統計・OR活用事典

編者●森村英典
牧野都治
真壁　肇
杉山高一

東京書籍

序

本書はビジネスマン，学生等で，業務や勉学の上で統計や OR の用語に出合い，その大略の概念や使い方を知りたい人を読者対象と考えている。

そのような要望が生じたとき，普通の辞書では，たとえその言葉が見つかったとしても，ほとんど字面だけの解釈に終わってしまうであろう。といって，ハンドブック形式の大辞典では読むのが大変で，どこまで読めばいいのかという点でも戸惑いを覚えることが多い。また，一体どんなことができ，どんなことは難しいのかということも分かりにくい。

それでは，どのようなスタイルの事典がよいのか。多分，ある程度のことが比較的短い分量で書かれていて，しかも数値を用いた例などで使い方の方向がうかがえるといったものが考えられよう。また，表現はできるだけやさしくして，誰でも一応の見当はつく，といったものが望ましいであろう。

そこで考えられたのが，見開き 2 ページで 1 項目というスタイルである。この程度の分量であれば，分からない用語に出合ったとき，とにかく読んでみようと思うであろう。それでも多いと感ずる人のためには，1 ページ，更には数行でもそれなりの知識は得られるという形をとれば，なお使いやすいであろう。

本書では，見開き 2 ページのうち，左ページの上の部分で「辞書的解説」，つまり見出し語のごく基本的な，特に字義的な解説をする。「国語大辞典」で扱う程度の解説であって，まず基本的なイメージを持っていただくためのものである。そして，左ページの下の部分ではその見出し語の概念や関連のある他の用語などについてやや詳しい解説を行っている。したがって，このページを読み終われば，たいていの項目では一応の理解は得ていただけると期待している。なお，ここで右ページのための準備としての解説を行うこともある。

しかし，本書の特色は「活用」をうたった右ページにある。ここでは，なるべく例題による計算法やグラフ・表による解説によって，使い方を知っていただくことを目指している。このページはもちろん原則的には 1 ページであるが，ごく基本的な手法でどうしても解説が長くなるものについては，例外的に 3 ページとしている。

このような意図で書かれているので，各項目を読み終わった段階で，何とかそれについてのイメージはつかめることが期待されるが，それを本当に使うためには，もっと深く勉強する必要が生ずるかもしれない。そのときのために，適切な参考書を項目別にあげるように努めた。ただ，行数の限られている場合は，巻末にあげた参考書のリストを見ていただく必要があろう。

本書の執筆に当たっては，できるだけやさしく書くことに留意した。このような事典の執筆態度には，理論的によく武装して，「難解は厭わず過誤を戒める」という態度をとることがままあるように感じられるが，これとは全く逆の書き方を意図している。モットーはむしろ「誤解を恐れず徹底的にやさしく」ということであった。数学については特に注意し，たとえば$\sum$の使用はなるべく避け，＋だけで表記するように心掛けるなどに配慮をしたつもりである。とはいっても，やはり行数の限られている場合には，$\sum$を使用している。これは1項目2ページという読みやすさとジレンマであった。

配列は，通常の辞書と同様に見出し語の50音順としている。これは，2ページで一応読み終わることを意図したためでもある。このため，関連項目のページを知りたいことも多く起こるであろう。左ページの右下の余白は，この目的のために設けられている。そこには，関連項目のページが大変見やすい形で示されている。

また，索引に1行解説を付したのも新しい試みである。これは，本書の見出し語では単語まではカバーしきれないため，それらを探すときの便利さをねらって考えられた。読者は，まず索引で調べたい語句を探し，その段階である程度の理解を持ってから，本書の該当個所を探せばよい。同種の語句の中で調べたい語句を特定するためにも，その語句に対して抱いていた理解を確認するためにも，この索引は役立つに違いない。

事典や辞書類の企画にとって，項目の選定はかなり重要な仕事である。しかし，項目の選定をする人と執筆をする人との間の意志疎通は概して困難であって，必要な項目が脱落することもあり得るように思われる。それを避けるため，本書では，項目の選定を「木」の形式で行い，執筆をしながら必要と思われた項目は，樹木が枝を伸ばすように順次追加していくようにしている。このため，解説に欠落が生じるという虞は減ったと思っている。たとえば，中項目主義のために大項目

に当たる語句の解説が抜けてしまうという事典も散見するが，本書ではそのような事態を避けることができた。

ところで，項目の選定に当たっては，本書の項目から見て，ビジネスマンの出合いそうな語はなるべく入れるように心掛けている。2ページ1項目というスタイルのため，それらの語句をすべて見出し語とすることは必ずしもできないが，多項目の解説の流れの中で適宜とりあげており，しかも上述のような索引の充実も考えているので，読者は容易に所用の語句に到達でき，あまり不便は感じられないであろう。また，学術的な意味での程度を高くすることは毛頭考えていないので，やや特殊と思われる概念や手法までとりあげることはむしろ避けている。しかし，それはむしろ使いやすさを重視したためであり，本書で見つからなかった用語は，やや特殊なものと思っていただいても大きな誤りにはならず，実際上の不便はあまり起こらないであろう。

とはいっても，時代の進展に伴って，ビジネスマンが出合う語句そのものにも変化があるから，現在の取捨の判断が必ずしも適切であり続けるとは思われない。また各執筆者の感覚の微妙な差から，分野ごとの濃淡に差が出ることも避けにくいので，将来改訂の機会が得られれば，そのへんの補正も行いたい。幸い，本書のスタイルはそのような改訂作業を容易にしている。是非とも読者各位のご批判，ご叱正，ご意見をお願いしたい。

本書の企画は，牧野，真壁，森村，杉山の4名に立教大学の小林竜一教授を交えて立てられた。一応の企画が立ったのち，4名が部分的に執筆を行ってそれを検討しあい，項目と執筆者を「木」の形式で順次増やしながら検討を繰り返した。このようにして本書はでき上がったので，後からお願いした執筆者の方々には何かとご無理をお願いしている。それにもかかわらず，各位からは快いご協力を得られたので，比較的短い日時で本書の完成を見ることができた。

小林竜一教授，執筆者各位のご協力に対し，深甚なる謝意を表する次第である。また，東京書籍株式会社の駒田弥太郎顧問をはじめ編集担当の各位のご苦労に対してもお礼を申しあげたい。

編者一同

凡　　例

1. 本書の各項目は，「定義的解説」，「やや詳しい解説」，「活用のページ」からなる。
2. 見出し語は50音順に配列した。欧文の略称は，基本的にアルファベットの通りに読む（例：CPM《シーピーエム》，ROC曲線《アールオーシーきょくせん》）。ただし，そうでない読み方についてはルビを付した（例：PERT（パート），GERT（ガート））。
3. 英語の略称や難読文字で検索できない場合は，6ページの，「見出し語の関連表」や巻末の索引の対象ページを参考にする。
4. 1項目は見開き2ページが原則で，見開きの末尾にはその項目の執筆者の名を(　)で付記した。もしこれが見開きの末尾にないときは，例外的な「4ページもの」であり，次の2ページには「活用のページ」が続くことを意味する。
5. 索引は，和文は50音順，欧文はアルファベット順にした。また，それぞれに「1行解説」を付した。「1行解説」は字数の制約上必ずしも正確な表現ではないが，その語句の単純化した解説や，その文脈・内容などを示している（ただし，字面から明らかな場合は除く）。
6. 索引には，原則としてその語句の英訳を付してある。また，索引の対象ページは，その主な個所（太文字で示す）のほかは，その語句や概念が使われている個所に限っている。単に，文中にその語句が存在するというだけのところはあえて掲げていない。
7. 前項の原則にもかかわらず，ごく例外的に，そのような語句を索引に入れたものがある。これは，その英訳と合わせ，ごく簡単な解説を索引中で行って，索引の「簡単な辞書」としての役割を補完したいと思ったためである。
8. 参考文献は一般に入手しやすいものを掲げた。本文に余裕がある場合は見開きページの末尾に，余裕がない場合は巻末にまとめてある。巻末にある場合には，本文では〔　〕で参照番号を記した。

ギリシャ文字

	字　名	英語読み	独仏語読み
A　α	alpha	アルファ	
B　β	beta	ビータ (米：ベータ)	ベータ
Γ　γ	gamma	ガンマ	
Δ　δ	delta	デルタ	
E　ε	epsilon	エプサイロン	エプシロン
Z　ζ	zeta	ジータ	ゼータ
H　η	eta	イータ	エータ
Θ　θ	theta	シータ	テータ
I　ι	iota	アイオタ	イオタ
K　κ	kappa	カッパ	
Λ　λ	lambda	ラムダ	ランブダ（独）
M　μ	mu	ミュー	ムー（独）
N　ν	nu	ニュー	ヌー（独）
Ξ　ξ	ksi	クサイ	クシ
O　o	omicron	オーマイクロン	オミクロン
Π　π	pi	パイ	ピー
P　ρ	rho	ロー	
Σ　σ, ς	sigma	シグマ	
T　τ	tau	トー	タウ（独）
Υ　υ	upsilon	ユプサイロン	ユプシロン
Φ　ϕ, φ	phi	ファイ	フィー
X　χ	khi, chi	カイ	キー
Ψ　ψ	psi	プサイ	プシー
Ω　ω	omega	オーメガ	オメガ

見出し語関連図

（統計）

- 日本の統計 282
  - 国民経済計算 120
    - 景気動向統計 92
    - 生産統計 204
    - 貿易統計 354
  - 労働統計 414
  - 家計統計 60
  - 企業経営統計 72
- 記述統計 74
  - 度数表 274
  - 方眼紙 356 — 正規確率紙 198
  - 特性値 272
    - 積率 210
      - 平均値 344
      - 分散 332
      - 変動係数 348
    - 相関関係 222
    - 比率 316 — 指数 154
- 統計調査 258 — 確率抽出法 54
- 統計的推測 260
  - 点推定 254
  - 区間推定 82
  - 最尤（さいゆう）法 134
  - 仮説検定 62
    - 母数に関する仮説検定 362
    - 2つの母集団の母数の差の検定 324
    - 適合度検定 248
    - 分割表による検定 328
    - ノンパラメトリック検定 288
- 実験計画法 160
  - 因子 20
    - 単一因子実験 232 — 分割法 330
    - 要因実験 396 — 直交表 238
    - 要因効果 394 — 分散分析 336
  - 構造模型 110
  - 無作為化 384 — 乱塊法 402 — ラテン方格法 400

- 数理計画法 186
  - ネットワーク・モデル 284
    - 最大流問題 130
  - 線形計画法 218
    - 単体法 234
    - 輸送問題 392
  - 整数計画法 208
    - 割当問題 420
  - 非線形計画法 312
  - 動的計画法 270
    - マルコフ決定過程 380
- ゲーム理論 102
  - ゼロ和2人ゲーム 216
    - ゲームの解法 100
  - 非協力ゲーム 310
  - 協力ゲーム 80
    - 協力 *n* 人ゲームの解 78
- ＯＲの実施理論 44
  - 分析に対する行動科学的知見 340
  - 行動科学的意思決定論 112
- 発想的ＯＲ 294
  - ワーク・デザインとKJ法 416
  - 決定の木とPDPC 106
  - メリット・デメリット分析とマトリックス分析 386
  - フェース法 322
  - 不平等度の計測 326
    - パレート図 300

☆見出し語187を樹形図の形にまとめ，本書の内容概観を与えた。各項目脇の数字はページ数を示す。

☆関連は必ずしも一意的でないものもあるが，それらを示すとかえってわずらわしくなるので，割り切って，どの語も1ヵ所にしか掲げていない。

☆大きく「統計」と「ＯＲ」とに分類したが，これは本書の題名のためでもあり，分類自体に異論もありうると考えられるので，本関連図の7～8ページあたりは両者の中間もしくは周辺という意味合いも込めて点線で両者を結んでおいた。

編者・執筆者一覧

市野省三	労働省政策調査部統計調査第二課長
新村秀一	住商コンピューターサービス株式会社
■杉山高一	中央大学理工学部教授　数学科
高橋幸雄	東北大学経済学部助教授　経営学科
中野文平	東京工業大学大学院総合理工学研究科助教授 システム科学専攻
中村善太郎	慶応義塾大学理工学部助教授　管理工学科
福川忠昭	慶応義塾大学理工学部専任講師　管理工学科
■真壁　肇	東京工業大学工学部教授　経営工学科
■牧野都治	東京理科大学理工学部教授　情報科学科
宮村鐵夫	茨城大学工学部助教授　情報工学科
武藤滋夫	東北大学経済学部助教授　経営学科
森　雅夫	東京工業大学工学部助教授　経営工学科
■森村英典	東京工業大学理学部教授　情報科学科

※50音順・編者は■印

装丁　東京書籍A. D.

統計・OR活用事典

ROC 曲線 (Receiver Operating Characteristic curve)

ROC 曲線は決定行列の True Positive を縦座標とし，False Positive を横座標として描いた成績評価のための曲線である。日本語では，受信者動作特性曲線などと訳され，人間の視聴覚による信号検知能力に関する研究に起源がある。

医療の分野では，X 線写真の読影精度の評価法に用いられた。ROC 曲線を判別手法の評価に用いれば，複数の判別成績の評価が簡単に行える。さらにこれを用いれば，判別境界点の決定方式の違いによる判別成績の差異を統一的に扱える。

★解説

■判別成績の評価 判別成績は，判別境界点を決めることにより，誤分類確率(両群で誤分類された数／全体の数)または 2×2 の決定行列で表すことができる。しかし，判別境界点の決定方法として尤度比方式，ベイズ方式などといろいろ提案されており，方式ごとに誤分類確率が異なるので，どの方式を用いたらよいかの判断が難しい。そこで，判別境界点を適当にいくつかとり，それを 1 枚の図として表したものが ROC 曲線である。

■ROC 曲線 興味のある方を G_1 群，他方を G_2 群として判別分析を行った。判別境界点 d を決めれば，G_1 群が正しく G_1 群と判別される率(TPR)と，G_2 群が誤って G_1 群と誤判別される率(FPR)が定まる。 判別境界点を動かすことにより TPR を 0 % から 100 % までの間で適当に動かせば，図 1 に示す判別成績の ROC 曲線が得られる。

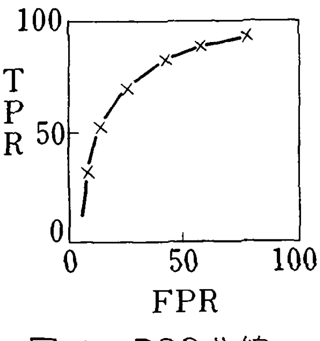


図 1　ROC曲線

関連ページ
決定行列　104
ベイズの定理を用いた判別　346
FUNCAT による判別　320
数量化Ⅱ類　190

判別点のサンプル数量	正診（True Positive）		誤診（False Positive）		判別的中率
	累積症例数	累積百分率	累積症例数	累積百分率	
7.20	7例	8.6%	0例	0.0%	96.5%
2.60	40	49.4	31	1.4	96.6 ←
1.40	47	58.0	57	2.7	95.7
0.20	60	74.1	208	9.7	89.1
0.00	63	77.8	389	18.1	80.6 ←
−0.20	70	86.4	954	44.3	54.0
−0.40	78	96.3	1722	80.1	17.7
−0.60	81	100.0	2015	93.7	3.9

表 1　電算機診断の判別点と悪性診断力（モデル2：内部標本）

■実用例　2096例の胃X線写真から，専門医が胃がんに関する所見を抽出した。これを説明変数とし，その後に行われた外科手術によるがん群と良性群の診断を外的基準として判別分析を行う。

表1は数量化Ⅱ類の判別成績を示す。サンプルスコア上で判別境界点を−0.60から7.20まで動かして，表1の成績が得られた。すなわち，判別境界点を7.20としてそれ以上をがん群とした場合，がん症例の8.6％が正しく判別され，良性例の0.0％が誤分類されたことを示す。この例では，良性群が2015例とがん群の81例に比べて極端に多い。このため，判別的中率の一番よい96.6％点をとれば，がん群の50％近くが良性と誤診されることになる。

同じデータを用いて，一般医が4段階表示(がん確診，がん疑，悪性否定しえず，良性疾患)で判定したものを医師診断とよぶことにする。この悪性度表現も一種の判別軸と考えられる。

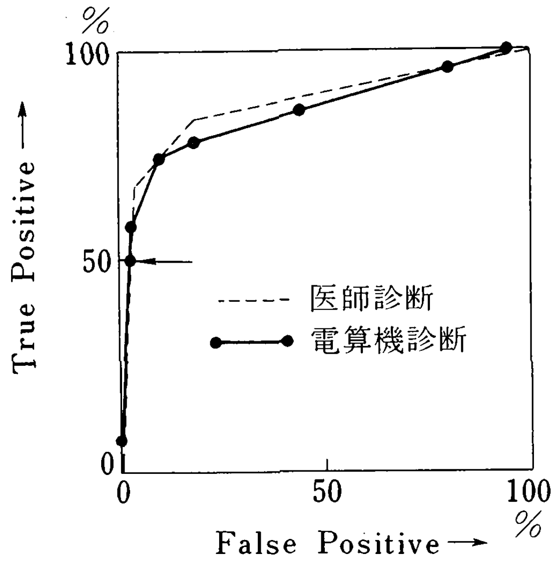


図 2　電算機診断と医師診断のROC曲線

図2は，以上の2つの判別成績をROC曲線で表した。医師診断が全体として数量化Ⅱ類よりよいことが分かる。また，両診断で判別的中率が最高の点は，図中の矢印の点に対応する。判別境界点は，α誤りとβ誤りの等しい点(電算機診断ではサンプルスコア0，医師診断では悪性否定しえず)の前後に決めればよい。すなわち，判別的中率80.6％を採用することにする。(新村)

アベイラビリティ

アベイラビリティは簡単にいえば，修理可能なシステムが所要の時点で機能を発揮しうる確率，またはシステムが所要の期間中に機能を発揮している時間の割合をいう。

アベイラビリティには瞬間，平均，固有，運用の各アベイラビリティがある。

アベイラビリティは稼動率，利用率ともいわれる。このときには，生産の都合などでシステムが停止する時間は非稼動，または，非利用の時間に含まれ，まぎらわしいので，区別して用いたほうがよい。

★解説

瞬間アベイラビリティとは，規定された所要の時点においてシステムが稼動しうる(つまり，要求された機能を維持している)確率をいう。

平均アベイラビリティは簡単にいえば，

$$\frac{(\text{要求機能を果たしうる累積時間})}{(\text{観測対象となった累積時間})}$$

である。これは，次の固有および運用アベイラビリティを含む。

固有アベイラビリティとは，

$$\frac{\mathrm{MTBF}}{\mathrm{MTBF}+\mathrm{MTTR}}$$

である。ここでMTTRは平均修復時間(Mean Time To Repair の略)である。また，**運用アベイラビリティ**は，

$$\frac{\mathrm{MUT}}{\mathrm{MUT}+\mathrm{MDT}}$$

である。ここで，MUT は Mean Up Time，MDT は Mean Down Time の略である。

アベイラビリティは信頼性工学の中では，**耐久性**(MTBF)と**保全性**(MTTR)を総合して示す尺度といえる。

関連ページ
MTBF 40
信頼性 184
耐久性・保全性 184

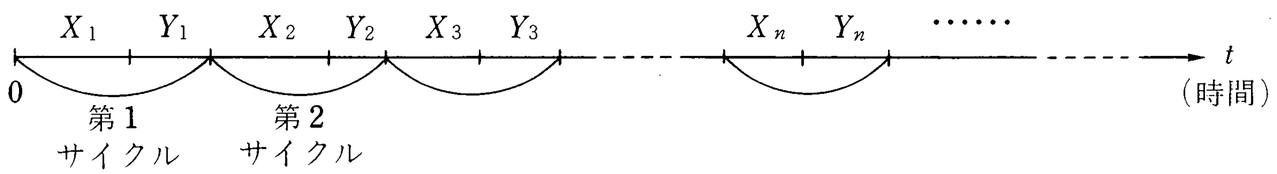


図 1

■平均アベイラビリティと固有または運用アベイラビリティ　システムが稼動（または機能を果たす）している期間を X_1, X_2, ……とし，この期間の間における不稼動（または機能を果たさない）である期間を Y_1, Y_2, ……とし，これを図1のように考える。ここで，X_i は互いに独立な同一分布にしたがう確率変数，Y_i も同様なものと仮定しておく。

このとき，X_1, Y_1 の組を第1サイクル，X_2, Y_2 を第2サイクルなどのようによぶが，ここで平均アベイラビリティは，

（要求機能を果たしうる累積時間）

$$=\sum_{i=1}^{n} X_i$$

（観測対象となった累積時間）

$$=\sum_{i=1}^{n} X_i+\sum_{i=1}^{n} Y_i$$

であるから，

平均アベイラビリティ

$$=\frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i+\sum_{i=1}^{n} Y_i}$$

となる。これを変形すれば，

平均アベイラビリティ

$$=\frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\frac{1}{n}\sum_{i=1}^{n} X_i+\frac{1}{n}\sum_{i=1}^{n} Y_i}$$

であり，ここで大数の法則により

$n\to\infty$ のとき

$$\frac{1}{n}\sum_{i=1}^{n} X_i \longrightarrow E(X_1)$$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \longrightarrow E(Y_1)$$

であることを見れば，

平均アベイラビリティは，

$$\frac{E(X_1)}{E(X_1)+E(Y_1)}$$

となる。固有アベイラビリティのときは，

$$E(X_1)=\mathrm{MTBF},\ E(Y_1)=\mathrm{MTTR}$$

とし，運用アベイラビリティのときは，

$$E(X_1)=\mathrm{MUT},\ E(Y_1)=\mathrm{MDT}$$

と考えていたと解すると，いくつかのアベイラビリティについての説明を統一できる。

■アベイラビリティの問題点　アベイラビリティはあくまで時間比で評価されている。しかし，システムなどでは機能が上昇したり，劣化することもあるので，これらをも含めて評価したほうが便利なこともある。〔25〕〔51〕（真壁）

移動平均法

時系列データ $x_1, x_2, \cdots, x_t, \cdots, x_T$ において，期間ごとの平均をとる方法が移動平均法である。たとえば，3項単純移動平均法は，

$$x'_2=\frac{x_1+x_2+x_3}{3},\ x'_2=\frac{x_2+x_3+x_4}{3},\ \cdots,\ x'_t=\frac{x_{t-1}+x_t+x_{t+1}}{3},\ \cdots$$

によって計算する。移動平均は長期間にわたる基本的な傾向変動を見るのに用いられる。また不規則変動を除去したり，移動平均の項数をうまく選ぶことにより短期の周期変動を取り除くことができる。

★解説

平均項数として5項をとった5項単純移動平均は，

$$\frac{x_{t-2}+x_{t-1}+x_t+x_{t+1}+x_{t+2}}{5} \qquad t=3,\ 4,\ \cdots,\ T-2$$

である。7項単純移動平均，9項単純移動平均，…は同様である。

季節変動のある月別時系列データは，時刻 t での中心化12項移動平均

$$\frac{1}{12}\left(\frac{1}{2}x_{t-6}+x_{t-5}+\cdots+x_t+\cdots+x_{t+5}+\frac{1}{2}x_{t+6}\right)$$

により，季節変動の波を除去することができる。

移動平均法は，次の2つの特性をもっている。

(1) m項移動平均により，時系列データの中の周期m, $m/2$, $m/3$, …の周期をもつ周期波動は除去される。

(2) m項移動平均により，時系列データの中の偶然変動の分散が，移動平均系列中では $1/m$ になる。

移動平均法の欠点は，移動平均をとることにより，実在しない周期変動をときには取り入れてしまうことがあることである。

関連ページ

季節調整の方法　76

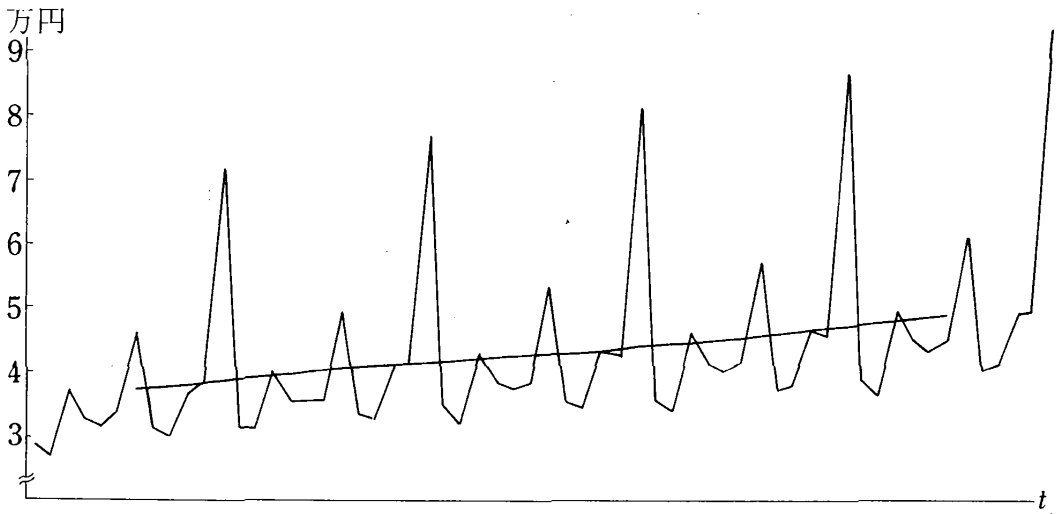


図1　原系列と中心化12項移動平均列

全国の百貨店の販売額(単位：円)について，1975年1月から1979年12月までのデータを図示したのが，図1の折れ線である。お中元とお歳暮のシーズンである6月と12月にピークがあるのに気づく。

月	1975年	1976年	1977年	1978年	1979年
1	2908.	3156.	3490.	3635.	3898.
2	2706.	3127.	3224.	3429.	3676.
3	3773.	4038.	4336.	4663.	5008.
4	3306.	3626.	3907.	4191.	4543.
5	3207.	3553.	3771.	4021.	4336.
6	3443.	3643.	3824.	4152.	4510.
7	4585.	4984.	5408.	5784.	6162.
8	3159.	3409.	3573.	3738.	4037.
9	3035.	3283.	3500.	3828.	4160.
10	3708.	4099.	4321.	4672.	4889.
11	3840.	4110.	4235.	4609.	4977.
12	7210.	7696.	8156.	8749.	9390.

