(n-p-1)/2}}{|\boldsymbol{\Sigma}|^{n/2}}\exp\left(-\frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right),$$
$$\boldsymbol{W}>0,\ \boldsymbol{\Sigma}>0,\ p\leqq n$$

をもつならば，$\boldsymbol{W}$ は尺度行列 $\boldsymbol{\Sigma}$ と自由度 n の非特異な p 次元ウィシャート分布に従うという．ただし，c は定数で

$$c=\left[2^{np/2}\pi^{p(p-1)/4}\prod_{j=1}^{p}\Gamma\left(\frac{n+1-j}{2}\right)\right]^{-1}$$

である．$\boldsymbol{W}>0$ および $\boldsymbol{\Sigma}>0$ は，各々の行列が正値であることを意味し，そうでない場合には，$f(\boldsymbol{W})=0$ である．また $n<p$ のときには，分布は特異(singular)となり，密度は存在しない．したがって，$\boldsymbol{W}=\{w_{ij}\}$, $\boldsymbol{\Sigma}^{-1}=\{\sigma^{ij}\}$, $i, j=1, \cdots, p$ とすれば，$\boldsymbol{W}>0$ に対して，

$$f(\boldsymbol{W})\propto\frac{|\boldsymbol{W}|^{(n-p-1)/2}}{|\boldsymbol{\Sigma}|^{n/2}}$$
$$\times\exp\left(-\frac{1}{2}\sum_{i=1}^{p}\sum_{j=1}^{p}w_{ij}\sigma^{ij}\right)$$

となる．ただし，$\propto$ は比例することを表す．

ウィシャート分布は，正規母集団からの標本分散共分散の同時分布であり，$p=1$ のときには，W の分布はカイ二乗変数の分布に比例する．すなわち，ウィシャート分布は，†カイ二乗分布（または†ガンマ分布）のp次元への拡張とみなすことができる．(Anderson, T. W., 1958)　〔渡部〕

ウィルクスのラムダ〔Wilks' lambda〕 k 組の p 変量標本の平均ベクトルの差を検定するときなどに用いられる統計量で，$\Lambda=|N\hat{\boldsymbol{\Sigma}}_\Omega|/|N\hat{\boldsymbol{\Sigma}}_\omega|$ によって定義される．ここで $\hat{\boldsymbol{\Sigma}}_\omega$ は帰無仮説のもとでの分散共分散行列 $\boldsymbol{\Sigma}$ の最尤推定量，$\hat{\boldsymbol{\Sigma}}_\Omega$ は仮説を設けない場合の $\boldsymbol{\Sigma}$ の最尤推定量である．$p=1$ のとき，Λ と F 統計量との間には，全標本数を N として，

$$\Lambda=\frac{1}{1+[(k-1)/(N-k)]F}$$

の関係がある．また，$k=2$ のとき，†ホテリングの T^2 との間には

$$\frac{1-\Lambda}{\Lambda}=\frac{T^2}{N-2}$$

の関係がある．Λ はまた，アンダーソンの U とも呼ばれる．→アンダーソンのU，平均ベクトルに関する尤度比検定 (*Rao, C. R., 1973)　〔石塚〕

ウィルコクスンの検定〔Wilcoxon's rank sum test〕 順位にもとづく†ノンパラメトリック検定の1つで，母集団分布の同一性の仮説を検定するもの．⇒順位和検定　〔芝〕

ウィンザライズ平均〔Winsorized mean〕 大きさ n の標本を小さいものから順に並べた順序統計量を $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ とする．これらのうち，はじめの g 個（$X_{(1)}$ から $X_{(g)}$ まで）と，あとの g 個（$X_{(n-g+1)}$ から X_n まで）を，それぞれ $X_{(g+1)}$ と $X_{(n-g)}$ とでおきかえる．すなわち，$X_{(g+1)}, X_{(g+1)}, \cdots, X_{(g+1)}, X_{(g+2)}, \cdots, X_{(n-g-1)}, X_{(n-g)}, \cdots, X_{(n-g)}, X_{(n-g)}$ としたものを α ウィンザライズ標本と呼び，この平均を α ウィンザライズ平均という．ただし，$\alpha=g/n$ である．(Barnett, V. & Lewis, T., 1978; 竹内啓・大橋靖雄, 1981)　〔芝〕

ウェルチの検定〔Welch test〕 2つの正規母集団 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ の平均値に関する仮説 $H_0: \mu_1=\mu_2$ を，等分散 $\sigma_1^2=\sigma_2^2$ の仮定なしに検定する方法．その中でも比較的簡単な近似法は，次のようにしておこなわれる．それぞれの母集団から大きさ n_1, n_2 の標本をとり，その平均を $\bar{x}_1, \bar{x}_2$，不偏分散を s_1^2, s_2^2 とする．これらを用いた統計量

$$t=\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$$

を考えると，これは仮説 H_0 のもとで，自由度 ν

$$\nu=\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2\Big/\left(\frac{s_1^4}{n_1^2(n_1-1)}+\frac{s_2^4}{n_2^2(n_2-1)}\right)$$

のティ分布に近似的に従う．このことを用いて，平均の差に関する検定をおこなうことができる．(竹内啓・大橋靖雄, 1981)　〔芝〕

ウォルシュの検定〔Walsh test〕 順位にもとづく†ノンパラメトリック検定の1つで，対になった2群の標本に関し，対の差 $d=y-x$ の母集団分布の中央値が0である，という仮説を検定するもの．まず，データ (x_1, y_1), $(x_2, y_2), \cdots, (x_n, y_n)$ のすべてについて，各対の差 $d_i=y_i-x_i$ を求める．これらを小さいものから順に並べる．添字をつけかえて $d_{(1)} \leqq d_{(2)} \leqq \cdots \leqq d_{(n)}$ とし，これらが検定に用いられる統計量となる．これらの一部の正負により仮説を検定するための数表が用意されている．(岩原信九郎, 1964)　〔芝〕

え

エーアイシー〔AIC〕 Akaike's information criterion の略．⇒赤池情報量基準