このデータに基づいて，中心化12項移動平均を計算する。これは，1年を周期とする季節変動があるという知識に基づいている。1975年の移動平均値は，

$$x_7'=\frac{1}{12}\Big(\frac{1}{2}\times2908+2706+3773$$
$$+\cdots+3840+7210+\frac{1}{2}\times3158\Big)$$
$$=3750$$

になる。以下同様に，x_8'，x_9'，…を求めると次のようになる。

月	1975年	1976年	1977年	1978年	1979年
1		3906.	4195.	4455.	4794.
2		3933.	4219.	4478.	4823.
3		3954.	4235.	4499.	4849.
4		3981.	4254.	4527.	4872.
5		4008.	4268.	4557.	4896.
6		4040.	4292.	4597.	4938.
7	3750.	4074.	4318.	4633.	
8	3778.	4092.	4332.	4654.	
9	3806.	4108.	4354.	4679.	
10	3831.	4132.	4380.	4708.	
11	3858.	4153.	4402.	4736.	
12	3881.	4170.	4426.	4764.	

これを図示したのが図1の変化の少ない曲線である。原系列には12カ月を周期とした規則性が見られるが，中心化12項移動平均をとった系列では，季節変動がほぼ除去されている。

この方法は系列の最初の6項と最後の6項が計算できないという不便さはある。しかし，最小2乗法などによる当てはめに比べれば計算が容易であり，傾向線の意味も理解しやすい利点がある。傾向変動が簡単な曲線で表せるときは，最小2乗法による曲線の当てはめでよいが，移動平均法はそうでない場合でも用いることができる。(杉山)

因子

実験結果に影響を及ぼすと思われるいろいろな原因系で，その実験でとりあげたものを因子とよぶ。効果的な実験を行うには，因子とそのとりうる条件の選定を技術的によく事前検討することが必要である。

★解説

因子のとりうる種々の条件を水準とよび，その水準指定の難易などにより因子をつぎのように分類できる。

(1) **制御因子**　実験結果をみて，最適な水準を自由に採択できる因子を制御因子とよぶ。たとえば，反応時間，反応温度，加工方法など実験でとりあげる因子の多くは制御因子である。ばらつきを評価するのが目的であるサンプリング実験を除き，この因子が1つもない実験は無意味で，実験因子のなかでもっとも重要なものである。

(2) **標示因子**　水準に再現性があって，その水準をとりあげる意味はあるが，最適水準を自由に選ぶことができない因子。他の制御因子との交互作用を求めるのが，この因子をとりあげる主な目的である。原材料をA，Bの2社から購入していて，これを1社にまとめることができないとき，原材料は標示因子である。

(3) **ブロック因子**　実験のすべてを1日では遂行できないときに，何日にもわたる実験全体を**無作為化**して行うと，日の違いによる変動が誤差に組み込まれ，誤差変動は非常に大きくなる。このようなときには，日を因子としてとりあげて，これによる変動を誤差変動から除去すれば誤差分散は小さくなる。このように実験の場をいくつかに分割して，実験の精度をよくする目的でとりあげられる因子をブロック因子とよび，分割された個々のものをブロックとよぶ。ブロック因子の水準には再現性がなく，したがって制御因子との交互作用を求める意味はない。日のほかには，ロット，作業者，地域などがブロック因子の例となる。

関連ページ

要因効果　394

無作為化　384
乱塊法　402
ラテン方格法　400

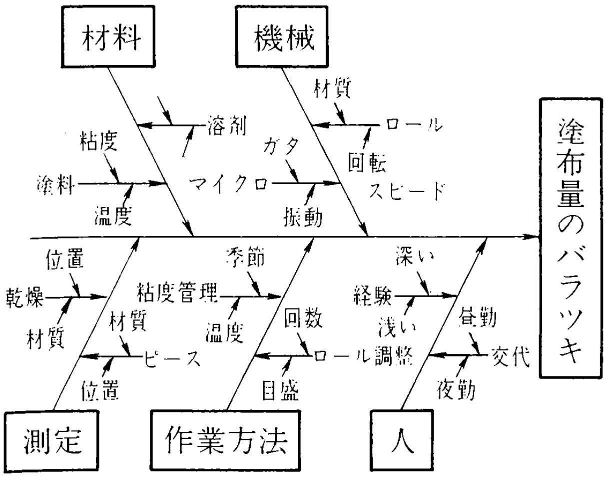


図1　特性要因図の一例

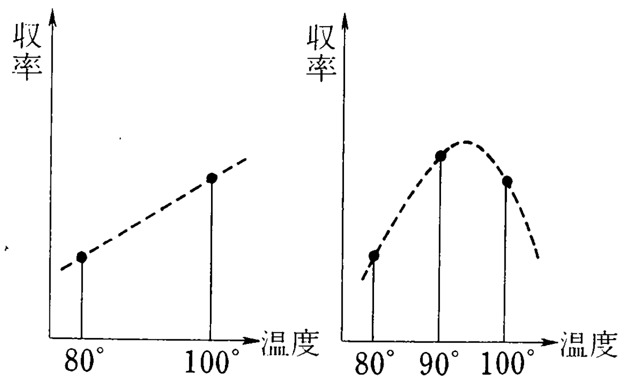


図2　応答曲線

■特性値の選定　実験を行うときには，その目的を明確にして，目的に直結した特性値を選ぶことが第一に重要である。たとえば，製品の収率を問題にしているときに，収率と炉の温度に一定の関係があることが明らかになっていれば，温度を収率のかわりの特性値と考えてもよい。そうでなければ，収率そのものを特性値に選ぶべきである。第二に，データ解析の立場からは，要因効果の加法性を満たすような特性値が望ましい。これにより無視できる交互作用の数が多くなり，直交表による実験回数を減らし，多くの因子をとりあげることができる。

■因子の選定　実験の結果を定量的に把握するのに特性値は必要であるが，実験を計画するにあたっては，結果に影響する原因系を明確にして，そのなかから重要なものを選定することがつぎに重要となる。原因系を明確にするには，特性要因図を用いるとよい。特性要因図は，図1のように特定の結果と原因系との関係を系統的に表した図のことで，これを書くことにより重要な因子を見落とすことを防ぐことができる。考えられる重要原因をすべて列挙した特性要因図を書いたのちに，過去の経験や技術情報などをできるだけ集めて，特性値の変動に大きな影響を与えると思われる要因を，因子としてとりあげる。

■水準の選定　原料の銘柄，種類，加工方法などの質的因子については，現在使用可能であるものが限られていることから，とりあげるべき水準はあらかじめ決まっていることが多い。反応温度や反応時間などの量的因子については，任意に水準数を選ぶことができるが，実験規模を大きくしないでなるべく多くの因子をとりあげるには，水準数をあまり多くするべきではない。水準の変化による特性値の応答が，直線的傾向，2次または3次曲線的傾向ならば，それぞれ少なくとも2，3，4水準を必要とする。しかし，水準の幅をあまり大きくしなければ，3水準とれば十分な場合が多い。　（宮村）

因子軸の回転

共通因子の解釈は因子負荷量の値から推察する。因子負荷量は共通因子 f_1, f_2, …, f_k の張る k 次元空間を決めるだけで，互いに直交しているその方向は任意である。そこで，解釈しやすいように直交回転を行うことになる。因子軸の回転の方法としてはバリマックス法，プロクラステス回転などいろいろと考えられている。

★解説

因子分析のモデルを，

$$\chi_i = \lambda_{i1}f_1 + \lambda_{i2}f_2 + \cdots + \lambda_{ik} + e_i$$

$$(i=1,\ 2,\ \cdots,\ p)$$

で表す。$h_i^2 = \lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{ik}^2$ を共通性(communality)という。バリマックス法は，

$$\frac{1}{p}\sum_{l=1}^{k}\left\{\sum_{i=1}^{p}(\lambda_{il}^2)^2 - \frac{1}{p}\sum_{i=1}^{p}(\lambda_{il}^2)^2\right\}$$

を最大化する直交回転を見つけることである。これは因子負荷量の絶対値を，各因子について0に近いものと1に近いものに分離し，因子の解釈を容易にしようとするものである。上式で λ_{il}^2 のかわりに共通性で割った λ_{il}^2/h_i^2 を代入した式を最大化する方法を，規準バリマックス法といい，こちらの方法を用いる場合の方が多い。

ある与えられた因子負荷量に近づけるような回転をプロクラステス回転という。実験者が考えている仮説因子負荷行列を $\{\lambda_{il}^{(0)}\}$ で表す。データから推定した因子負荷行列 $\{\lambda_{il}\}$ が，この $\{\lambda_{il}^{(0)}\}$ に近いかどうかを調べたいというようなときに，この回転は有用である。

関連ページ
因子分析　26
因子負荷量　24

中学3年生137人の2学期の成績データの相関行列に主成分分析を行い，得られた第1主成分，第2主成分(→176ページ)に，係数が0に近い値と絶対値の大きい値に大別するQuartimax回転を行ったのが下記の数値である。

	Comm.	第1主成分	第2主成分
国語 x_1	.840	.916	.012
社会 x_2	.868	.932	−.013
数学 x_3	.777	.881	.026
理科 x_4	.880	.937	.045
音楽 x_5	.785	.885	.040
美術 x_6	.641	.782	−.174
体育 x_7	.970	.617	.768
技家 x_8	.789	.880	.124
英語 x_9	.846	.919	.039

表1　主成分分析

相関行列を出発点とし，共通因子の数 $m=2$ として最尤法により求めた因子負荷量に，同じくQuartimax回転を行うと次のようになる。

	Comm.	第1因子	第2因子
x_1	.812	.901	.003
x_2	.870	.930	−.072
x_3	.753	.862	−.102
x_4	.900	.946	−.070
x_5	.736	.857	.038
x_6	.514	.716	.041
x_7	.407	.634	.068
x_8	1.198	.896	.629
x_9	.853	.918	−.100

表2　因子分析

主成分分析で得られた2つの主成分(表1)を，因子分析の2つの因子(表2)へ近づけるようなプロクラテス回転を行うと以下のような結果を得る。

	Comm.	第1主成分	第2主成分
x_1	.839	.916	−.023
x_2	.869	.931	−.048
x_3	.777	.881	−.007
x_4	.880	.938	.009
x_5	.785	.886	.006
x_6	.642	.775	−.204
x_7	.971	.646	.744
x_8	.790	.884	.091
x_9	.846	.920	.004

表3　プロクラステス回転

これを見ると，主成分分析の第1主成分と，因子分析の第1因子とはよく似たものであることがわかる。第2主成分は体育 x_7 の係数が0.744と大きく，また第2因子は技家 x_8 の係数が0.629と大きくなっていて，一致していない。

上の数値はプロクラステス回転により，表1の2つの因子が表2の2つの因子に最も近づいた場合である。これは最小2乗法の原理を用いた回転である。

(杉山)

因子負荷量

主成分 $y=h_1x_1+h_2x_2+\cdots+h_px_p$ において，主成分と説明している変数との相関関係が因子負荷量である。係数 h_i が大きければ大きいほど因子負荷量は大きく，x_i はその主成分と強く相関していることになる。そのことは，x_i が主成分をよく説明している変数であることを意味している。

★解説

主成分 y の固有値を l で表すと，相関行列に基づいて主成分分析を行った場合の y と x_i との相関関係は，

$$\sqrt{l}\,h_i$$

である。これが因子負荷量である。固有ベクトル$(h_1,\ h_2,\ \cdots,\ h_p)$のかわりに，因子負荷量$(\sqrt{l}\,h_1,\ \sqrt{l}\,h_2,\ \cdots,\ \sqrt{l}\,h_p)$の値を示すことがよくある。

分散共分散行列から主成分分析を行った場合の因子負荷量は，

$$\frac{\sqrt{l}\,h_i}{\sqrt{x_i\text{の分散}}}$$

により得られる。

固有ベクトルは，

$$h_1^2+h_2^2+\cdots+h_p^2=1$$

という制約条件があることから，変数の数 p が増すと，1つ当たりに割りふられる h_i^2 の平均は $1/p$ と小さくなるが，因子負荷量の場合は主成分と x_i との相関関係であるから，そのような動きはない。相関係数の意味を正確に把握していれば，数値としては固有ベクトルよりも読みやすい。

関連ページ
主成分分析における固有値　178
主成分分析における固有ベクトル　180
相関関係　222

標準化した変量を z_i で表すと，中学3年生137人の9教科の成績の相関行列に基づく第1主成分は次のようであった(→176ページ)。

$$y_1=0.350z_1+0.355z_2+0.337z_3 +0.359z_4+0.339z_5+0.292z_6 +0.263z_7+0.340z_8+0.352z_9$$

主成分 y_1 の固有値は6.810であるから，国語 z_1 と y_1 との相関係数は，

$$\sqrt{6.810}\times 0.350=0.913$$

である。社会 z_2 と y_2 との相関係数は，

$$\sqrt{6.810}\times 0.355=0.926$$

になる。以下同様の計算で主成分 y_1 と各教科との相関係数を計算し，第1主成分の因子負荷量を求める。

	第1主成分 y_1 の因子負荷量
国語 z_1	0.913
社会 z_2	0.926
数学 z_3	0.879
理科 z_4	0.937
音楽 z_5	0.885
美術 z_6	0.762
体育 z_7	0.686
技家 z_8	0.887
英語 z_9	0.919

美術 z_6 と体育 z_7 の因子負荷量は多少小さめであるが，他の教科は0.9前後であり，比較的大きい。符号は正で，どの教科の成績が上がっても，y_1 は大きくなるという傾向がある。

主成分の意味を解釈するとき，係数の大きい変数に注目して，その主成分の意味づけを考えた。係数が大きいということは，主成分とそれを説明している変数 z_i との相関が強いということである。言いかえれば，その変数は主成分をよく説明していることになるからである。

第2主成分 y_2 の固有値は0.586であり，因子負荷量は，

0.074(z_1)	0.043(z_4)	0.707(z_7)
0.100(z_2)	0.044(z_5)	0.041(z_8)
0.057(z_3)	0.246(z_6)	0.047(z_9)

である。y_2 と比較的高い相関をもっているのは体育 z_7 だけである。このことから第2主成分は体育の因子と解釈される。

因子負荷量は主成分ともとの変数 z_i との相関係数そのものであり，解釈が容易であることから，固有ベクトルのかわりに因子負荷量で結果を示す人も多い。

相関行列に基づく因子負荷量で説明したが，分散共分散行列に基づく因子負荷量の場合も解釈はまったく同じである。相関行列の場合は，主成分と標準化した変数 z_i との相関係数であったが，分散共分散行列の場合は，主成分と標準化する前の変数 x_i との相関係数を与えることになる。　(杉山)

因子分析

因子分析の目的は，多変量 x_1, x_2, …, x_p の相関行列の構造を，少数の因子 f_1, f_2, …, f_k により説明することである。変数を x で記すと次の構造

$$x_i=\lambda_{i1}f_1+\lambda_{i2}f_2+\cdots+\lambda_{ik}f_k+e_i \qquad (i=1,\ 2,\ \cdots,\ p)$$

を考える。ここで，e_1, e_2, …, e_p は互いに独立で，かつ共通因子 f_1, f_2, …, f_k とも独立であると仮定する。これは各変数 x_i をいくつかの基本特性 f_1, f_2, …, f_k によって記述するモデルであり，e_i は共通因子では説明できない残差を表している特殊因子である。

★解説

因子分析のモデルは**重回帰モデル**とよく以ているが，大きく異なる点は共通因子 f_1, f_2, …, f_k が観測不可能なことであり，因子数 k も未知であることである。係数 λ_{i1}, λ_{i2}, …, λ_{ik} $(i=1,\ 2,\ \cdots,\ p)$ を因子負荷量といい，特殊因子 e_i の分散とともにデータから推定したい母数である。共通因子の効果が除かれたあとは，e_i は互いに無相関になっていることが要求される。

1つの**共通因子**により，

$$x_i=\lambda_{i1}f_1+e_i \qquad (i=1,\ 2,\ \cdots,\ p)$$

としたとき，データからの推定量 $\hat{e}_1$, $\hat{e}_2$, …, $\hat{e}_p$ の相関が0に近ければ，因子数は1ということになる。いくつかの共通因子を用いたときに，$\hat{e}_1$, $\hat{e}_2$, …, $\hat{e}_p$ の相関係数がすべて0とみなせるかが分析のポイントになる。これは**因子数 k の決定**とも関連している。共通因子 f_1, f_2, …, f_k のそれぞれが具体的にどういう意味あいの因子であるかは，データーから推定した因子負荷量 λ_{ij} の値の大きさから推察することになる。

関連ページ
重回帰モデル 170
相関関係 222

中学3年生137人の2学期の成績のこの相関行列は次のようであった。この相関行列が因子分析の出発点となる。

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
国語	x_1	1								
社会	x_2	.84	1							
数学	x_3	.76	.78	1						
理科	x_4	.83	.91	.82	1					
音楽	x_5	.78	.78	.78	.80	1				
美術	x_6	.67	.70	.61	.66	.61	1			
体育	x_7	.57	.57	.55	.60	.55	.45	1		
技家	x_8	.81	.79	.71	.80	.79	.67	.61	1	
英語	x_9	.86	.84	.83	.87	.78	.62	.58	.76	1

因子数 $m=1$ としたとき，つまり共通因子は1つであるとしたときの因子負荷量を相関行列から最尤法により計算すると次のようになる。

国語	x_1	0.905
社会	x_2	0.931
数学	x_3	0.863
理科	x_4	0.945
音楽	x_5	0.859
美術	x_6	0.719
体育	x_7	0.636
技家	x_8	0.860
英語	x_9	0.918

つまり，それぞれの教科の成績は，f_1 という共通因子により，

$$x_1=0.905f_1+e_1$$
$$x_2=0.931f_1+e_2$$
$$x_3=0.863f_1+e_3$$
$$x_4=0.945f_1+e_4$$
$$x_5=0.859f_1+e_5$$
$$x_6=0.719f_1+e_6$$
$$x_7=0.636f_1+e_7$$
$$x_8=0.860f_1+e_8$$
$$x_9=0.918f_1+e_9$$

と表されることになる。同じデータに主成分分析を行ったときの第1主成分は次のようであった(→176ページ)。

$$y_1=0.350x_1+0.355x_2+0.337x_3+0.359x_4+0.339x_5+0.292x_6+0.263x_7+0.340x_8+0.352x_9$$

第1主成分の係数に0.905/0.350をかけたものは次のようになる。これは，因子分析の x_1 の因子負荷量に値をそろえただけであり，それ以外の特別な意味はない。

国語	x_1	0.905
社会	x_2	0.918
数学	x_3	0.905
理科	x_4	0.928
音楽	x_5	0.876
美術	x_6	0.755
体育	x_7	0.680
技家	x_8	0.879
英語	x_9	0.910

上の数字は，左側の因子分析における因子負荷量の数字によく以ている。これは比較的単純な構造をしている相関行列に基づく分析で，このように似かよった結果が出ることもある例である。主成分分析は分散指向の分析法であり，因子分析は相関係数を指向した分析である点で，一般に結果は異なると考えられる。　(杉山)

因子分析における因子負荷量の推定

因子負荷量と特殊因子の分散の推定法としては，多変量正規分布を前提とした最尤(さいゆう)法や，非線形最小2乗法などがある。最尤法は分散共分散行列に，仮定した因子構造を代入し，尤度関数を最大にする点を求める方法である。実際には，適当な初期値を与えて反復解法を用いる。

★解説

関連ページ

最尤推定値を得る方法としては，Jöreskog（1967年，Psychomerika）による方法，Jennrich and Robinson（1969年，Psychomerika）による方法などが知られている。Jöreskogによる方法は，それまでに知られていた最尤解を求める方法より解への収束が速いことなどの特徴をもっている。これまでの反復法では特殊因子 e_i の分散の推定値が負になると計算が続行できず，その後の様子を知ることができなかった。実際には反復解法のある段階で e_i の分散の推定値が負になってから，さらに計算を続行すると，真の解へ収束することがある。Jennrich and Robinsonの方法は，この点を克服している。

最尤法は変数 x_1, x_2, …, x_p が多変量正規分布に従うことを前提としているが，**最尤推定量**であるので推定量の統計的性質（データが十分多いとして）がわかるという利点がある。また，因子数に関する検定を行うことができる。他方，非線形最小2乗法は，多変量正規分布に従うという前提条件は必要ない。

最尤法　134

共分散　229

変数 x_i と x_j との**標本共分散**を s_{ij} とで表す。因子数を k としたときに，最尤法による x_i と x_j の共分散の推定値を $\hat{\sigma}_{ij}$，特殊因子 e_i の分散の推定値を ψ_i で表すと，

$$\chi^2 = n'\sum_{i<j} (s_{ij} - \sigma_{ij})^2/\psi_i\psi_j$$

は自由度 $1/2\{(p-k)^2-(p+k)\}$ の χ^2 **分布**で近似される。ここで，$n'=$ 標本数 $-1+(2p+5)/6-2k/3$

χ^2 分布　314

中学3年生137人の2学期の成績データから，共通因子は1つであることの仮説検定を行うと，

$$\chi^2=61.70$$

になる。これは，χ^2分布の有意水準0.1％点55.48より大きく，有意である。

共通因子は2つであるとして仮説検定を行うと，

$$\chi^2=38.89$$

になり，このときもχ^2分布の有意水準1％点36.19より大きく，有意である。

因子負荷量は，

	comm.	第1因子	第2因子
国語 x_1	0.813	0.901	0.003
社会 x_2	0.870	0.930	−0.072
数学 x_3	0.754	0.862	−0.102
理科 x_4	0.900	0.946	−0.070
音楽 x_5	0.736	0.857	0.038
美術 x_6	0.514	0.716	0.041
体育 x_7	0.407	0.634	0.068
技家 x_8	1.197	0.896	0.629
英語 x_9	0.853	0.918	−0.100

である。国語x_1のcomm.(communalityの略)は，

$$0.901^2+0.003^2=0.813$$

であり，

$$1-0.813=0.187$$

が

$$x_1=0.901f_1+0.003f_2+\hat{e}_1$$

における$\hat{e}_1$の分散である。同様に社会x_2のcommunalityは，

$$0.930^2+(-0.072)^2=0.870$$

であり，$\hat{e}_2$の分散は0.130となる。

技家x_8に注目すると，そのcommunalityは，1.197と1.0を超えており，このことは$\hat{e}_8$の分散が負であることを意味している。このことから共通因子の数 $m=2$ のときは，適切な解が得られないことがわかる。

いま，共通因子の数 $m=1$ のときの解に，1回のテストの試験のみに影響する因子(specific factor)が付加した形になるように回転してみよう。

	comm.	第1因子	第2因子
国語 x_1	0.813	0.900	0.036
社会 x_2	0.870	0.932	−0.038
数学 x_3	0.754	0.865	−0.070
理科 x_4	0.900	0.948	−0.035
音楽 x_5	0.736	0.855	0.069
美術 x_6	0.514	0.714	0.067
体育 x_7	0.407	0.631	0.091
技家 x_8	1.197	0.872	0.661
英語 x_9	0.853	0.921	−0.066

共通因子として意味をもつのは，因子負荷量の絶対値の大きい係数が2つ以上あるときである。解として安定しているのは，大きな値をとる係数が3つ以上ある場合であるともいわれている。上の数値結果からは，第2因子は意味のある因子ではなく，specific factor(特殊因子)と考えられる。

(杉山)

インダストリアル・エンジニアリング

インダストリアル・エンジニアリング(Industrial Engineering)とは，簡単にいえば，生産工場やオフィス(たとえば，事務所や病院など)で人間，設備・機器および材料などあわせて最適な生産や事務処理サービスなどを遂行するための科学的な手法の体系をいう。インダストリアル・エンジニアリングはIEとも略称する。

★解説

インダストリアル・エンジニアリング(以下，IEと略)の発展の源は，テイラー(F. W. Taylor)までにさかのぼる。テイラーは19世紀の末から，今世紀のはじめにかけて，科学的管理の考え方を提唱した。この考え方は，作業を行ったときに，よい結果が得られれば必ずよい原因が，悪い結果が生じたときには悪い原因が存在するはずであるから，結果をよく科学的に分析して原因を究明し，よい原因はよく再現するように，悪い原因はこれを取り除くように手段を講ずることを基本とする，とした。

テイラーの科学的管理(scientific management)は，今世紀に入ってガント(Henry L. Gantt)やギルブレス(F. G. Gilbreth)などによって引きつがれ，これが次第に近代化された。さらに1910年前後に米国の産業の発展にともない，フォードシステムが完成したのに続いて，科学的管理の考え方は，インダストリアル・エンジニアリング(IE)として新しい進展をとげるようになった。

IEはオペレーションズ・リサーチや統計的手法とも密接な関係がある。たとえば，IEの中でとりあげられるラインバランスの問題には待ち行列などのOR手法が，時間研究に必要なワークサンプリングには統計手法が必要とされる。

関連ページ

ガント図　280

待ち行列　366
オペレーションズ・リサーチ　48

■IEの歴史　IEの起源は1890年ごろにさかのぼる。このころ，米国のベスレヘム製鋼所において職工，職長として勤務していたテイラーは1人1人の作業量にバラツキが大きいことに注目し，一日の科学的に公正と考えられる仕事量を定めることの必要性を認識した。この作業研究を続けて，テイラーは，ついに，1910年に『科学的管理法の原理』という本を出版した。

その後，この考え方は，単なる労働強化のためのものではないかとの批判もあり，紆余曲折はしたが，1920年代を過ぎてから，米国政府の標準局やドイツのRefa（作業時間設定全国委員会）などの協力の下に，それぞれ独自の発展をとげるようになった。

また，IEは1920年代に米国のシュハートによって提唱されて，1950年代にわが国でも組織的に推進され，発達した品質管理との関係も深い。最近は計算機の発達にともなって，IEの中身も近代化の歩を速めている。

■IEの技法　IEには，作業研究，工程分析，動作分析，時間分析，プラント・レイアウト，エンジニアリング・エコノミーなど多くの分科がある。このうちの一部をのべることにする。

(1)**作業研究**　作業のやり方について現状分析をし，ムリ，ムダ，ムラのないように，これを改善するための研究である。

(2)**工程分析**　生産の現場を流れる加工品について，その工程や作業のやり方，加工，運搬，検査，停滞の4つに分類し，その流れ方を工程分析表に整理し，工程や作業方法の改善や工程のレイアウトなどを改善する手法である。

(3)**動作分析**　作業者の行動や動作を分析し，これをサーブリック記号によって表現し，動作の改善をはかるものである。この分析を進めるには目視によって動作を分析するもののほかに，フィルムやビデオおよびメモモーションなどを用いることもある。

(4)**プラント・レイアウト**　工場や工場の中の設備の配置などを科学的に分析する手法である。工場の配置には，材料や製品の運搬コストなどを最適化するために，数理計画法などのOR手法が用いられる。

■IEのための組織　生産企業では，技術，生産，品質管理，営業，経理などの部門のほかに，IE部や生産管理部を組織の中に設けている。この部門は，企業や工場の生産に関する人員配置や生産コストの最適化をはかるためにIE技法を用いて分析，改善などの仕事を進めている。

米国などではIE技術者（industrial engineer）などの専門職種が存在し，広い分野で活躍しているが，わが国の場合には，とくにこの種の専門職種を設けず，生産部門の技術者がこれに当たることも少なくない。　（真壁）

AID(Automatic Interaction Detector)

AID は，ミシガン大学サーベイ・リサーチ・センターにおいて開発された。マーケティングなどのアンケート集計によく用いられる。外的基準のある質的データの解析手法である。外的基準を最もよく分離する項目で，データを逐次的に2分割して最終的にいくつかのグループに分ける。外的基準が連続変数の場合，外的基準値の大小順にデータをグループ分けする説明項目がみつかる。外的基準が 0/1 で表される質的データの場合，各グループにおけるカテゴリー1の占める割合の大小順にグループを並べ換え，各グループを特徴づける説明項目のカテゴリー値を比較検討すればよい。その形状が枝分かれになるので，外的基準のない場合のアソシエーション分析とともに，「ツリー分析」ともよばれる。

★解説

▩アルゴリズム　説明項目 a のカテゴリー値の同じものを1グループとして，外的基準 y の平方和($S_T(a)$ と略す)と級間平方和($S_B(a)$)を計算し，比 $S_B(a)/S_T(a)$ を求める。この比の値の一番大きな項目が，最も説明力のある項目として第1段階で選ばれる。ただし，3カテゴリー以上をもつ項目は，カテゴリーを2組に分ける組み合わせの中で比を最大にする組を新しい2カテゴリーとして扱う。

第1段階で選ばれた項目のカテゴリーで，サンプル全体を2分割する。第2段階では，分割後のデータに対して残りの項目の中から第1段階と同様に比を最大にする項目を選ぶ。第3段階以降も同じ手順をふんで行う。

第 k 段階で，分割後のデータ数 N_k が一定数以下になったり，第1段階と第 k 段階の次の平方和の比，S_B^k/S_T^1 または S_T^k/S_T^1 が一定値以下になれば停止する。多くの項目の中から，説明に有効な項目を選択してくれる便利な手法である。外的基準が2値カテゴリーの場合，外的基準に0と1を与えて各平方和を計算する。

関連ページ

クロス集計　90

枝分かれ法による判別　34

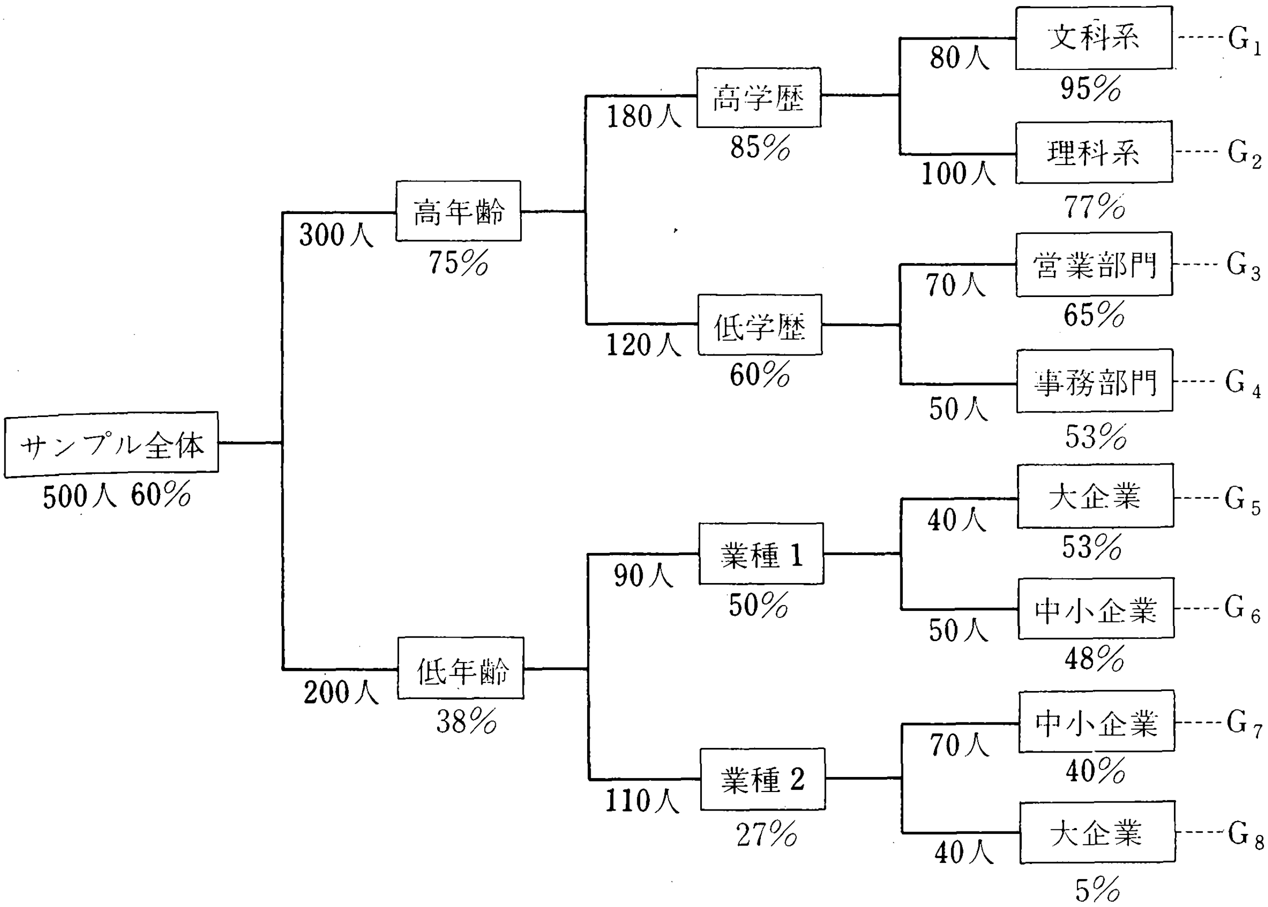


図1　所得格差のアンケート調査の分析

いま，500人のサラーマンに，所得，年齢，学歴，業種，出身学部，仕事の内容，会社規模などの項目についてアンケート調査を行ったとする。いま，説明を簡単にするため，年収400万円以上を1，未満を0とする2値カテゴリーに分ける。サンプル全体でカテゴリー1は60％を占めていた。

この0/1を外的基準として，各項目ごとに相関化 S_B^1/S_T^1 を計算した。この結果，年齢項目が最も高い値を示した。年齢項目は，20歳台から50歳台までの4カテゴリーであったが，偶然に20歳台と30歳台そして40歳台と50歳台の2組に分かれたので，一方を高年齢群，他方を低年齢群と名づけた。年齢項目によりデータは300人と200人に2分割され，年収400万円以上の割合は75％と43％になった。

高年齢群に対しては，第2段階で学歴が選ばれた。低年齢群では業種項目が選ばれた。

第3段階で，元のデータが8個のグループに分割された。第1グループ(G_1)は，高年齢·高学歴·文科系というカテゴリー属性をもつ80人のサラリーマンであり，95％が年収400万円以上であった。低年齢·業種2·大企業の属性をもつグループは，一番所得の少ないグループを形成している。(新村)

枝分かれ法による判別

ベイズ診断と同様に，質的データの判別手法が未発達な時代に主として医学分野で問診や各種の診断体系の作成に使われた。Yes と No の 2 カテゴリーをもつ質問項目を逐次的に 10 回質問すれば 1024 個のカテゴリーの組み合わせが決まり，対象を分類できる。質問(論理判断)が 3 値以上の多値の場合の枝分かれ法も現実には多い。

★解説

■枝分かれ法　逐次的な論理判断の流れは，図 1 のように樹木が次々に枝分かれするのと同じ形で表せるので枝分かれ法という。◇ は論理判断を示し，Yes の場合には Y の枝を，No の場合には N の枝を選ぶ。次々に枝分かれを繰り返した後で，いくつかのグループ分けができる。これらの 1 つ 1 つに病名などの分類名を与えればよい。図 1 は論理表として表すこともできる。また，IF 文などの条件判定文を用いてコンピュータ・プログラム化できる。

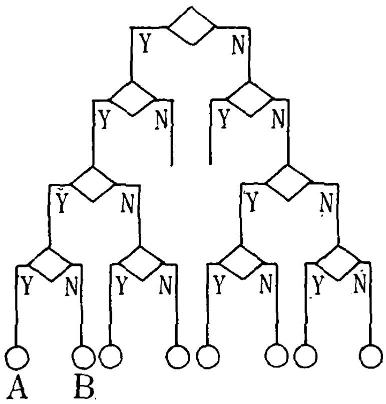


図 1　枝分かれ法

■心電図診断　心電図波形から各種の連続な計測値が測定され，これらの値を用いて心電図所見といわれるグループ分けの「心電図診断体系」が確立されている。この診断体系を枝分かれ法を用いてコンピュータシステムとして実用化したものが心電図自動解析システムである。枝分かれ法の代わりに多変量解析手法の適用も試みられたが失敗した。この理由としては，①心電図は多くの所見と計測値をもつ複雑な対象であること，②経験知識が計測値を用いて既に「心電図診断体系」として体系化されている場合，枝分かれ法はそれを利用できること，などがあげられる。

関連ページ
質的データの解析　162
ベイズの定理による判別　346
決定行列　104
ROC 曲線　14
数量化Ⅱ類　190
AID　32

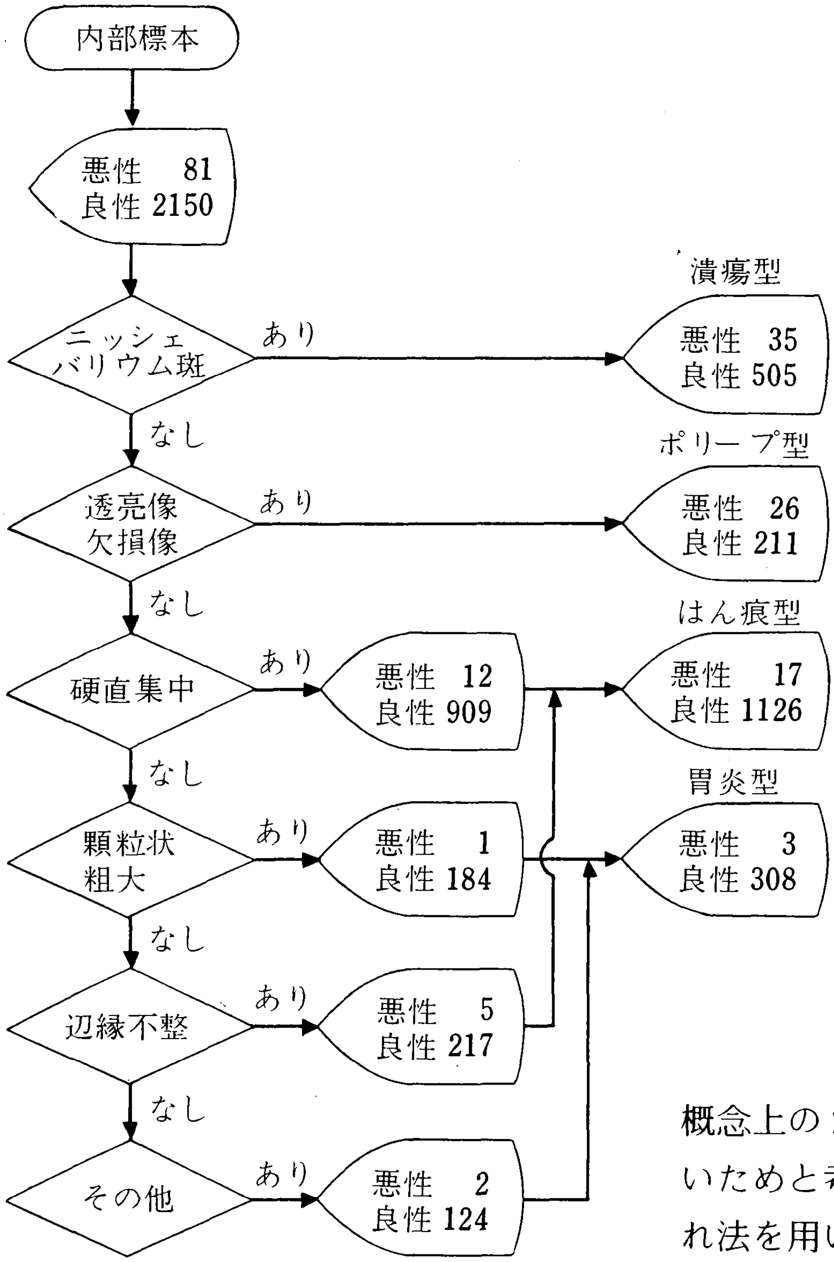


図2　胃X線写真の病型分類

図2は，大阪府立成人病センター鈴木医師らによって作成された胃X線写真の病型分類である。

胃X線検査を受けた2231症例を対象としている。追跡調査の結果，胃X線検査よりも上位の病理診断などの検査で悪性群と良性群に分け，これを外的基準とする。悪性群とは胃がん，良性群とは胃潰瘍などの良性疾患や正常症例と考えればよい。この結果81例が悪性群，2150例が良性群に分類された。

鈴木医師らは，当初この2群を胃X線写真から抽出した項目を用いて判別を試みた。説明変数として用いる項目は，胃X線写真の図形情報である。たとえば，バリウムの貯りがあるか否かとか，顆粒状の像が見えるか否かなどである。

しかし，2群判別で求められたカテゴリースコアが固有知識をよく説明しないことがわかった。この理由としては，悪性群と良性群に分けられた概念上の2群が実情をよく表していないためと考えられた。そこで，枝分かれ法を用いて4病型に層別することにした。たとえば，ニッシェ・バリウム斑のある症例は，良性の胃潰瘍であるかそれと極めて似た症状をもつ胃がんであることを示す。胃炎型の悪性群と良性群は潰瘍型と異なり，いくつかの項目の論理判断で初めて決められるので，枝分かれで後ろの方にきている。以上の4病型ごとに2群判別を行ったところ，固有知識に合致するカテゴリースコアが得られ，判別成績も少し改善された。　　　　(新村)

FMEA と FTA

FMEA はFailure Mode and Effects Analysis の略で，故障モード影響解析という。FTA はFault Tree Analysis を略したもので，故障の木解析という。

FMEA は部品レベルの故障モードを調べ，これがシステムに及ぼす影響を評価し，危険優先数の高いものから，設計・製造上の改善案を講ずるものである。FTA はプラントなどの重大事故をトップ事象としてとりあげ，この原因となる事象群を木を用いて系統的に図示し(この図を FT 図という)，重大事故の防止策を講ずるものである。

★解説

FMEA の一例は次の通りである。

部品	故障モード	故障原因	影響度			危険優先数	対策	備考
			頻度	きびしさ	検出度			
ボルト	折損	応力過大	2	10	3	60	設計改善	
		締付不足	3	10	5	150	—	
		腐食	1	7	3	21	点検により検出	

表 1

FTA の一例を次に示す。

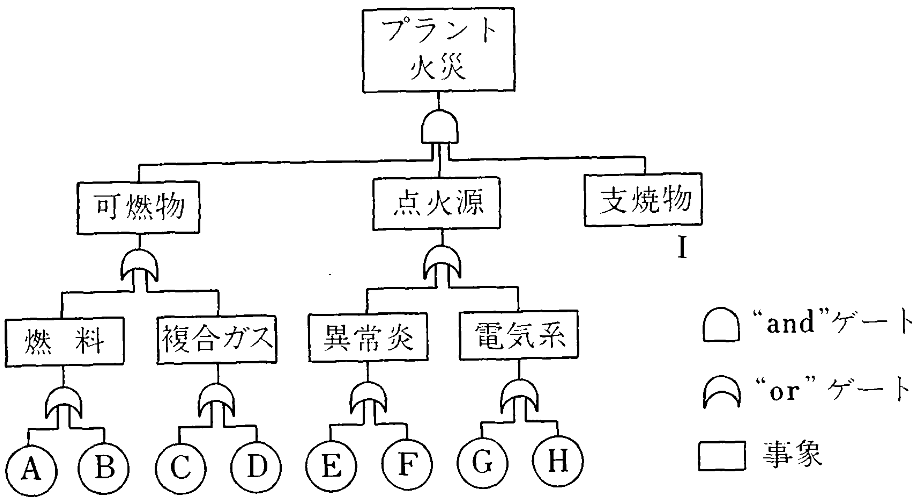


図 1

関連ページ
信頼性 184

■FMEA の誕生　FMEA は一番最初に，米国の航空機会社グラマン社で1960 年前後に用いたといわれているが，詳細は不明である。しかし，1960 年代に入って米国で人工衛星の開発が始まってから，にわかに注目されるようになり，次第に一般の産業界でも活用されるようになった。

■FMEA の作り方　FMEA の手法は“bottom up”的であるといわれているが，このことは部品レベルから，この故障のシステムへ及ぼす影響を考察することに由来するものと考えられる。FMEA の作り方は次の手順による。

(1) システムを構成する部品または部位の中で，クリティカルなものを選び出す。
(2) この部品についてすべての故障のモードを書き出す。
(3) 各部品の故障モードの原因と，そのシステムに及ぼす影響を評価する。
(4) これらの評価を総合して，危険優先数を定める。一般には，
　　危険優先数＝(頻度)×(きびしさ)×(検出度)
　として計算する。
(5) 危険優先数の高いものより逐次改善対策を立てる。

FMEA には目的によって，設計，生産準備，工程の FMEA などがある。

■FTA の誕生　FTA は一番初めにはベル研究所において，1960 年前後に完全解析のための手段として用いられたという。それ以来，次第に一般の分野に広まり，とくに，MIT のラスムッセン(Rasmussen)報告が原子力発電所の安全性解析に FTA を用いた例として有名である。

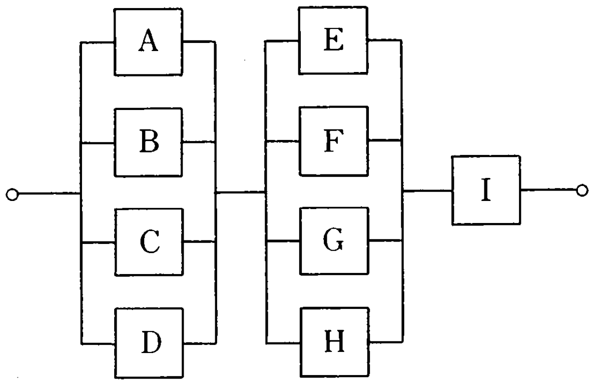


図 2　直並列系

■FTA による解析例　いま，左ページにある FTA 図を直並列系に書き直すと，図 2 のようになる。

ここで，各事象の生起確率を，それぞれ

$$p_A = p_B = 0.1,\quad p_C = p_D = 0.05$$
$$p_E = p_F = 0.2,\quad p_G = p_H = 0.1$$
$$p_I = 0.2$$

とすれば，トップ事象の生起確率は

$$P = \{1-(1-0.1)^2(1-0.05)^2\} \times \{1-(1-0.2)^2(1-0.1)^2\} \cdot 0.2 = 0.0259$$

となる。

FTA はその形状より，品質管理の特性要因図によく似ていることも注目される。参考文献〔25〕〔49〕〔51〕

(真壁)

MRP

Material Requirements Planning の略称で，生産と在庫を一貫して管理するため，必要資材を必要な時期に合わせて用意するようなスケジュール作成方式。必要な時間ごと，たとえば毎週必要となる資材の量を部品ごとに算定し，毎週変化する状況に応じて常に新しい計画を提供する。その中に各部品の発注行動も含まれる。電算機を活用して実現可能となった極めて現場的な計画作成方法である。

★解説

■在庫モデルとの違い　発注点方式などのふつうの在庫モデルでは，

(1) 対象品目の需要は他品目と独立である。

(2) 発注した品物は調達期間後に納入されるが，納入を早めるための督促などは特に考えていない。

(3) 需要は，時間的に見てほぼ一様な割合で発生する。

などの仮定を置いている。しかし，多くの現場では，これらの仮定が明らかに満たされていないため，発注点方式で発注された注文とは無関係な形で現実の調達が行われがちである。この点をふまえ，どんな資材がいつ必要かを常に明確な形で把握しようというのが MRP の基本的立場であるといわれている。

■プライオリティ計画　初期の MRP は発注方式と考えられていたが，むしろ，何をいつ必要とするかを常に把握しているため，資材の発注を適当な時点で行うとともに，その正しい納期を定め，更にそれを最新の情報を用いて訂正することが可能であることに気がついた。その結果，どの仕事を優先して扱うべきかが定まるので，この機能をプライオリティ計画とよぶようになった。

■ MRP の構造　右ページで例示するが，製造予測を基に生産計画，能力計画などをリンクして利用する形をとる。

関連ページ
発注点方式　296

最終製品 A（たとえば自動車）を作るために必要な部品 B（たとえばエンジン）には更に多くの部品が使われている。このとき，A の製造計画（マスタースケジュール）ができると，部品構成表を利用して，その下のレベルの部品 B などの所要量が算定される。これをたとえば週ごとに算定し，それを基に，更に下位レベルの部品 C などの所要量を算出する。そして，手持量と受入確定量を勘案して，その部品の生産命令を出すときを定める。1 回の発注量（ロットサイズ）はここでは B は 60，C は 100 に固定しておく。調達期間もそれぞれ 3 週，4 週としてあるが，これらの量はいずれも可変であるとすることもできる。

右上の表1の上段にはAの製造計画を例示してある。B，C の製造計画はその下に続けて書いてあるが，これは例示のためであって，本来は別々に作られる。そのため，所要量の欄は上の発注量と等しいが別記してある。もちろん複数個の部品を必要とすることもあるので，常に同じとは限らない。

B の所要量は第 1 週で 30 である。現在 40 の手持ちがあるから，これを使っても 10 だけの在庫が残る。第 2 週には 60 の納入があるので 70 が在庫となり，第 3 週に 50 を使って 20 が残る。第 4 週はそのまま推移して，第 5 週で在庫が −5 となることがわかる。そのため，第 5 週で在庫切れを起こさ

週		（現在）	1	2	3	4	5
Aの製造計画			30	0	50	0	25
B	所要量		30	0	50	0	25
	納入量			60			
	手持量	40	10	70	20	20	−5
	発　注			60			
C	所要量			60		60*	
	納入量						(←100)
	手持量	80	80	20	20	−40	(←60)
	発　注	100					

表 1　部品所要量の計算

ないように 3 週前の第 2 週に発注をしておかなければならない。こうして第 2 週に部品 B を 60 発注するという計画が作られる。

部品 C に対しても同様の手順で計算を進めればよい。ここでは，B 以外の部品に対しても部品 C が使われていると考え，そのための需要が第 4 週に 60 あるとしている（＊印でそれを示す）。ここでも，第 4 週で在庫切れを生ずるため，4 週前，すなわち現在すぐに発注すべきことが示されている。

A の製造計画は，近い将来のものほど確度は高いが，遠い将来になるにつれ，確度は低くなり，変更される度合いが増える。B や C にしても何らかの偶発的原因のため，予定された品物の納入が遅れたりすることもある。そのようなとき，直ちに再計算が行われ，計画の手直しがされなければならない。参考文献〔11〕〔59〕　（森村）

MTBF と MTTF

MTBF は Mean Time Between Failures の略で，平均故障間隔という。JIS では，これを「修理系の相隣る故障間の動作時間の平均値」としている。また，MTTF は Mean Time To Failure の略で，JIS では，これを「非修理系アイテムの故障寿命の平均値」と定めている。MTTF は平均故障寿命という。

★解説

MTBF は，修理系が故障を修理して回復し，動作を始めてから，次の故障に至るまでの時間の分布の確率密度を $f(u)$ とすれば，

$$\mathrm{MTBF}=\int_0^{\infty} uf(u)du$$

によって与えられる。また，故障率が時間に関係なく一定のときに，これを λ とすれば，

$$\mathrm{MTBF}=1/\lambda$$

となる関係をもつ。たとえば，エンジンが 5000 時間を稼動し，この間に 4 回故障したとすれば，MTBF は 5000 時間÷4＝1250 時間と推定され，λ は 4/5000 回/時間＝0.0008 回/時間＝0.08 ％/時間と推定される。

MTTF は，アイテム(ここでは部品またはシステムと考えてよい)の寿命分布の確率密度を $f(t)$ とすれば，

$$\mathrm{MTTF}=\int_0^{\infty} tf(t)dt$$

と与えられる。MTTF は通俗的には平均寿命ともよばれる。したがって，アイテムが人間であれば，MTTF は文字通り平均寿命に当たると考えてよい。

アイテムの寿命分布として用いられる**ワイブル分布**のときには，$f(t)=mt^{m-1}/t_0\cdot exp(-t^m/t_0)$ とおいて計算すると，$\mathrm{MTTF}=t_0^{1/m}\cdot\Gamma(1+1/m)$ となる。ここで $\Gamma(\cdot)$ はガンマ関数である。

関連ページ

ワイブル分布
418

■**MTBFの推定法**　いま，ある工場の機械は故障の都度直ちに修復して，これを稼動されていたが，この故障の間隔を測定したところ，次のように記されていたという。

170，83，105，255，72，

320，270　　　　（単位：時間）

これよりMTBFを推定すると，

$$\mathrm{MTBF}=(170+83+105+255+72+320+270)\div 7$$

$$=182.1\ （時間）$$

となる。

■**MTBFと故障率との関係**　故障率曲線は図1の中ほどのところでは，時間に関係なく一定の値λをとる。すなわち，

$$\lambda(t)=\lambda$$

である。

このとき，一定の値λを故障率とする寿命分布は，確率密度が

$$f(t)=e^{-\lambda t},\ 0\leqq t<\infty$$

となる指数分布となる。この分布に故障間隔の長さが従うとすれば，

$$\mathrm{MTBF}=\int_0^{\infty} te^{-\lambda t}dt=1/\lambda$$

となる。すなわち，MTBFと故障率λの間には，逆数関係のあることが分かる。

たとえば，ある航空機のエンジンの

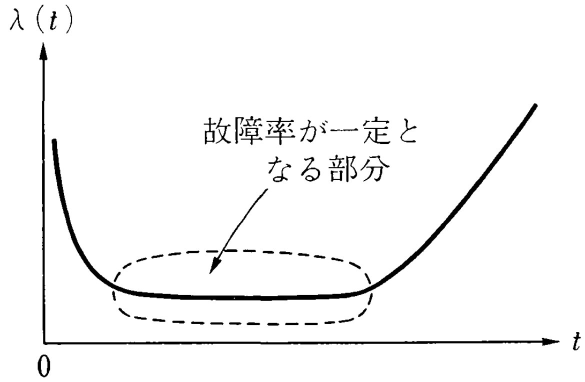


図1　故障率曲線

故障率を0.05％/Hとすれば，このエンジンのMTBFは，

$$\mathrm{MTBF}=\frac{1}{0.05\%/\mathrm{H}}=\frac{1}{0.0005}\mathrm{H}$$

$$=2000\mathrm{H}$$

となる。

■**MTTFとセーフライフ**　MTTFは寿命分布の平均値であるから，寿命分布関数を$F(t)$とすれば，

$$F(t)=\int_0^t f(x)dx$$

であるから，部分積分を用いれば，

$$\mathrm{MTTF}=\int_0^{\infty} tf(t)dt$$

$$=\int_0^{\infty}\{1-F(t)\}dt$$

となることが分かる。一方，セーフライフは，たとえばB_{10}（be ten life）は，

$$\int_0^{B_{10}} f(t)dt=F(B_{10})=0.10$$

となるようなB_{10}の値によって定められる。参考文献〔25〕〔49〕〔51〕

（真壁）

MDS (MultiDimensional Scaling)

MDS（多次元尺度法）は，数量化Ⅳ類や因子分析などと同じく，n 個の要素間の（非）類似度を行列の要素とする（非）類似度行列を入力データとして入力し，類似性の高い要素間が近くなるように n 個の要素に p 次元空間の座標点（布置という）を与える方法である。

類似度が比尺度や間隔尺度で与えられた場合を計量 MDS といい，順序尺度の場合を非計量 MDS という。非計量 MDS は，心理学・社会学・経済学・マーケティングなどの分野において特に有効な手法である。このほか，個人差 MDS がある。

★解説

■概略　地図から3都市 i,j,k 間の直線距離（非類似度）δ_{ij}，δ_{ik}，δ_{jk} を読みとることは容易である。逆に2都市間の距離のみが与えられ，これから都市間の地図を復元することは，都市の数が増えたり測定値に誤差がある場合にはそれほど容易ではない。MDS はこの地図を作成するために都市の座標点を計算してくれる。都市を空間上に位置づけることを布置といい，布置から都市間の潜在構造を探すことを目的とする。布置上の座標点から再度2都市間の距離 d_{ij}，d_{ik}，d_{jk} が計算できる。δ が図1のように計量値の場合を計量 MDS という。この例では，d は δ に等しくなる。

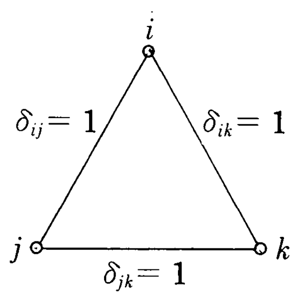


図1　3都市間の直線距離

一方，δ が2要素間の何らかの類似性を表す順序尺度の場合，非計量 MDS が用いられる。この場合には，漠然とした尺度 δ や $\delta_{ij} \neq \delta_{ji}$ のような非対象な距離が与えられ，出力として計量値 $d_{ij}=d_{ji}$ が求まる。心理学などの精密な計測データが得られない領域で，非計量値を計量値に変換（数量化）してくれる有効な手法である。

関連ページ

質的データの解析　162

林の数量化理論　298

数量化Ⅲ類　194

数量化Ⅳ類　196

因子分析　26

SAS　136

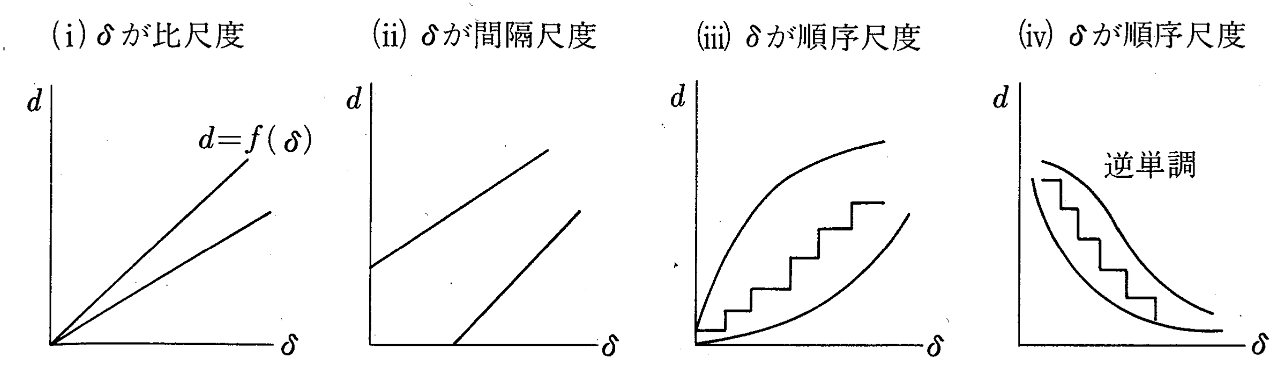


図2　入力データδ(類似度)と布置上の距離dの関係

■アルゴリズム　MDSも数量化Ⅳ類も(非)類似度行列を入力し，要素iの布置(座標点)を計算する。しかし，数量化Ⅳ類は類似度行列の固有値問題で求まる固有ベクトルを布置とするアルゴリズムであるのに対し，MDSは繰り返し計算で布置を求める。このため，MDSの方が計算時間がかかる。

図2は，入力データの類似度δを横軸に，繰り返し計算で求まった布置から計算された距離dを縦軸にプロットした場合の模式図である。

δが比尺度である場合，dとδの関係は関数 $d=f(\delta)=a\delta$ に従うものとする。ここでaは回帰係数である。δが順序尺度である場合，繰り返し計算の基準は次のストレスSを用いる。

$$S=\sqrt{\frac{\sum_i\sum_j(f(\delta_{ij})-d_{ij})^2}{\sum_i\sum_j d_{ij}^2}}$$

このストレス値がある値以下になれば，繰り返し計算は停止する。

左ページの2都市間の直線距離のような場合，関数関係は $d=f(\delta)=\delta$ になる。すなわち，入力データの位置関係がそのまま復元される。δが誤差なく測定されていれば，ストレスSはゼロに収束する。

δが間隔尺度の場合には，関数関係は直線 $a\delta+b$ になる。以上の(i)と(ii)が計量MDSとよばれる。

δが順序尺度でその値が大きいほど非類似性を表している場合，(iii)のような単調増加関数になる。(iv)はδが大きいほど類似性が高い場合に距離は小さくなるので単調減少関数になる。以上の(iii)と(iv)が非計量MDSとよばれる。

■2元MDSと3元MDS　商品間の類似性を100人のパネラーにアンケート調査した結果を，1つの類似度行列にまとめ分析する手法が2元MDSである。3元MDSはパネラーごとに類似度行列を作成し，これを分析する。個別の布置は標準的な布置の座標軸を拡大縮小したものになる。2元MDSのプログラムにはM-D-SCAL，3元MDSではINDSCALがある。SASに収録されているALSCALは両方を扱うことができる。　(新村)

ORの実施理論

ORの理論・解法などの定量的側面の発展は特に著しいものがあり，ある領域ではすでに多くの成果をあげ，またわれわれの気づかない形で実務の中に定着しているものが多い。しかしより広い，多部門間にまたがる，または戦略レベルの応用においては数多くの問題点を残し，まだ十分活用されるレベルまでには至っていない。特に問題となるのが，管理者とOR研究者の関係改善という行動科学的・非定量的側面の理解不足にあるといわれている。ORの実施理論は，OR活動を成功に導くための非定量的側面からの考察を行うものである。

★解説

ORの使命は，他の経営科学的諸手法と同じく，業務，管理・経営の変革・改善をOR特有の高度な方法論を用いて行おうとするものであるが，ORの実施論はORの問題を**変革への抵抗**あるいは**分析への抵抗**(resistance to change, resistance to analysis)という形で問題を受けとめている。

OR実践のためには，単に定量的技法だけでなく，組織や意思決定の基本的過程，技術革新が組織へ及ぼす影響，および管理者と分析者(OR研究者)の置かれている仕事の性質や状況の相異など深く理解することが必要である。ORの実施論はさまざまな観点からさまざまな説が提案されているが，大別すると，①管理者によるORの中身及びOR活動そのものに対する理解不足が問題だとする説，②ORの分析結果の売り込み失敗説あるいは売り込み不足説，③トップ・マネジメントのOR活動への支持・参加が不十分であるとする説，④管理者の革新への態度，組織風土がOR活動の成否を左右するとする説，⑤管理者による政治的行動障害説，がある。①から⑤への順番に政治的要因を強調している。

関連ページ

行動科学的意思決定論　112

分析に対する行動科学的知見　340

■ORの実施問題　ORや統計などを含めた分析的手法のめざましい理論的(定量的側面の)発展に比して，ORの幅広い適用ならびに成果という面で不十分であるという指摘がこれまでにもしばしばなされてきた。実際的問題にORを適用する際に発生する諸問題を，行動科学的・組織論的な観点(非定量的側面)から論ずる分野があり，**実施論**(implementation theory)とよばれている。

■OR実践における2つの誤謬　OR実践において発生しうる可能性のある誤謬(error)には，大きく分けて2つ考えられる。1つは**科学的誤謬**(scientific error)で，もう1つは**行動的誤謬**(behavioural error ないしは political error)である。科学的誤謬とは，モデル構築ならびにモデル解析などの誤りで，主として手法の採択ならびに計算上の誤りなどである。すなわち定量的アプローチの未熟さからくる誤謬である。実施論が主として問題にしてきたのは，後者のタイプの誤謬で，いかに方法や計算が適切であっても，分析の推進の仕方いかんによって失敗する可能性がある。すなわち，意思決定者を含めたさまざまな関与者との間の不十分なコミュニケーションや結果の提示(presentation)の仕方などに起因して発生する誤謬である。いわば**もって行き方論**である。

■部門内ORと部門間OR　OR問題が1つの部門内に局限された，局所的・技術的な問題に適用される限りは，ORは比較的実践しやすく，このような**ORを部門内OR**(intra-divisional OR)という。これに対して問題が部門間にまたがり，OR的分析結果が著しく部門間関係のあり方に影響を及ぼしたり，情報のタイプや流れを変更したり，強いては権限関係に変化が生じたりする場合も多い。このような部門間にまたがる問題を扱うORを**部門間OR**(inter-divisional OR)という。この際，部門間の**調整**(coordination)が適切になされなければならない。

■行動科学的・組織論的考察の重要性

OR活動や分析の結果は単に技術的合理性だけに影響するものでなく，広く仕事の中身や組織上の問題へと波及していくことが多い。このような問題に対して行動科学は寄与する。以上のような意味から，意思決定者(マネージャー)と分析者の関係(「分析に対する行動科学的知見」の項参照)や意思決定の実態論(「行動科学的意思決定論」の項参照)など，OR・分析活動における行動科学的知見は，定量的理論のように，具体的で確実な知識や手法・手続きを提供するまでには至っていないが，OR実践の際に考慮すべきチェック項目にはなりうるであろう。

■OR実施論の出現　ORの実施問題は，1965年のチャーチマン(C. W. Churchman)とシャインブラット

(Scheinblatt)の論文(Management Science誌)において提起され，1つの研究分野にまでなった。これを**実施論**(または実施理論)という。チャーチマンらは，OR関係の論文を6年間にわたって調査・分析し，その結果OR研究結果の採用の形跡も少なく，また利用されたモデルも現実への影響もほとんどないとし，その理由を模索した。その影響を受けて，行動科学者や組織論者がこの問題をとりあげるようになり，ORに限らず経営科学的技術全般(たとえば，コンピュータ利用技術も含めて)の実施問題へと広がっていった。

■相互理解と説得　チャーチマンらは，管理者(Manager，以下Mとする)とOR研究者(Researcher，以下Rとする)の間の関係を以下のような図式を用いて考察した。MとRがお互いの仕事の内容やスタイル，置かれている立場，状況の異同を理解しあっているときを**相互理解**(mutual understanding)といい，RがMをよく理解し，MがRを理解しないとき必要となるのが，**説得**(persuation)であるとした。逆にMがRをよく理解し，RがMをよく理解してない(たとえばMのかかえている真の問題を正確にとらえていないとか，Mの置かれている状況をRのそれと異ならないと考えたりする)とき必要となるのが，**コミュニケーション**であるとし，両者が互いに相手を理解しない状態を，**分離された機能**(separate functionalist)とよび，この場合は両者の一致はありえないとした。OR活動がよい結果を出すには，相互理解が理想であるとした。

■実施論の基本的スタンス　実施論の立場は，ORの方法は基本的には間違っておらず，多くの場合分析結果はきわめて優れており，ORの実施がうまくいかないのは，むしろM側の態度その他に問題があるとしている。すなわち，実施論は行動的誤謬のみを問題と考える。

■OR実施論の代表的見解

(1)　**MによるOR理解不足説**　OR活動の失敗の主要因は，MがOR活動ならびにORそのものの内容を十分に理解していないことからくる説。ORは手法が高度で，抽象的(モデルの利用，現場にある人々からはモデルという概念が分かりにくいとよくいわれている)なため，十分な専門的訓練を受けていない人々には分かりにくいとされてきた。Mは内容の理解できない解答よりは，解けない問題の方を好む傾向があるという調査もあり，理解できない解答を提示されても，それを採用する気になれないとするものである。チャーチマンの**相互理解**はこの説に立つ。

(2)　**売り込み不足説**　分析結果の提示(presentation)の仕方がまずかったり，分析結果の売り込み不足が，OR活動の失敗の主要因であるとする説。

チャーチマンの**説得**はこの立場に立つ。この説は，分析結果の形式的報告で事足れりとしたり，部厚過ぎる報告書，難しい報告書などをいましめ，発表技術の工夫の必要性を強調する。

(3) トップ・マネジメントの不参加説 OR 活動にどれくらいトップが熱心か，および M が問題の設定，モデル構築，分析結果の検討などにどれくらい深くかかわってくるか(management involvement の程度)が，OR 活動の成否を決めるという説。トップの参加度と OR の成功度の間に正の相関があるとする調査報告もあり，トップの OR 活動に対する明確な支援と OR プロジェクト推進のための財政的援助などが重要だとし，チャーチマンの二者関係(M–R 関係)モデルを拡大し，トップ・マネジメントも含めた三者関係の関係改善を強調する。

(4) 組織風土説 組織や M が，革新的なこと，新しい仕事のやり方に積極的な(未来指向型組織とか，事実的データを尊重する風土であるとかの)組織風土(organizational climate)ないしは態度(management attitude)を有しているかどうかが，OR 活動の成否を決めるとする説。

(5) 管理者の行う政治的行為による阻害説(政治悪説) 分析結果の利用を台なしにするのは，M の行う政治的行為(politics)にあるとする説。この説は，組織とは①自己の権力の増大のみを画策する人々(管理者)の集団であること，②したがって M は組織目標より，部門目標ないしは自己の目標を追求する人々からなる，とする組織観に立つ。すなわち組織＝権力ゲーム(power game)装置とみて，M はそのプレーヤーであるととらえる。この立場では，分析結果は政治の道具としてしか取り扱われず，M にとって不利(たとえば分析結果を採用すると自分の権限の範囲が縮小される懸念があるなど)ならば，それが組織にとってよいものであっても拒否されるとする。

■実施論の問題点 実施論の基本的スタンスは，OR の理論や方法論には間違いはなく(科学的誤謬はなく)，悪いものはもって行き方にあるとするものであるが，ややもするとわれわれ OR 研究者はちゃんとやっているのに，それをダメにするのは M であるとする M 批判(政治悪説など)になりやすい。基本的には M の協力が必要とされるが，それ以上に R の努力が要請される。特に R が数理的・理論的専門家であると同時に，非定量的な側面(行動科学的技法)にもある程度精通すること(**両刀使い**)が要請される。また近年の傾向として，システム技法や OR 技法などが人間を含むシステムをアプローチするには，方法が堅すぎるとして，方法そのものをもっとソフトなものにすべきだとする意見が増大している。 (中野)

オペレーションズ・リサーチ

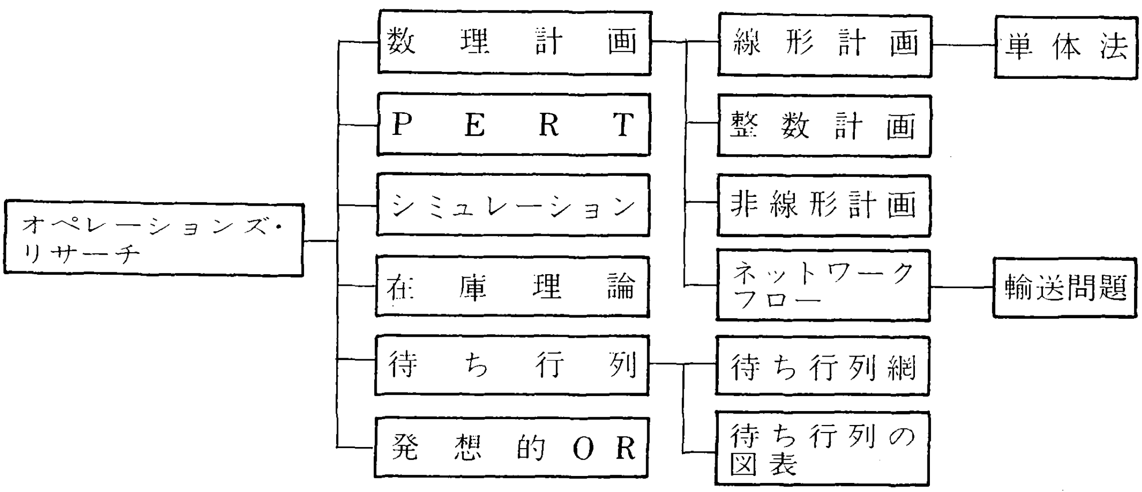
システムの運用(operations)に関する最善の方策を見出したり，計画を立てたり，あるいは何らかの決定を下すための判断資料を提供する科学的方法。JIS の用語(Z－8121)では，「科学的方法および用具を体系の運営方策に関する問題に適用して方策の決定者に問題の解を提供する技術」と定義し，運営研究という訳語も与えている。OR と略す。

★解説

言葉の起こりや簡単な歴史およびその特徴については右ページに述べる。また下図に示すように，OR にはいろいろな手法がある。これは，OR を進めるうちに，似たような種類の問題に多く出合い，それらを解決するのに適した数学的方法が次第に体系化されたものである。

現在のところ，大別すると，①確定的，②確率的，③戦略的の 3 つに分けられ，このほか**発想**を助けるための手法も利用されている。①の分野では，**数理計画法**(特に**線形計画法**)，**PERT**，**スケジューリング**，**経済性工学**など，②の分野では，**在庫**，**待ち行列**，**シミュレーション**など，③の分野では**ゲーム理論**などが主なものであろう。また，統計の技法も OR の問題解決に多用されている。

関連ページ
発想的 OR　294
数理計画法　186
線形計画法　218
PERT　292
スケジューリング　280
経済性工学　94
在庫　128
待ち行列　366
シミュレーション　166
ゲーム理論　102



注／手法とよばれるものはこのほかにも多いが、ここでは本書の項目としてあげられるものに限って記してある。

図 1　OR 関係事項の関連図

■言葉の起こり　1935年，イギリスで防空のためのレーダー・システムが開発され，実験室ではうまくいったのに演習で失敗した。このため，レーダーの操作について研究する必要が認識され，その研究に対して operational research という名が与えられた。

■歴史　第二次世界大戦中，上記研究グループのメンバーであったブラケット(P. M. S. Blackett, 1897～1974, 後にノーベル物理学賞を受賞)の率いる科学者グループが，軍隊の作戦に関するさまざまな問題の解決に活躍した。それに刺激されてアメリカでも同種の活動が盛んになり，ORの名が定着した。

戦後，アメリカORの中心であったモース(P. M. Morse, MIT物理学教授)が本(Mathematical Methods in Operations Research)を書き，企業での活用を提唱した。日本ではこの本の講読が導入の始まり(1952)といわれている。企業における活用とあいまって，operationsという言葉はシステムの運用を指すようになり，近年は，計画の面が一層重視されるようになった。

■特徴　歴史上，さまざまな専門家を集めてORチームが作られた経緯もふまえ，学際的であることをORの第一の特徴にあげる人もある。しかし，現状では，①システム全体としての「良さ」を追求する，②科学的，特に数理的モデルを利用してアプローチする，③何らかの決定のための判断資料となることを目標とする，の3点が主な特徴と考えてよいであろう。もちろん，解析にあたっては，衆知を集めたり，専門家の持つノウハウをよく吸収したり，データを利用して客観的に評価することも大切である。

■よいモデル　OR活動の決め手は，必要な時に，必要な人に，正しい解を提供することである。そのためにはよいモデルを作って解析することと，一度作ったモデルをよりよく使いやすいものにする努力が必要になろう。

よいモデルとは，使いやすく，かつできるだけ大勢の人に理解されやすいものでなければならない。それで，モデルの3s(small, simple, steady)を心がけるとよいといわれている。第一の小さい，ということは規模のことで，第二の「単純」ということも当然関係するが，本質を損なわない限り，できるだけ細かい条件を無視して，簡単かつ安価に計算が繰り返せる程度に単純化したものが望ましい，ということである。それは，モデルの中に当然含まれているパラメータをいろいろ変えてその結果を見ながら方策を考える必要があるためである。第三のsteadyとは少々条件が変わっても結果が変わらないことを指している。また，モデルの中で精粗まちまちであるのも感心しない。よいモデルはバランスのとれたものである。　(森村)

GERT

Graphical Evaluation and Review Techniques の頭文字を並べたことばで，第一の語以外は PERT と同じであることからもうかがえるように，PERT の拡張として考えられた手法である。PERT においては矢線図上の仕事はすべて行われなければならないが，複数の仕事のうちどれか 1 つが実施されればプロジェクト自体は完成するという不確定要素のはいった場合を取り扱うことを目標とする。

★解説

■GERT の特徴　矢線図についていえば，①ノードの種類が複数個あり，たとえば，あるノードに至るアークのうちでただ 1 つが実現すればそのノードに達したこととなり，その他のアークで表される仕事はその時点で打ち切られるとか，そのノードに至るすべてのアークが実現したとき，そのノードに到達したことになるとかの区別がつけられる。②上のアークの実現確率やそのアークに付随した時間や費用を与えるが，多くの値は確率変数であってもよい。③フィードバックのループを入れることが許される。④あるアークが実現して 1 つのノードに達したとき，その影響が後に残って，その後のパスが規定されることも一部では許される。こういう矢線図を扱うので，PERT とは本質的に異なる構造を扱うことになる。

■ネットワークの解析　右ページの図 1 に例示するように，何本かのアークに分かれているネットワークをより簡単なネットワークで置き換えても，ノードの実現確率や期待時間などは変わらない。このような置換を通し，基本的な 3 つの型のネットワーク要素によって全体のネットワークを構成し直した後，積率母関数を利用して，つぎつぎに等価な母関数を求める方法で，全体のネットワークにおける期待時間などを計算することができる。

関連ページ
PERT　292

ネットワークモデル　284

■ノードの種類　GERT Iでは入力側3種，出力側2種を組み合わせてできる6種のノードを考えていたが，GERT IIではこれを3種類に統合する代わりに，パラメータを増やして表現力の増大を計っている。しかし，構造的には前者の方がわかりやすいので，それについて解説する。

入力側の第一のタイプはexclusive-orとよばれ，そのノードに至るアークのどれか1本のみが実現したときに限り，そのノードが実現する。その時点で他のアークが表す行動は打ち切られる。第二のタイプはinclusive-orとよばれ，そのノードに至るアークのいずれかが実現すればそのノードは実現する。もちろん，同時に2本以上のアークが実現しても差し支えない。

第三のタイプはandとよばれ，そのノードに至るすべてのアークが実現して初めてそのノードは実現する。

たとえば，入札で品物を購入するという業務で，購入を示すノードに至るアークが各社の納入を意味するならば，納入を実現するのはどれか1社に限られるので，exclusive-orがこのノードの入力側の性格を与える。これに対し，複数の技術開発が並行して進められ，いずれが成功しようとも次の開発に進めるのであれば，inclusive-orで表現されることとなろう。

出力側は確定的と確率的の2種である。前者は，このノードから出るすべてのアークは実現することを意味する。後者はどれか1つだけが実現するという意味である。

図1のノードSは出力側のみ確率的，ノード1は入力側がinclusive-orで出力側が確定的であることを示している。また，ノード3はexclusive-orとなっている。

■例題　ある製品を作るのに残された技術開発がある。定められた期限まで，2つのプロジェクトを同時に進めて開発したい。いずれか一方が成功すれば製品化にかかれるが，どちらも成功しないときには，他社の特許を買って製品化する。この仕事をGERTで表現したのが図1である。ノード1は自社開発，ノード2は技術導入，ノード3は製品化を示している。

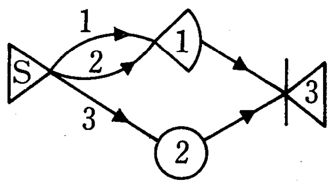


図1

■ネットワークの変換　図1でアーク1, 2, 3の実現確率をそれぞれp_1, p_2, p_3, そのときの費用をc_1, c_2, c_3とおく。ノード1の実現確率はp_1+p_2，期待費用は$c_1p_1+c_2p_2$であるから，アーク1, 2を1本のアークで置き換え，そのアークに$p=p_1+p_2$, $c=c_1p_1+c_2p_2$の値を与えれば，等価なネットワークが得られたことになる。このような考えで，より簡単なネットワークに順次変換する。　(森村)

回収期間法

投資をリターンで回収しきる期間(→268ページ)により，投資案の合否を判断する方法である。基準になる回収期間を設定し，その期間より短い回収期間の投資案を合格としたり，回収期間のより短い投資案をよしとする考えに基づく方法である。

回収期間法は投資案から得られる正味の得(利益)を大きくすることをねらうものではなく，資金繰りがきびしい場合や将来のリターンの獲得が非常に不確実な状況のもとでの安全性をねらった方法である。

★解説

投資案の回収状態を年度の経過でみるには，年度別の回収残高を計算してみるのがわかりやすい。たとえば，1000万円を投資し，5年にわたり毎年末に600万円のリターンがあがる投資案についてみてみよう。$i=10\%$ とする。

年度	回収残高の計算(万円)
0	-1000
1	$-1000\times(1+0.1)+600=-500$
2	$-\ 500\times(1+0.1)+600=\ 50$
3	$50\times(1+0.1)+600=\ 655$
4	$655\times(1+0.1)+600=\ 1320$
5	$1320\times(1+0.1)+600=\ 2052$

この計算は，第 k 年度のリターンを R_k，回収残高を S_k とおくと，次式で行ったものである。

$$S_k=S_{k-1}\times(1+i)+R_k$$

最終年度の回収残高は**正味終価**になる。

回収残高の計算より，その値がはじめて正になる年数，上記の例では2年，が回収期間になる。この案では，3年以後のリターンが予想通りあがらなくても，2年でもとがとれることになる。

関連ページ
投資案の経済性指標　268
投資案の感度分析　266
正味終価　268

回収期間法の性格を数値例でみてみよう。

案	投　資	リターン	寿命
A	1000 万円	600 万円	5 年
B	5000 万円	1800 万円	5 年

この表の A 案は左ページの数値例である。A 案の年度別回収残高はすでに求めてあるので，B 案のそれを計算する(i=10% のもとで)。

年度	回収残高
0	−5000
1	−5000×(1+0.1)+1800=−3700
2	−3700×(1+0.1)+1800=−2270
3	−2270×(1+0.1)+1800= −697
4	− 697×(1+0.1)+1800= 1033
5	1033×(1+0.1)+1800= 2937

この結果，両案の回収期間と正味終価は次のようになる。

案	回収期間	正味終価
A	2 年	2053 万円
B	4 年	2973 万円

このように，回収期間が短い投資案がかならずしも正味の得(利益)を大きくするわけではない。

もし，基準になる回収期間が 3 年と設定されていると，A 案は合格で B 案は不合格になる。この判断は，3 年後の回収残高(正味終価)の値でもできることに注意しておこう。

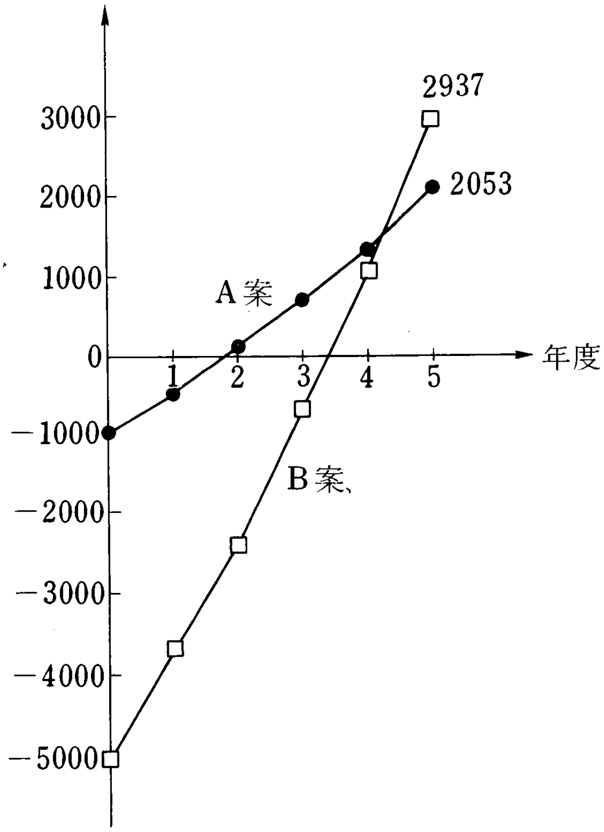


図 1　回収残高の推移

投資案の各年度における回収残高は，その年で投資の効果が停止したもとでの投資から生じる正味の得を意味している。回収残高の推移を図 1 のようなグラフに示してみると，投資案の特徴が見やすくなる。

このグラフからわかることは，A 案の最終的な利益は B 案よりも小さいが，初期投資の規模は B 案よりも少なくてすみ，かつ 1 ～ 4 年間の範囲でみれば利益も B 案より有利になっていることである。

このように，回収期間法は資金繰りがきびしいとき，あるいは将来の不確実性が大きいときの判断指標として役立つものである。　(中村)

確率抽出法

標本調査の目的は，母集団の一部分を標本として抽出し，それに対する調査・実験結果に基づいて，母集団に関する推測を行うことである。したがって標本調査では，標本抽出の方法が問題になるが，ふつうよく用いられるのは確率抽出法である。この方法によると，母集団の特性がよく反映され，また標本値から算出される各種の推定値のよさが，客観的に評価されて都合がよい。

★解説

標本を選ぶとき，それが**母集団**の正確な縮図になっていることが望ましい。しかし，適当に選んだのでは，その評価が難しくなる。そこで，標本に基づく推論の過程を客観性あるものにするために，ふつう次の原理に基づく抽出法を用いる。それは，母集団を構成しているすべての単位に対して，それがある確率で(必ずしも等確率でなくともよい)標本に選ばれるようにすることである。このようにして得られた標本のことを任意標本，無作為標本，またはランダム・サンプル(random sample)という。これに対して，**有意抽出法**による標本のことを有意標本という。

また，抽出の仕方に，復元抽出と非復元抽出との別がある。母集団を構成している，ある特定の単位が重複して標本に選ばれることを許す抽出方式が**復元抽出**であり，重複を許さないのが**非復元抽出**である。

母集団を構成している単位のことを**母集団の大きさ**といい，それから抽出される標本の数のことを**標本の大きさ**という。母集団の大きさNが十分大きな値であるとき，標本に基づく統計的推論を行うには，その標本が復元抽出によってとられたものであるかどうかを，あまり問題にしなくてよい。しかし，Nがあまり大きくないときには，復元抽出か非復元抽出かによって，異なった結果が導かれる。

関連ページ
標本　260
母集団　260
有意抽出法　259

■枠　標本抽出を行うには，ふつう抽出単位を設けておき，そのリスト(またはこれにかわるもの)を用いる。これを枠という。

■単純無作為抽出法　確率抽出法のなかで，もっとも基本的なものが，単純無作為抽出法である。これは，母集団を構成しているすべての単位に対して，等しい確率を与えて標本を抽出する方式である。したがって，単純無作為抽出法によって標本を抽出するには，原理的には，1つ1つの標本を，その都度，乱数を発生させて求める。しかしこの操作は面倒なので，最初の標本だけを乱数によって抽出し，あとは等間隔で抽出する，という**系統抽出法**によって代用することが多い。

■層別抽出法　標本を用いて母数を推定するにあたり，できるだけ精度のよい推定をしたいわけであるが，推定精度は，標本の大きさと母分散の値とに依存する。標本数が大きいほど，よい推定がなされるのは当然であるが，母分散についていえば，それが小さいほどよい。ところで，母分散の大小は，母集団を構成している単位の異質度の大小であるということができる。そこで，母集団をなるべく等質的ないくつかの集団に分割して，その集団ごとに標本を選ぶという方法が，よく用いられる。このとき，分割された集団のことを**層**といい，層ごとに標本をとる抽出の仕方を層別抽出法という。層別抽出法を用いると，層別しないときよりも，一般に推定の精度が向上する。

確率抽出法
- 単純無作為抽出法
- 層別抽出法
- 集落抽出法
- 多段抽出法
- 確率比例抽出法
- その他

表1　確率抽出法

■集落抽出法　母集団を構成する単位のいくつかを含む大きい単位のことを**集落**といい，これを抽出単位とする抽出の仕方を集落抽出法という。集落抽出法による調査では，はじめにいくつかの集落をランダムに抽出し，つぎにその集落内の単位を全部調べさせるという方法をとる。

■多段抽出法　母集団を，抽出単位であるいくつかの集落に分け，最初にこのような集落からいくつかの集落を抽出する。つぎに，本来の調査個体を抽出する。このような抽出法を**2段抽出法**というが，一般にそれが多段になっているとき，**多段抽出法**という。このようにすると，第1段階で抽出された集落についてだけリストをつくればよいので，コストや手間が軽減される。

■確率比例抽出法　たとえば，大きさの不揃いな抽出単位から標本抽出をする場合，その大きさに比例する確率で抽出すると，一般に精度が向上する。これを確率比例抽出法という。(牧野)

確率変数と確率分布

われわれは偶然現象を調べるとき，偶然に左右されて定まるある量と，その値の生じやすさの程度に関心がある。考察の対象とする偶然現象の観測，調査や実験を行うことを**試行**とよび，試行の結果，値(実数値)が定まるような変量を**確率変数**という。その確率変数のとりうるすべての値に確率を対応させたものを**確率分布**とよぶ。

★解説

確率変数には大きく分けて，1カ月間の計算機の故障回数のようにとりうる値が**離散的**な場合と，部品の寿命のようにとりうる値が**連続的**な場合とがある。確率分布を，それぞれの場合に応じて扱いやすいように表現する。

■離散的な場合　確率変数Xのとりうる値が a_1, a_2, …, と番号をつけられるとき，Xが値xをとる確率

$$p(x)=\mathrm{P}\{X=x\} \qquad (x=a_1,\ a_2,\ a_3,\ \cdots)$$

をXの確率分布という。$\mathrm{P}\{\cdot\}$ は $\{\cdot\}$ となる確率を意味し，$(x=a_1,\ a_2,\ \cdots)$ はXのとりうる値を明示する。確率分布の総和は1である。とりうる値が $x=0, 1, 2, \cdots$ と整数値のとき p_x と書き表す。**離散分布**の項を見よ。

■連続的な場合　ある部品の寿命をXとすると，Xがちょうどある値xをとることはありそうもない。つまり，どんなxに対しても，$\mathrm{P}\{X=x\}=0$ と考えてよい。この場合，上のやり方で確率分布を定めても意味がない。しかし，Xがある範囲 $(a,\ b)$ の間に入る確率は定まる。その確率に対して，

$$\mathrm{P}\{a<X<b\}=\int_a^b f(u)du$$

と表される $f(x)$ をXの**(確率)密度関数**とよぶ。右辺は図2の斜線部の面積を表す。$f(x)$ は“確率”ではないが，xの生じやすさの程度を表す。$f(x)\Delta$ はXが微小区間$(x,\ x+\Delta)$ に入る確率と考えてよい。$f(x)$ とx軸で囲まれる全面積は1である。**連続分布**の項を見よ。

関連ページ

離散分布　404
連続分布　410

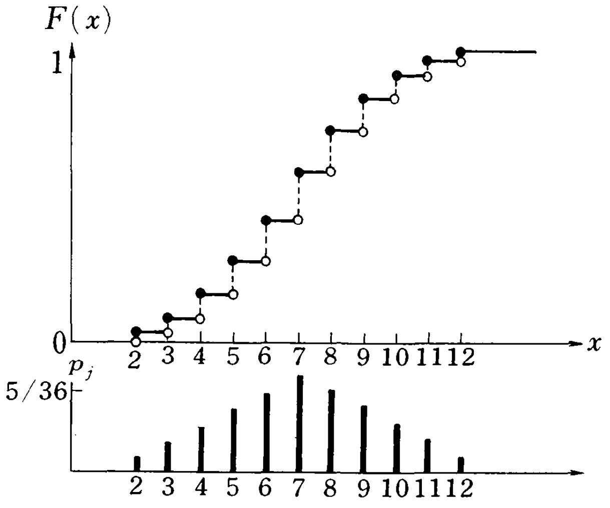


図1　離散的な場合の分布関数と確率分布

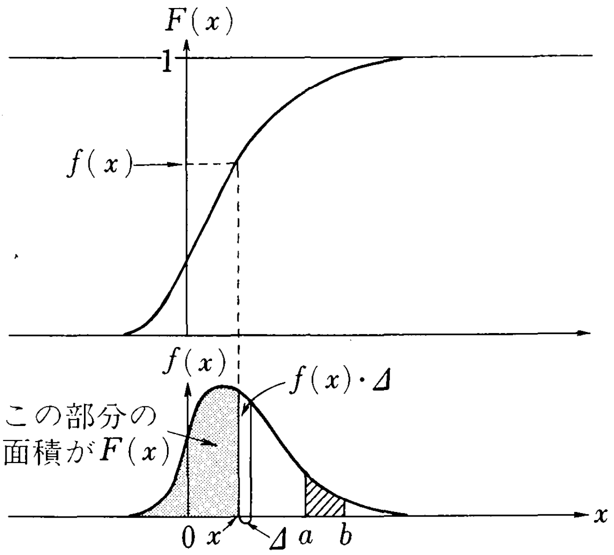


図2　連続的な場合の分布関数と密度関数

■分布関数　確率変数Xがx以下の値をとる確率，すなわち，

$$F(x)=\mathrm{P}\{X \leqq x\}$$

をXの**分布関数**とよぶ。離散的な場合は，図1のように，$F(x)$は確率分布$\{p(a_i)\}$の$a_i \leqq x$の範囲での和である。連続的な場合は，$f(x)$とx軸で囲まれる領域のx以下の範囲の面積である（図2)。逆に$F(x)$から$f(x)$も求まる。

また，分布関数は，

① xに関して非減少

② $F(-\infty)=0,\ F(+\infty)=1$

という性質をもつことは明らかだろう。

ここでは1つの変数Xの確率分布を考えたが，2つ以上の変数を同時に考える場合は**多変数の分布**(→228ページ)を見よ。

(例1)　2つのさいころの目の和の分布　目の和をXとすると，Xのとり得る値は，2, 3, …, 12 であり，その分布は

$$p_2=p_{12}=1/36,\qquad p_3=p_{11}=2/36,$$
$$p_4=p_{10}=3/36,\qquad p_5=p_9=4/36,$$
$$p_6=p_8=5/36,\qquad p_7=6/36$$

である。これを1つの式にまとめて，

$$p_i=(6-|i-7|)/36\ (i=2, 3, \cdots, 12)$$

と書いてもよい。

この分布関数$F(x)$は次のように求まる。$x=4.2$とおくと，$i \leqq 4.2$の範囲での$\{p_i\}$の和，つまり，

$$F(4.2)=p_2+p_3+p_4$$
$$=1/36+2/36+3/36=6/36$$

となる。逆に$F(x)$が分かればp_iも求められる。たとえば，

$$p_5=F(6)-F(5)$$

となる(図1を参照)。

■分布の特性　確率変数Xの確率分布がだいたいどんな形をしているか，どのあたりの値が出やすいか，値の散らばり具合はどうか，など，分布のおよその特徴をとらえることも大切である。

多くの場合，密度関数は図3(a)のように1つの山を持つ(**単峰性**という)。ときには，図3(b)のようにいくつか山を持つものもある。大学の試験の点数では

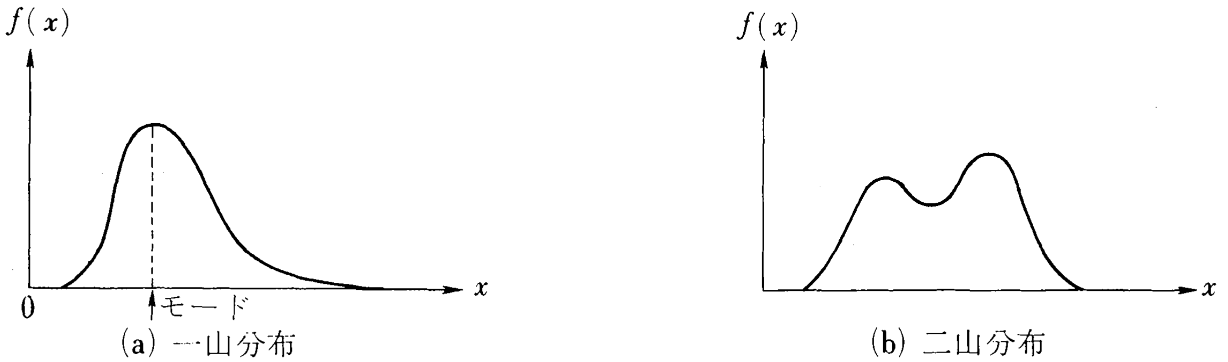


図3　典型的な密度関数

よく2つの山が表れる。勉強するグループとしないグループとに分かれるためである。いちばん峰の高い位置をその分布の**最頻値**(**モード**)とよぶ。

分布の中心的位置を表す量として図4のように，左右50％ずつ分ける**中央値**(**メディアン**)がある。さらに細かく25％ずつに切り分けた**四分位**が分かると，分布の広がり具合がおよそつかめる。これらと同様な情報を与えるものに，次に述べる平均値と分散がある。

■**平均値**　テストが行われると，答案を返すとき平均点が発表される。N人のクラスの平均点は，各自の成績をx_1, x_2, …, x_N とすると，

$$\overline{x}=(x_1+x_2+\cdots+x_N)/N$$

と計算される。i点をとった人の数をf_iとすると，これは

$$\overline{x}=(1\cdot f_1+2\cdot f_2+\cdots+100\cdot f_{100})/N$$
$$=1\cdot\frac{f_1}{N}+2\cdot\frac{f_2}{N}+\cdots+100\cdot\frac{f_{100}}{N}$$

と計算してもよい。f_i/N はi点をとった人の割合である(→344ページ)。

これと同じように，確率変数Xの**平均**(**値**)も定義できる。Xが離散分布$\{p(a_i)\}$ に従う場合，その平均値は，

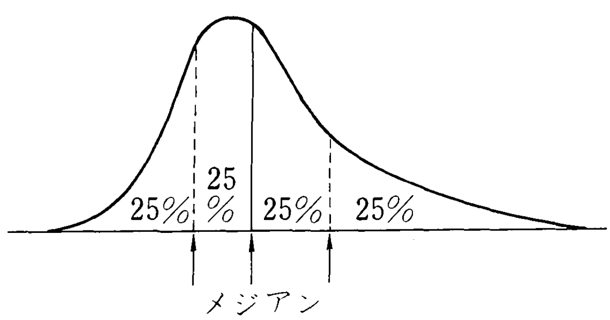


図4　分布のメジアン，四分位

$$E(X)=a_1p(a_1)+a_2p(a_2)+\cdots$$

と定義される。このように，Xのとりうる値にその確率をかけて和を求める計算を，Xの**期待値**をとるという。

Xが連続的で密度関数 $f(x)$ をもつとき，平均値は，

$$E(x)=\int_{-\infty}^{\infty}x\cdot f(x)dx$$

と計算される。上と同様にXのとる値xに，xをとる確率 $f(x)dx$ をかけて和をとっていることになる(図5)。Xの平均値 $E(X)$ を(原点のまわりの)**1次モーメント**ともいう。太さの均質でない棒の重心(あるいはモーメント)を求めるのと同じやり方である。

■**分散**　平均値は分布の中心がだいたいどのあたりにあるかを示しているが，その周りに値がどのように散らばっているかを見る量として分散がある。平

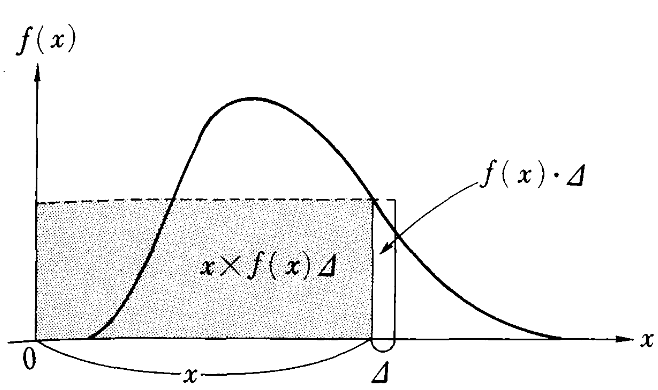


図5　連続的な場合の平均値の計算

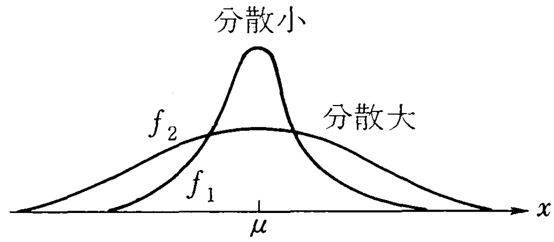


図6　分散の大小

均値を μ とおくと，**分散**は，

$$\mathrm{Var}(X)=E((X-\mu)^2)$$

と定義される。これは

$$\mathrm{Var}(X)=E(X^2)-\mu^2$$

と計算してもよい。$E(X^2)$ を（原点まわりの）2次モーメント，分散を μ のまわりの2次モーメントともいう。

X が離散的なときは，分散は

$$\mathrm{Var}(X)$$
$$=(a_1-\mu)^2p(a_1)+(a_2-\mu)^2p(a_2)+\cdots$$

あるいは

$$=[a_1^2p(a_1)+a_2^2p(a_2)+\cdots]-\mu^2$$

と計算する。連続的なときは

$$\mathrm{Var}(X)=\int_{-\infty}^{\infty}(x-\mu)^2f(x)dx$$

あるいは

$$=\int_{-\infty}^{\infty}x^2f(x)dx-\mu^2$$

と計算する（→332ページ）。

分散は平均値からのずれを2乗しているので，図6の2つの密度関数を比べると，μ のまわりに値の集中している f_1 の方が分散は小さい。またディメンジョンをもとの測定単位と同じにした $\sigma=\sqrt{\text{分散}}$ を**標準偏差**とよぶ。X が非負値をとるとき，バラツキの大きさを平均値と比べた**変動係数** CV

$$CV=\text{標準偏差}/\text{平均値}$$

を調べることも役に立つ。

■期待値と分散の性質　2つの確率変数 X，Y を考える。このとき，任意の定数 a，b，c に対し，期待値は

$$E(aX+bY+c)=aE(X)+bE(Y)+c$$

という性質をもつ。つまり，期待値は**線形性**をもつ。さらに，もし X，Y が**独立**（→229ページ）ならば，分散は，

$$\mathrm{Var}(aX+bY+c)$$
$$=a^2\mathrm{Var}(X)+b^2\mathrm{Var}(Y)$$

と計算できる。

（例2）　例1の分布の平均値は，

$$E(X)=2\cdot p_2+3\cdot p_2+\cdots+12p_{12}$$
$$=2\cdot 1/36+3\cdot 2/36+\cdots+12\cdot 1/36=7$$

となる。分散は，

$$\mathrm{Var}(X)=(2-\mu)^2\cdot p_2+\cdots$$
$$+(12-\mu)^2\cdot p_{12}$$
$$=5^2\cdot 1/36+\cdots+5^2\cdot 1/36=35/6$$

となる。これはまた，

$$E(X^2)-\mu^2$$
$$=(2^2\cdot p_2+3^2\cdot p_3+\cdots+12^2\cdot p_{12})-\mu^2$$
$$=1974/36-7^2=35/6$$

と計算しても同じである。　（森）

家計統計

世帯は生活する上で賃金などの所得を得，食料，衣類などのモノや医療費，タクシー代のようなサービスを購入し，あるいは貯金する。これらの経済行動を「家計」という。国民経済規模の中で家計消費は約6割を占めることから，その水準と動向は重要である。

家計の実態を明らかにする統計には，都市の勤労者世帯と一般世帯(自営業世帯と無業世帯)については家計調査(総務庁統計局)が，農家世帯については農家経済調査(農林水産省)があって毎月調査を行っている。

★解説

家計調査からは，①世帯収入の水準と構造，②勤労所得税や社会保険料負担，③家計消費の水準と構造，④家計収支と貯蓄純増，の諸実態を知ることができる。

また調査は毎月行われているから，収入と支出の月々の増減額，生活費の動向，あるいは地域別・職業別・所得階級別などの格差を知ることができる。とりわけ，消費支出の動きは消費行動をみる指標として使われている。費目別の消費支出額は，消費者物価指数によって実質化され，実質消費の動向を示すものとして使われている。

このほか，つぎの分類や概念が家計収入や支出の特徴を示すものとして多く用いられている。

(1)可処分所得　「実収入」－「非消費支出」をいい，実際に消費や貯蓄に回すことができる所得額。

(2)消費性向　「消費支出」÷「可処分所得」をいい，消費態度の傾向を示す指標として使われている。

(3)エンゲル係数　「食料費」÷「消費支出」をいい，消費生活の豊かさの指標として使われている。

(4)消費水準指数　家族数4人，30.4日(365日÷12カ月)の消費支出額を調査結果から推計し，これを消費者物価指数で割って実質化したもの。

関連ページ

主な経済指数
425

項　目	実　数 (円)	対前年 増加率 (%)	構成比 (%)
実収入	405,517	3.2	100.0
うち　世帯主収入	337,395	3.1	83.2
妻の収入	31,960	7.4	7.9
実支出	333,603	3.1	—
消費支出	272,199	2.3	100.0
食料	72,099	1.5	26.5
住居	12,929	2.6	4.7
光熱・水道	15,774	3.6	5.8
家具・家事用品	11,216	1.4	4.1
被服及び履物	18,910	0.0	6.9
保健医療	6,436	3.0	2.4
交通通信	25,729	7.3	9.5
教育	10,414	4.3	3.8
教養娯楽	23,462	3.1	8.6
その他の消費支出	75,230	1.3	27.6
非消費支出	61,404	6.8	—
可処分所得	344,113	2.6	—
黒字	71,914	3.5	—

資料　総理府統計局調査部消費統計課「家計調査年報」

表1　家計収支(勤労者世帯、58年)

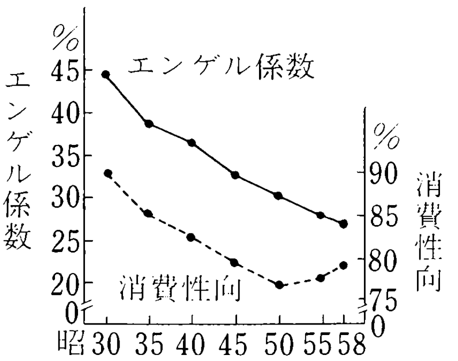


図1　家計指標の推移

家計調査によって，昭和58年の勤労者世帯の1カ月平均の実収入をみると40万6千円で，前年に比べ3.2％増加した。内訳は，世帯主の定期収入が3.1％増であったのに対し，主婦の就業化が進んだため妻の収入は7.4％増と大きく，世帯収入の増加に寄与した。

実収入から税金や社会保障費などの非消費支出を差し引いた可処分所得は，34万4千円で2.6％増となった。非消費支出は，45年以降毎年10％を超える増加が続いている。この結果，実収入に占める非消費支出の割合は年々高まり，45年の8.2％から58年には15.1％となった。

消費支出は27万2千円で2.3％増加した。消費性向の動向を長期的にみると，昭和30年の90.8％から，年々低下を続け，49年には75.7％となったが，以後，上昇に転じ，58年には79.1％となった。58年について消費費目別にみると，交通通信，光熱・水道，保健医療，家具・家事用品，教養娯楽は，前年に比べいずれも実質増加となったが，被服及び履物，食料，教育，住居は物価上昇によってそれぞれ実質減少し，消費構造の変化が進んでいる。

エンゲル係数は昭和30年の44.5から40年36.2，50年30.0，57年26.7と着実に低下し，58年には26.5に低下した。

家計調査は標本調査であるから，職業別や市町村別など標本世帯数が少ない区分，購入回数が少ない品目は誤差が大きいので注意が必要である。また，亭主の小遣いや酒，たばこなどの消費は捕捉が難しいので，過小に計上されているといわれている。　(市野)

仮説検定

母集団の統計的性質に対して，ある仮定を設け，その仮定が正しいかどうかを，有限個の標本によって調べる。このとき，あらかじめ設けた仮定のことを仮説といい，このような推論の仕方を仮説検定という。

★解説

関連ページ

二項分布　276

■仮説検定の考え　「ある貨幣を投げて表の出る確率を p とする。10回投げてみたところ，表が9回出た。$p=1/2$ と考えてよいだろうか」という問題を考えてみる。もし，$p=1/2$ とすると，10回の試行(実験)で9回(またはそれ以上)も表が出る確率は，**二項分布**の公式を用いて，11/1024しかないことがわかる。ところが実際には，9回も表が出たわけであるが，それは $p=1/2$ であっても，ほんとうにたまたま，そのような事柄が実現したのかもしれない。しかしこれを，「この貨幣は表のほうが出やすい」，つまり $p>1/2$ であると考えたほうが自然であろう。仮説検定というのは，これを次のように形式化する。

「10回中，9回も表が出た。$p>1/2$ と考えてよいかどうかを検定せよ。」

このようなとき，$p=1/2$ という仮説をたてる。これを**帰無仮説**という。これに対して内心では，$p>1/2$ といったほうがよいのではないかと期待しているわけであって，$p>1/2$ のことを**対立仮説**という。これらを通常，次のように書く。

帰無仮説　H_0 ; $p=1/2$

対立仮説　H_1 ; $p>1/2$

上の例では，$p=1/2$ のとき，偶然9回以上も表の出る確率は11/1024でしかなかった。そこで，帰無仮説 $p=1/2$ を棄却して $p>1/2$ であると判断しても，誤りをおかす確率は5％以下である。このことを「**有意水準**(または**危険率**ともいう)5％で**帰無仮説を棄却する**」という。

第1種の過誤と第2種の過誤　前ページの投銭実験で，表が9回以上も出る確率は5％以下であるとして，「帰無仮説 $p=1/2$ を，有意水準5％で棄却する」ことを述べた。ここでは有意水準を5％としたが，有意水準として，5％か1％のどちらかをとるのがふつうである。有意水準5％で H_0 が棄却されたとき，**“有意である”** といい，1％で棄却されるならば，**“極度に有意である”** という言い方をしたりする。また，帰無仮説を棄却するということは，対立仮説を採択することでもある。前の例では，危険率5％で帰無仮説を棄却した。しかしもし，表が5回出たとした場合，p＝1/2であるとして，積極的に帰無仮説を採択するわけにはいかない。このようなときには，「$p=1/2$ でないとはいえない」という言い方をするのが正しい。帰無仮説については，表1に示すように，「棄却する」か，「棄却しない」かのいずれかの判断を下すべきであって，採択するとか採択しないという表現は適当でない。ところで，前ページでは表が9回も出たので，帰無仮説を棄却したわけであるが，実は仮説 H_0 が正しいにもかかわらず，誤ってこの仮説を棄却してしまったのかもしれない。このような誤りのことを，**第1種の過誤**という。これに対し，表が5回出たので，帰無仮説を棄却しないと判断したが，実は $p=0.8$ が正しい値であったとすると，ここでも判断の誤りをおかすことになる。このように，「帰無仮説が正しくないのにかかわらず，これを棄却しない誤り」のことを，**第2種の過誤**という。

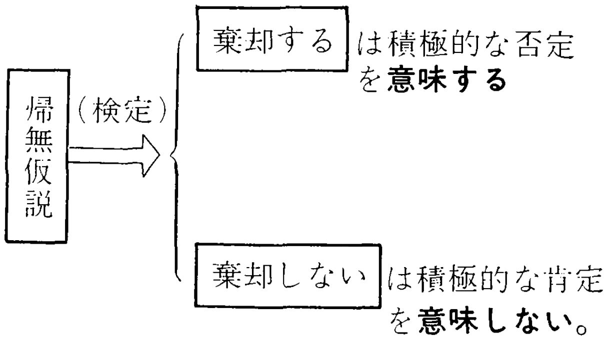


表1　仮説検定問題での判断

疑わしきは罰せず　裁判では被告に対して，ふつう，

帰無仮説 H_0；被告は無罪である

という仮説をたてる。これが明らかにくつがえされるに足る事実が判明しない限り，有罪とは認められない（ただし，公害裁判のような例外もある）。この場合「仮説（被告は罪を犯していない）が正しいのにかかわらず，これを棄却する誤り（有罪と判決する誤り）」が第1種の過誤であり，「仮説が正しくないのにかかわらず，それを棄却しない誤り」が第2種の過誤になる。裁判では，第1種の過誤を小さく抑えようとするので，第2種の過誤が大きくなることは避けられない。

生産者危険と消費者危険　抜取検査（→108ページ）での生産者危険は第1種の過誤に相当し，消費者危険は第2種の過誤に相当する。

■両側検定と片側検定　先の投銭実験の例では，

帰無仮説 H_0；$p=1/2$

対立仮説 H_1；$p>1/2$

として検定したが，これとは別に，

帰無仮説 H_0；$p=1/2$

に対し

対立仮説 H_1；$p\neq1/2$

を立てて検定することがある。前者のような検定の仕方を**片側検定**といい，後者を**両側検定**という。それでは，どういうときに片側検定を行い，どういうときに両側検定を用いたらよいかというと，大まかに次のような目安によればよい。すなわち，前の例のように，「$p>1/2$（表が出やすい）かどうかを検定せよ」という問題に対しては片側検定を行い，「$p=1/2$ かどうかを検定せよ」という問題に対しては両側検定を行う。有意水準 α を固定したとき，片側検定のほうが，帰無仮説を棄却しやすくなる。

■棄却域　投銭実験の例で，表が何回以上出たら帰無仮説を棄却するかという領域を定める。このような領域のことを**棄却領域**という。図1は，「平均値が未知である正規母集団から抽出した標本に基づき，母平均 μ が μ_0 であると考えてよいかどうか」を検定する問題を想定したときの棄却域を示している。横軸で太く塗った部分が棄却域である。

■検定力　第1種の過誤を小さくしようとすると第2種の過誤が大きくなり，逆に第2種の過誤を小さくしようとすると第1種の過誤が大きくなってしまう。そこで，これをどのように調和させるかが問題になる。これを一般には，次のような考えで対処する。すなわち，もし帰無仮説 H_0 が正しくないならば，実験結果がなるべく棄却域にはいりやすいようになっている検定法を選ぶべきである。ここで，H_0 が正しくない場合に，実験結果が棄却域にはいる確率のことを**検定力**とよんでいる。検定法は1つとは限らずいろいろ考えられるのであるが，それらの中で検定力の大きいものを選ぶのがよい。有意水準 α の検定法式の中で検定力を最大にするものを，有意水準 α の**最強力検定**という。

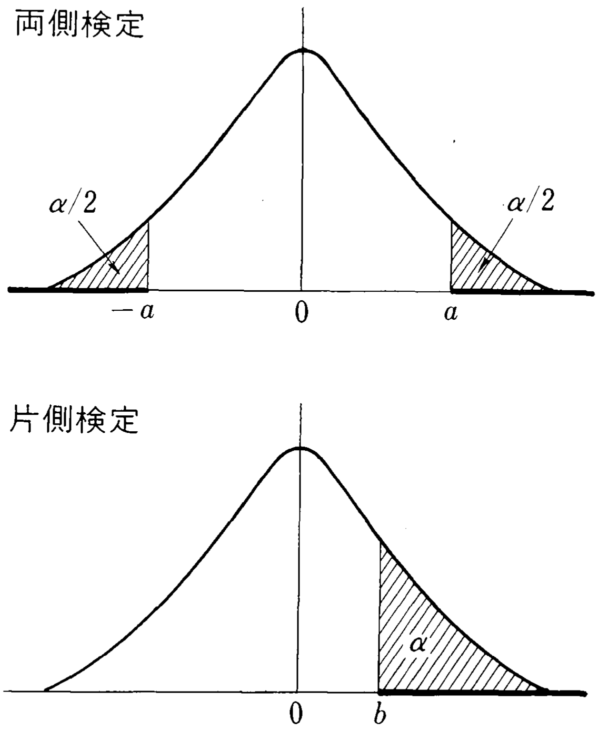


図1　有意水準 α での棄却域

表2は，ふつう行われている検定を

- 仮説検定
  - 母数の検定
    - 母平均・母比率の検定（正規分布表や t 分布表をつかう）
    - 2つの母集団の平均の差の検定（t 分布表をつかう）
    - 2つの母集団についての等分散の検定（F 分布表をつかう）
  - 分布の検定
    - 適合度検定（χ^2 分布表をつかう）
    - 分割表による検定（χ^2 分布表をつかう）

表 2　各種の検定

まとめて表示したものである。母数の検定に用いられる統計量は，推定(→260ページ)のときのそれと同じである。

■母平均の検定　分散 σ^2 が既知である正規母集団からとられた大きさ n の標本 $(X_1, X_2, \cdots, X_n)$ に基づいて，母平均 μ について，

帰無仮説 H_0 ; $\mu=\mu_0$

対立仮説 H_1 ; $\mu\neq\mu_0$

を検定するには，次の考えによる。まず，標本平均 $\overline{X}$ を算出すると，これは正規分布 $N(\mu, \sigma^2/n)$ にしたがう。よって帰無仮説のもとで，統計量

$$(\overline{X}-\mu_0)\Big/\left(\frac{\sigma}{\sqrt{n}}\right)$$

が標準正規分布 $N(0, 1)$ にしたがう。このことを用いて，上の仮説を検定すればよい。しかし，母分散が未知のときには，統計量

$$(\overline{X}-\mu_0)\Big/\left(\frac{S}{\sqrt{n-1}}\right)$$

が，自由度 $(n-1)$ の t 分布にしたがうことを用いて検定する。ここで，S^2 は標本分散である(→332ページ)。

■2つの母集団の平均値の差の検定

ふつう分散が大きく異ならない2つの正規母集団からとられた2組の標本に基づいて，母平均 μ_1, μ_2 の間に差があるかどうかを検定する。検定の結果帰無仮説 $\mu_1=\mu_2$ が棄却されるようであれば，2つの母平均の間に**有意差**(意味のある差)があるといい，このような検定を**有意差検定**という(→324ページ)。

■2つの母集団についての等分散の検定　上の，平均の差の検定を行うには，2つの正規母集団の分散が，あまり大きく異ならないことを確かめておく必要がある。このようなとき，等分散についての検定を行う(→325ページ)。

■適合度検定と分割表による検定　観測された資料を場合分けして，それがある特定の分布からとられたものであるかどうか，などを検定することを適合度検定という。また，資料をA，B 2つの属性によって分類して，AとBとが独立かどうかを検定したりする。これを分割表による検定という(→328ページ)。　(牧野)

ガンマ分布とアーラン分布

連続型の非負の確率変数の分布で，トランジスタなど電子部品の寿命分布，待ち行列のサービス時間の分布，所得分布などによくあてはまる。

★解説

ガンマ分布の**密度関数**や**特性値**は次のとおりである。

$$f(x)=\begin{cases}\dfrac{\beta^{\alpha}x^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta x} & (x\geqq 0 \text{ のとき}) \\ 0 & (x<0 \text{ のとき})\end{cases}$$

平均　$E(X)=\alpha/\beta$　　分散　$\mathrm{Var}(X)=\alpha/\beta^2$

ここで，$\Gamma(\alpha)$ は**ガンマ関数**といい，

$$\Gamma(\alpha)=\int_0^{\infty}y^{\alpha-1}e^{-y}dy$$

で与えられ，$\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$ という性質をもつ。とくに，$\alpha=k$ (整数)のとき，$\Gamma(k)=(k-1)!$ である。また $\Gamma(1/2)=\sqrt{\pi}$ である。

図 1 に見るように，α によって密度関数の形が異なる。α を**形パラメータ**という。β は**尺度パラメータ**である。**故障率**は図 2 に見るとおり，単調に減少する($\alpha<1$ のとき)場合から，単調に増加する($\alpha>1$ のとき)場合も含む。**ワイブル分布**とあわせていろいろな寿命分布によくあてはめられる。ガンマ分布は正規分布やワイブル分布，指数分布などと違ってどのような場合に生成されるかという物理的メカニズムがはっきりしていない。

次の分布はガンマ分布の特別な場合にあたる。

①**指数分布**；$f(x)=\beta e^{-\beta x}$　　($\alpha=1$ のとき)

② **k 次アーラン分布**；$f(x)=\dfrac{\beta^k x^{k-1}}{(k-1)!}e^{-\beta x}$ ($\alpha=k$ のとき)

③**自由度 k の χ^2 分布**；$f(x)=\dfrac{x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}e^{-\frac{1}{2}x}$

$\left(\begin{matrix}\alpha=k/2\\ \beta=1/2\end{matrix}\text{ のとき}\right)$

関連ページ
密度関数　56
特性値　272
故障率　40
ワイブル分布　418
正規分布　200
指数分布　156
χ^2 分布　314

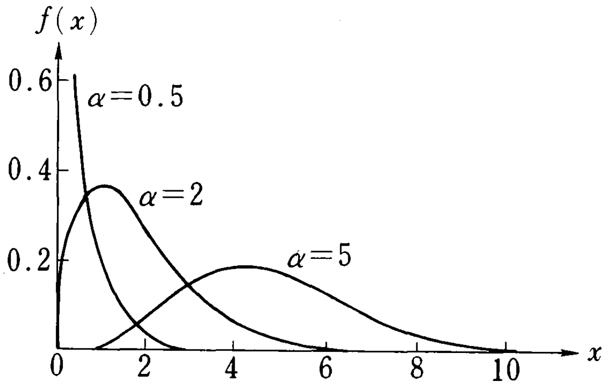
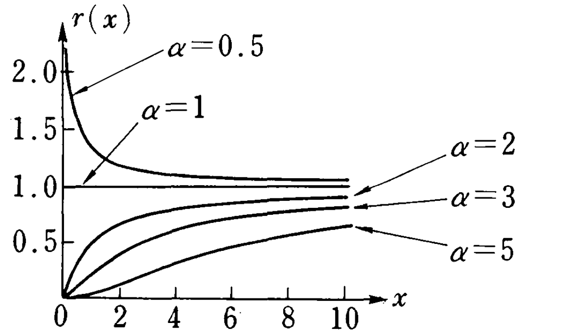


図1 ガンマ分布の密度関数($\beta=1$の場合)



(注) 一般にガンマ分布の故障率はxを大きくするとβに近づく。

図2 ガンマ分布の故障率 ($\beta=1$の場合)

■**所得の分布** 図3は1969年度の米国の1家庭当たりの所得の分布に，ガンマ分布と**対数正規分布**(→410ページ)をあてはめたものである(調査数40,746家庭)。所得分布は所得を10階級に分類して調査したため，ヒストグラムの幅が異なったところもある。適合度検定(→248ページ)を行ったところ，対数正規分布よりガンマ分布のほうがχ^2値がかなり小さく，あてはまりがよい。

■**アーラン分布** 1920年ごろデンマークの電話会社のアーラン(A. K. Erlang)が考えついた。電話の通話時間の分布は指数分布があてはまる場合も多いが，図4のように平均より小さいところに山を持つ場合も多い。図

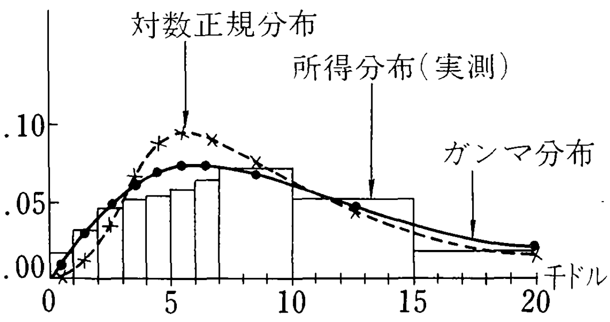


図3 所得分布

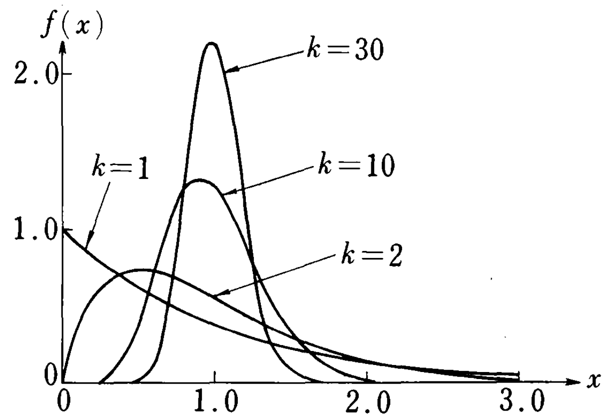


図4 k次アーラン分布(平均値=1の場合

4は平均値 (k/β) が次数kによらず，一定＝1 となるように描いてある。

アーランはこのようなサービス時間の分布を,次のようなメカニズムで説明した。1つのサービスは仮想的なk個の相(フェーズ)から成っており，k個の相を順に通過して初めて1つのサービスが完了する。各相の通過時間は互いに独立で，いずれも平均 $1/\beta$ の指数分布に従うとき，その総和となるサービス時間の分布は平均 k/β のk次アーラン分布となる。各相の通過時間が指数分布という性質が待ち行列モデルの解析にものをいう。$k=1$のときは指数分布,$k\to\infty$ のときは平均値に集中する単位分布となるなど,サービス分布の諸相が表せる。 (森)

管理図

管理図とは，管理限界線を有するグラフに工程より抽出された時系列データを記入し，もし，これが管理限界線の外に出れば，これを管理はずれの点とし，この点のはずれた工程上の不具合原因を究明して対策を講ずることによって工程を安定な状態に維持することを目的に作られたグラフをいう。

★解説

管理図はJISでは次のように定めている。

工程が安定な状態にあるかどうかを調べるため，または，工程を安定な状態に保持するために用いる図。

管理図を用いるときに，次の2つの過誤がある。

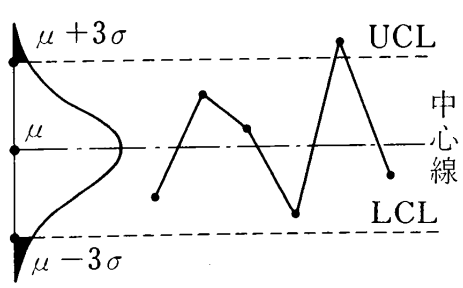
(1) **あわてものの誤り**　**第1種の過誤**に相当する。工程が安定しているのにかかわらず，データが管理限界線を外れて，工程が不安定であると誤認する誤りをいう。

(2) **ぼんやりものの誤り**　**第2種の過誤**に当たる。工程が安定していないのに，データが管理限界線の内側に入っているために，工程は安定しているとする誤りをいう。

管理図における管理限界線は，下図に示すようにデータの標準偏差を σ とするとき，データの平均値 μ の両側に 3σ の距離にとるので，$\mu+3\sigma$，$\mu-3\sigma$ のところにある。これらを順に**上方，下方管理限界線**(UCL，LCLと略記する)とよぶ。これは工程が安定してくるときに，これを不安定と誤るあわてものの誤りの確率を

$P(|x-\mu|\geqq 3\sigma)=0.0027$

となるようにしたもので，この考え方を 3σ 法という。



関連ページ

第1種の過誤　63
第2種の過誤　63

月日	4/1	4/2	4/3	4/5	4/7	4/8	4/9	4/10	計
伝票数	150	200	200	150	100	150	150	200	1300
誤記枚数	7	2	6	5	15	3	4	8	50

表 1

■管理図の種類 管理図には目的によって，

(1) 解析用

(2) 管理用

(3) 記録用

の3種があり，一般には，工程が十分に安定していないときには解析用，安定したならば管理用に用いる。

また，データの種類によって，$\bar{x}$-R 管理図，p 管理図(不良率管理図)および c 管理図(欠点数管理図)などがある。

■管理図の原理と作り方 管理図の中でも，簡単な p 管理図を用いて原理と作り方を説明する。

いま，某市において作成する伝票の中における誤記のあるものの枚数を毎日調査して，上のような表1を得たものとしよう。

表1より毎日の誤記率(不良率)を求めると表2のようになり，総不良率は

$$\bar{p}=\frac{50}{1300}=3.84\%$$

となる。

一般に，毎日の不良率は近似的に正

月日	4/1	4/2	4/3	4/5
誤記率	0.0467	0.01	0.03	0.033
月日	4/7	4/8	4/9	4/10
誤記率	0.15	0.02	0.0267	0.04

表 2

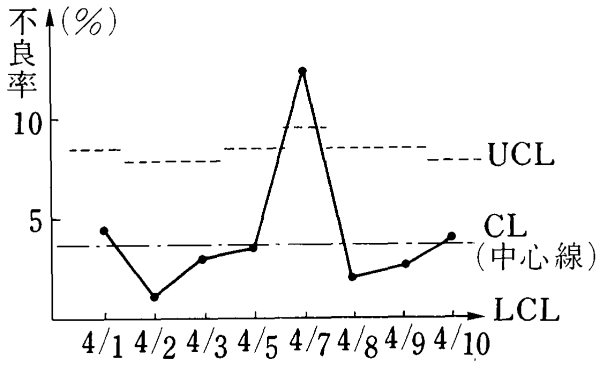


図1 p 管理図

規分布 $N\left(\bar{p}, \frac{\bar{p}(1-\bar{p})}{n}\right)$，ここで n は日々の伝票枚数にしたがうことが分かっているので，3σ 法の考えによって管理限界線は，

$$UCL=\bar{p}+3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$LCL=\bar{p}-3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

となる。この値は，たとえば，4月1日の場合は，

$$p_1=7/150=0.0467$$

となり，管理限界線は，

$$UCL=0.0384+3\sqrt{\frac{0.0384\times0.9616}{150}}$$

$$=0.0855$$

$$LCL=0 \text{（マイナスとなる）}$$

となる。これを p 管理図に示すと，図1のようになる。〔4〕〔5〕〔41〕〔48〕（真壁）

機会損失

ひとつの方策を採ると，他の方策の利益(得)を得る機会が失われる。この失う利益(得)のことを機会損失(opportunity loss)という。機会損失は機会費用(opportunity cost)ともいう。

★解説

ひとつの方策を採るということは，他の方策を捨てることを意味している。他の方策を採ることによってなんらかの利益が得られるのなら，その利益を得る機会を失うことを意味する。機会損失の概念は，このように比較の対象となる他の方策があってのものである。したがって，ひとつの方策の機会損失を明確につかむには，比較の対象となるもう1つの方策をはっきりさせておかねばならない。このことは**経済性の比較の原則**で述べられていることと同じである。それゆえ，方策の経済性を分析する際に，経済性の比較の原則を正しく適用すればおのずと比較の対象が明確になるので，機会損失という用語を知らなくても正しい計算をすることができる。

機会損失には，品切れ，材料不足，人と設備の能力不足などにより得られるはずの収益や粗利益を逃がしてしまうものや，安い材料や設備の入手し損ないなどにより費用の削減を逸してしまうものがある。また，収入の時期が遅れたり，支出の時期が早まると，金利の機会損失が発生する。

機会損失の概念の生きた使い方についてひと言ふれておく。検討の対象になっている方策や出来事(たとえば不良の発生)の機会損失は何か，と問いかけてみるのは価値がある。まだ気づいていないもっとよい機会(方策)はないだろうか。この出来事がもし起きなかったら，どんなよい機会(結果)が得られるだろうか。このような問いかけによって，新しい道がひらかれることになる。

関連ページ

経済性の比較の原則　98

■不良損のとらえ方　そば屋があり，人も設備も遊んでいるという時間がかなりあって，お客が増えればそれだけ売り上げも増すという**手余り状態**になっている。もりそば1個の費用は次のように計算されている。

売　価	500円
材料変動費	140円
人件費	200円
経　費	70円
利　益	90円

ある日のこと，店員さんがもりそばを1個客に渡すとき落としてしまった。そこですぐ1個余分につくり，客に出した。「1個落としたための損」はいくらになるだろうか。

落とさない場合と比べて何が変わるだろうか。収益は，客が帰らないので変わりない。人件費，経費も変わらない。変わるのは材料変動費の140円相当である。したがって，不良損は140円になる。

つぎに，このそば屋さんのもりそばは大へんおいしくて評判がよい。材料に限度があり，食べられない客がでる状態つまり**手不足状態**だとしよう。このときの不良損はいくらになるだろうか。

落とさない場合と比べると，変わるのは1個500円の収益が減ることである。その他の費用は変わらない。したがって，500円の損になる。売れる機会を逃したために，500円の**機会損失**が発生したことになる。

■在庫損失のとらえ方　ある工場でつくっている製品の在庫が倉庫にたくさんある。在庫による損失を計算しようとして，その製品の1個当たりの売価と製造原価を調べたら次のようになった。

売　価	10000円
変動費	4000円
製造経費	3000円
一般管理費	1000円
利　益	2000円

費用のなかで，変動費以外は短期的にみて，生産量や販売量が変化しても総額としては変わらない費用である。

在庫損失を計算するうえで大切なことは，問題状況を明確にし，比較の対象を正しくつかむことである。

もともと売れない製品を余分に製造して在庫になり，その製品を製造しない場合と比べるときは変動費の4000円がまるまる損になる。さらに，はやく処分しないことによる**機会損失**(処分収入，在庫管理費など)も発生する。

売れる製品でも，はやく製造しすぎたことによる在庫損を問題にしなければならない場合は，期当たりの金利がiのとき，1期当たり$(4000\times i)$円の損失が生じる。一方，1期売る時期をおそくしたために生じた在庫とみるべきときは，売る機会を1期おくらせたことによる**機会損失**として$(10000\times i)$円 が生じることになる。　(中村)

企業経営統計

現代の経済社会は，資本を提供する出資者，労働力を提供する労働者，並びに経営管理を行う経営者の協働によって営まれている。企業経営統計は，この協働活動の結果を統計として取りまとめたもので，企業会計における貸借対照表や損益計算書，製造原価明細書などを統計の母体としている。統計は，産業及び資本金階級別に示されているので，その区分ごとの経済性や企業経営の健全性，生産要素の効率性(たとえば労働の生産性)発展性の視点から分析することができる。『主要企業経営分析』(日本銀行)，『法人企業統計』(大蔵省)，『わが国企業の経営分析』(通商産業省)などがある。

★解説

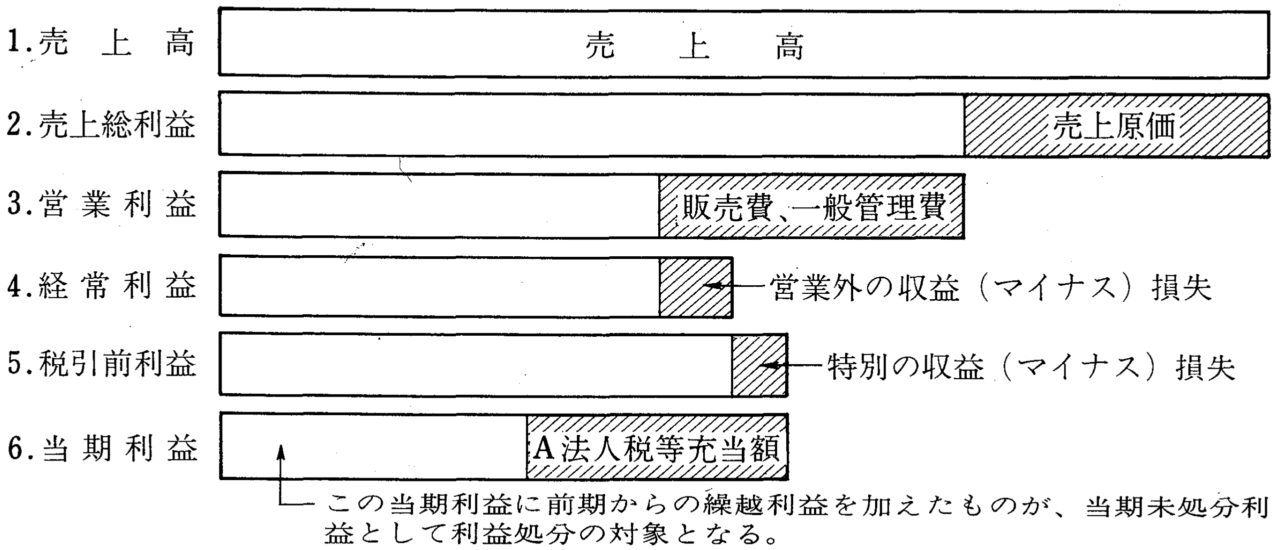
『主要企業経営分析』は，日本銀行が企業経営の実態を多面的に把握して金融政策に役立てるため，わが国の代表的な大企業約520社について調査している。原資料は各社の決算状況を示す有価証券報告書によっている。売上高，経常利益，付加価値率，付加価値構成，利益処分，資金需給，財務構成などのほか，利用者の理解を深めるため1人当たり売上高，売上高営業利益率など74種にのぼる経営指標をあわせて作成，提供している。

『法人企業統計』は，企業経営の分析，現状把握をねらいとして，大蔵省が昭和23年以来実施してきたもので，年1回調査(年報)と四半期調査(季報)とからなっている。季報では，資産・負債及び資本，損益，人件費，売上高などを調査している。

『わが国企業の経営分析』は，通商産業省所管業種を代表すると認められている資本金10億円以上の約380社について，収益動向，費用動向，設備投資と在庫投資，資産・資本の構成，財務の状況などについて調べている。

関連ページ

日本の統計　282



（注）売上原価………生産のための原材料費、工場労働者の賃金など
販売費、一般管理費………販売のための運送費、広告宣伝費、事務職員の賃金など
営業外の収益・損失………受取（支払）利息配当金など
特別の収益・損失…………土地や株式を売って一時的に得た利益や不良資産の処分によって発生した一時的損失
法人税等充当額……………国に納付する法人税など

表1　売上高と利益の関係

企業経営統計によって企業業績の動向をとらえるには，3つの視点が必要である。第一は，どれくらい売り上げが伸びたか(売上高向上)であり，これによって企業活動の成長の度合い，すなわち発展性を知ることができる。第二は，それによる収益の向上(収益拡大)であり，売り上げとの関係では企業の収益性，効率性をみることができる。第三は，これらの売り上げ，収益を生み出している企業体質の状況であり，自己資本と他人資本がどんなバランスになっているか，あるいは，総資本に対してどれほどの収益をあげたかなど企業の健全性である。

売上高と利益の関係は，表1のとおりである。

法人企業統計(年報)によると，昭和50年度以降，売上高の伸びは大幅に鈍化している。また，収益性の指標である売上高経常利益率は，40年代の3％前後から，50年代前半には2％以下に落ち込んだが，後半に入ると減量経営で企業業績は持ち直してきており，ほぼ石油ショック(48年秋)前の水準に持ち直してきた。総資本に占める自己資本の割合(自己資本比率)は，昭和35年の20.7％から年々低下し，昭和50年には13.9％となったが，その後はやや高まってきている。（市野）

記述統計

現象の認識の1つの方法は，統計資料に基づく記述的方法である。社会現象，経済現象など，さまざまな現象の実態を把握することは，たいへん重要であるが，困難なことでもある。一般には，実態を知る手がかりを得るために，データをとる。統計分析者は，調査対象(集団という)から得られたデータをもとに，集団の実態についての認識を得ようとする。これを，統計的記述といい，その方法論を記述統計という。

★解説

統計的記述を行うには，右ページの図1に示すように，さまざまな方法が用いられる。大別して，**度数表**や**グラフ**によるものと，**特性値**によるものとがある。度数表として，ふつうよく用いられるのは**度数分布表**である。これは，集団の各単位についての観察値を**階級**分けし，各階級区分に属する単位数を求めて，それを1つの表にまとたものである。このような度数分布表を図示するのに，**ヒストグラム**が用いられる。しかし，時には通常の度数分布表よりも，**累積度数表**を用いたほうが，都合がよい場合もあるし，**パレート図**，**確率紙**などによる表し方も有用である。

2つの変量の間の関係の強さなどを知るには，**相関表**や**散布図**が役立つ。**相関係数**は，これを数値的に表す尺度の1つである。

大量データを数値的に表すのに，集団の特性値である**代表値**，**散布度**などが用いられる。代表値としては，集団の中心的傾向を示す尺度である**平均値**，**メジアン**，**モード**などが用いられる。また，散布度はデータの散らばりの度合いを示す量であって，**標準偏差**，**分散**，**範囲**などがよく用いられる。さらにすすんだ分析をするために，**積率**，**歪度**，**尖度**などを算出することもある。

関連ページ
度数表　274
特性値　272
ヒストグラム　275
累積度数表　275
パレート図　300
確率紙　198
散布図　222
相関係数　223
代表値　272
平均値　272
メジアン　272
モード　272
標準偏差　272，332
分散　272，332
範囲　272
積率　210
歪度　211
尖度　211

	(度数表)	(グラフ)
図表による記述	度数分布表	ヒストグラム
	累積度数表	累積図表
	累積相対度数表	パレート図，確率紙
	相　関　表	散布図
特性値による記述	代　表　値	平均値，メジアン，モード，比率
	散　布　度	標準偏差，分散，四分偏差 範囲
	高次の積率	歪度，尖度

図 1　集団の統計的記述

■**記述統計学と推測統計学**　統計学は，個々の要素(単位)を対象とするものでなく，要素の集まりである集団を研究対象とする。統計学には，直接観察の対象となる集団の記述を研究目的とする**記述統計学**と，調査対象についてだけでなく，さらにすすんで一般的な場合への適応を考えようとする**推測統計学**とがある。そのために，推測統計学では，**母集団**と**標本**とを明確に竣別して用いる。記述統計学には，それがない。しかし，推測統計学においても，データ整理の段階，最終的に判断を下す段階で用いる道具は，記述統計学におけるそれと何ら異なることがない。

■**熟練した統計家**　自分に都合のよいデータをとり，都合のよい結論を引き出せるのが，熟練した統計家ではない。熟練した統計家とは，

「どのようにしたら適切なデータをとることができるか，またそのデータを使って，どこまで分析を進めることができるか，そして，どこでやめねばならぬかを知っている」

といわれている。統計では，通常多くのデータをとるが，せっかく入手し得たデータであるから，とことん分析しようとしがちである。しかし，データのもつ約束を無視してまで，分析をすすめてはいけない。　(牧野)

季節調整の方法

月次データや四半期別データの系列に，1カ年を周期とする季節変動が存在することがある。季節変動の除去の目的は，データから季節変動の影響を取り除いた季節変動調整済系列を求めることである。よく用いられる方法には，**CENSUS** 局法(米国商務省センサス局で開発されたもの)，**EPA** 法(経済企画庁で開発されたもの)，**MITI** 法(通産省で開発されたもの)などがある。ときには，季節変動の影響の大きさの度合いをみるために計算することもある。

★解説

季節変動の型が時の経過とともに少しずつ変化するとして，季節調整する方法の手順をのべる。原時系列が3つの変動，①長期にわたる規則的な変化を表す傾向変動，②1年を周期とした規則的な季節変動，③それ以外の不規則変動が重なりあったもの，で表されるということを根拠にしている。さらに数カ年を単位とする景気変動(循環変動)を加えることもある。ここでは時系列の構造として，**傾向変動**，季節変動，不規則変動のそれぞれの和で表された加法モデル，あるいは積で表された積のモデルに基づいて記す。

(1) **中心化12項移動平均**により，原系列から季節変動および不規則変動を取り出す。
(2) 季節変動と不規則変動を分離する。
(3) 各年の季節変動の合計が0になるように修正する。
(4) 原時系列から季節変動を引いて季節変動調整済系列を求める。

CENSUS 局法，EPA 法，MITI 法などは，この手順を基礎にしてより精密にしたものであり，この手順を何回か繰り返したり，異常値を取り去る方法を組み込んだりしている。

関連ページ

中心化12項移動平均　18

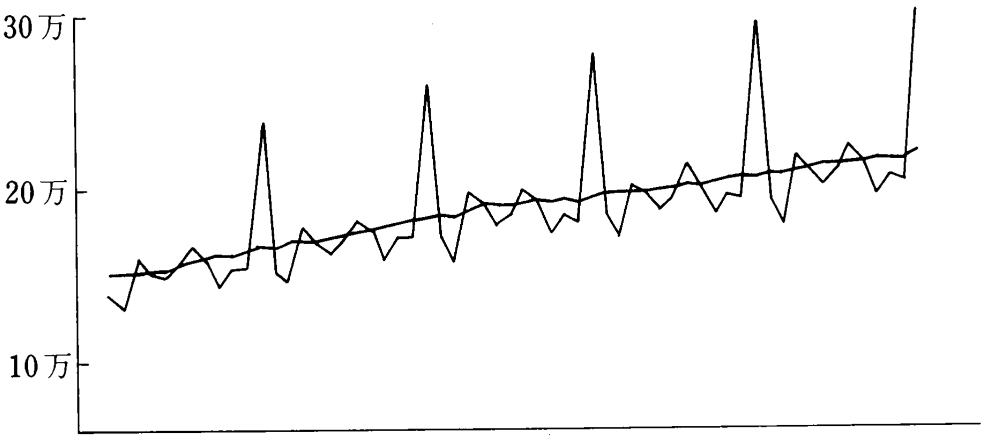


図1

全世帯消費支出の原系列 x_t が図1の実線である。いま，x_t は TC_t（長期傾向と数年を周期とする循環変動を区分しない傾向循環変動），S_t（季節変動），I_t（不規則変動）により，乗法モデル

$$X_t = TC_t \times S_t \times I_t$$

によって説明されるとする。EPA 法で求めた季節変動 S_t は次のようになる。

月	1975年	1976年	1977年	1978年	1979年
1	91.6	92.0	92.5	93.0	93.2
2	86.4	86.3	86.2	86.2	86.0
3	105.6	105.5	105.3	105.0	104.7
4	99.1	99.5	99.7	99.8	99.9
5	94.7	94.7	94.6	94.6	94.6
6	98.1	97.9	97.8	97.7	97.7
7	104.0	104.1	104.1	104.1	104.1
8	99.2	99.2	99.2	99.1	99.0
9	89.0	89.2	89.6	90.1	90.6
10	94.7	94.9	94.9	95.0	95.0
11	93.8	93.7	93.6	93.5	93.4
12	143.3	142.5	141.9	141.4	141.3

センサス局法や MITI 法も同じであるが，移動平均を適用することにより，季節変動の調整や不規則変動の除去を行い，季節変動調整済系列を求めている。季節調整済指数を求めるさいに用いる移動平均の式や手順が多少異なるが，それらの間に本質的な違いはない。当然，結果は微妙に異なる。EPA 法にも何種類かあり，ここでは EPA4 による季節変動調整済系列を記す（〔26〕参照）。

月	1975年	1976年	1977年	1978年	1979年
1	149357	164929	185162	197418	208071
2	150814	169427	183538	198471	208232
3	151989	169003	187608	197745	210161
4	151262	169684	191691	198330	211092
5	155952	171455	189934	198222	214024
6	159366	172676	189577	199329	215771
7	158698	174462	191248	203170	214883
8	159976	176741	193629	201771	216324
9	161287	178618	192310	204030	218699
10	161540	180706	193255	205937	218154
11	164607	182960	191609	207262	218363
12	166637	183097	194052	206744	220032

これを図示したのが図1の太線である。ここでの季節変動の調整は，季節による変動の型が年月を経るにしたがい，少しずつ変化すると想定して調整している。季節変動 S_t の値が，1975年12月の143.3と1976年12月の142.5では異なる数値であるのは，この考え方によって求めているからである。

ここでは，加法モデル $X_t = TC_t + S_t + I_t$ の季節変動 S_t の値を計算することもできる。経済学の分野では，和のモデルよりも積のモデルのほうがよく用いられている。（杉山）

協力 n 人ゲームの解

協力 n 人ゲームにおいて，妥当と思われる何らかの基準により選ばれた配分(→81ページ)，もしくは配分の集まりを解という。基準のとり方により，コア，仁，シャープレー値などさまざまな解がある。

★解説

■コア　配分 $(x_1, \cdots, x_n)$ が与えられたとき，たとえば提携 $\{1, 2\}$ は2人合わせて x_1+x_2 の利得を得ている。$\{1, 2\}$ は彼らだけで $v(\{1, 2\})$ を獲得できるから，この配分に対して $v(\{1, 2\})-(x_1+x_2)$ という不満の量をもつと考えられる。どのような提携に対しても，この不満の量が正にならない配分の集まりをコアという。

■仁　配分について，各提携がそれに対してもつ不満の量のうち最大のものをとる。各配分について最大不満の量を考え，それを最小にする配分をとる。このような配分が複数個存在する場合には，そのうち2番目に大きな不満の量を最小にする配分をとる。以下同様の手続をくり返し，最終的に得られる配分を仁という。仁は必ず唯一存在し，かつコアが存在する場合にはコアに含まれる。

■シャープレー値　プレーヤー i と i と含まない提携 S をとるとき，$v(S\cup\{i\})-v(S)$ は S に i が加わることによる獲得可能な値の増分，つまり i の S に対する貢献度と考えられる。この貢献度の平均値として与えられる配分をシャープレー値という。シャープレー値は，仁と同じく必ず唯一存在するが，必ずしもコアに含まれるとは限らない。

これらの解は，費用分担，便益分配などの計画問題に応用例が多い。右ページで，費用分担ゲーム(→81ページ)のコア，仁，シャープレー値を詳しく解説する。

なお，協力 n 人ゲームには，これ以外にも安定集合，交渉集合，カーネルなどの解がある。協力 n 人ゲームの解について，詳しくは参考文献〔30〕を参照されたい。

関連ページ
協力ゲーム　80

■**費用分担ゲーム**　このゲームの特性関数は，$v(\phi)=v(\{1\})=v(\{2\})=v(\{3\})=0$，$v(\{1, 2\})=v(\{2, 3\})=8$，$v(\{1, 3\})=1$，$v(\{1, 2, 3\})=20$ である(→81ページ)。

■**コア**　コアは，$x_1+x_2+x_3=20$，$x_1+x_2 \geqq 8$，$x_1+x_3 \geqq 1$，$x_2+x_3 \geqq 8$，$x_1, x_2, x_3 \geqq 0$ を満たす (x_1, x_2, x_3) の集まりである。したがって，線形計画問題

$$x_1+x_2+x_3 \rightarrow \text{最小化}$$

$$\text{制約条件}\quad \begin{aligned} &x_1+x_2 \geqq 8,\\ &x_1+x_3 \geqq 1,\\ &x_2+x_3 \geqq 8,\\ &x_1, x_2, x_3 \geqq 0 \end{aligned}$$

において，最小値が $v(\{1, 2, 3\})$ の値20以下であればコアが存在する。この問題の最適解は $(0.5, 7.5, 0.5)$ で，最小値は8.5，したがって，このゲームにおいてはコアは存在する。一般の n 人ゲームにおいても，同様にしてコアが存在するか否かが確かめられる。

■**仁**　まず，最大不満量の最小化に相当する線形計画問題

$$\varepsilon \rightarrow \text{最小化}$$

$$\text{制約条件}\quad \begin{aligned} &x_1+x_2+x_3=20,\\ &8-(x_1+x_2) \leqq \varepsilon,\\ &1-(x_1+x_3) \leqq \varepsilon,\\ &8-(x_2+x_3) \leqq \varepsilon,\\ &-x_1 \leqq \varepsilon,\ -x_2 \leqq \varepsilon,\ -x_3 \leqq \varepsilon \end{aligned}$$

を解くと，最小値は -6，最適解は唯一で $(6, 8, 6)$ である。したがって，このゲームの仁は $(6, 8, 6)$ であり，仁に基づく費用分担額は $(64, 47, 59)$ (万円)となる(→81ページ)。もし，この線形計画問題の最適解が複数個存在する場合には，これらについて，上の制約条件のうち狭義の不等号で成り立つものに関して，同様の最小化問題を考える。3人ゲームでは，この問題の最適解が仁である。一般の n 人ゲームにおいても，同様の最小化問題を繰り返し解いて仁が求まる。

■**シャープレー値**　プレーヤー1を考える。1を含まない提携は，ϕ，$\{2\}$，$\{3\}$，$\{2, 3\}$ の4個であり，それぞれに対する1の貢献度は，0，8，1，12 である。いま，プレーヤーが1人ずつ加わっていって3人提携がつくられる方法を考えると，全部で $3!=6$ 通りある。このうち，$\phi \leftarrow 1$ が起こるのは，$(\phi \leftarrow 1 \leftarrow 2 \leftarrow 3)$，$(\phi \leftarrow 1 \leftarrow 3 \leftarrow 2)$ の2通り，$\{2\} \leftarrow 1$ が起こるのは，$(\phi \leftarrow 2 \leftarrow 1 \leftarrow 3)$ の1通り，$\{3\} \leftarrow 1$ が起こるのは，$(\phi \leftarrow 3 \leftarrow 1 \leftarrow 2)$ の1通り，$\{2, 3\} \leftarrow 1$ が起こるのは，$(\phi \leftarrow 2 \leftarrow 3 \leftarrow 1)$，$(\phi \leftarrow 3 \leftarrow 2 \leftarrow 1)$ の2通りである。したがって，プレーヤー1の貢献度の平均値は，$(2/6) \times 0 + (1/6) \times 8 + (1/6) \times 1 + (2/6) \times 12 = 5.5$ であり，これがプレーヤー1のシャープレー値である。同様にして，プレーヤー2，3のシャープレー値は9，5.5と求まり，シャープレー値に基づく費用分担額は $(64.5, 46, 59.5)$ (万円)となる。　(武藤)

協力ゲーム

プレーヤー同士の話し合いが可能で，各プレーヤーの合意の上でそれぞれのとる戦略を決定するようなゲーム。非ゼロ和２人ゲーム(→ 310 ページ)，および３人以上のプレーヤーからなるゲームは，協力ゲームとなりうる。

★解説

関連ページ
非協力ゲーム　310

▩協力非ゼロ和２人ゲーム　非ゼロ和２人ゲームにおいて，もしプレーヤーが協力して互いのとる戦略を決定すれば，利得双行列のすべての利得の組の凸結合の全体が，実現可能な結果となる。逢い引きのジレンマ(→ 311 ページ)においては，実現可能な結果は図１の斜線部である。たとえば，(3/2，3/2)は，(手１，手１)，(手２，手２)をそれぞれ 1/2 ずつの確率でとることにより実現される。プレーヤー間の交渉により，実現可能な結果のうち，いずれが実現されるかを分析するのが，協力非ゼロ和２人ゲームの中心課題である。協力非ゼロ和２人ゲームは，２人交渉ゲームとよばれることが多い。

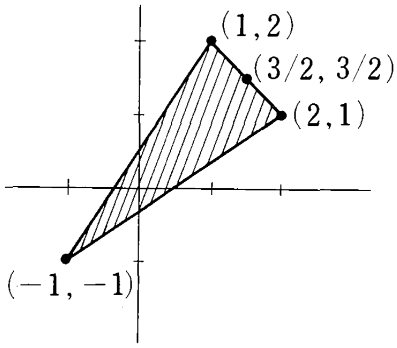


図１

▩協力 n 人ゲーム　３人以上のプレーヤーから成る協力ゲームにおいて，プレーヤーの全体を $N=\{1, \cdots, n\}$ とし，Nの部分集合Sを提携とよぶ。提携Sについて，Sに属するプレーヤーのみの協力で獲得可能な利得の値を $v(S)$ と表し，各提携Sに $v(S)$ を与える関数 v を特性関数という(空集合 ϕ に対しては，$v(\phi)=0$ とする)。(N, v) を協力 n 人ゲームの提携形ゲーム表現，または特性関数形ゲーム表現という。各プレーヤーがいかに協力しあい，また協力によって得られた利得をどのように分配しあうかを分析するのが，協力 n 人ゲームの中心課題である。

■ナッシュの交渉解　2人交渉ゲームにおいては，まず交渉を始めるにあたっての基準となる両者の利得 (c_1, c_2) を明確にしておかねばならない。これにはさまざまな方法があるが，一般には各プレーヤー自身の利得行列に関するマックスミニ値（→216ページ）をとることが多い。実現可能な結果の全体 R と基準点 (c_1, c_2) が与えられたとき，ナッシュは，交渉の妥結点 (u_1^*, u_2^*) の満たすべき性質としてパレート最適性など5つのものを考え，これらの性質を満たす (u_1^*, u_2^*) は唯一存在し，$(u_1-c_1)\times(u_2-c_2)$ を R 上で最大にするものであることを示した。この (u_1^*, u_2^*) をナッシュの交渉解という。逢い引きのジレンマでは，両者のマックスミニ値 $(1/5, 1/5)$ を基準点としてとれば，ナッシュの交渉解は，$(1/5, 1/5)$ から右上方に引いた45°線と，$(1, 2)$, $(2, 1)$ を結ぶ線分との交点 $(3/2, 3/2)$ で与えられる（図2）。

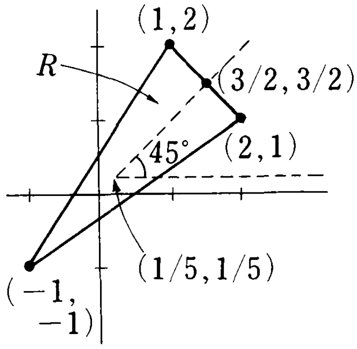


図 2

■配分　協力 n 人ゲーム (N, v) において，各プレーヤーが交渉の結果得る利得を，$x_1, \cdots, x_n$ と表す。このとき，① $x_1+\cdots+x_n=v(N)$, ②すべての $i=1, \cdots, n$ に対して，$x_i \geqq v(\{i\})$, の2つの条件を満たす $(x_1, \cdots, x_n)$ を配分といい，配分の全体を各プレーヤーの交渉の場と考える。①は，すべてのプレーヤーが協力したときに得られる値は余すところなく各プレーヤーに分配されることを表し，②は，各プレーヤーに対して，彼が単独で獲得できる以上の利得を与えることを表す。①を全体合理性，②を個人合理性とよぶ。

■費用分担ゲーム（例）　3人のプレーヤー1, 2, 3がある事業を行おうとしているとする。おのおのが独自に行えば，70万円，55万円，65万円の費用がかかるが，協力して行えば費用を軽減でき，1, 2が協力すれば117万円，1, 3が協力すれば134万円，2, 3が協力すれば112万円，全員が協力すれば170万円かかるとする。この170万円を各プレーヤーにいかに分担させればよいかを考える。この状況は，$N=\{1, 2, 3\}$, $v(\{1\})=v(\{2\})=v(\{3\})=0$, $v(\{1, 2\})=8$, $v(\{1, 3\})=1$, $v(\{2, 3\})=8$, $v(\{1, 2, 3\})=20$ という協力3人ゲームとして表される。ここで，$v(\{1, 2, 3\})=20$ となっているのは，1，2，3が協力すると，おのおのが独自に行った場合に比べ，$(70+55+65)-170=20$（万円）の費用を軽減できることによる。他も同様である。もしこのゲームの解（→78ページ）が (x_1^*, x_2^*, x_3^*) と与えられたとすれば，1，2，3の費用分担額は，$(70-x_1^*)$ 万円，$(55-x_2^*)$ 万円，$(65-x_3^*)$ 万円と与えられる。　（武藤）

区間推定

得られた標本に基づいて，母集団分布の母数の値を推定するのに，「いくらいくらの間にある」といったとき，的中率が何％である，という推し測り方をする。このような推定の仕方を区間推定という。

★解説

■母数　θ を推定するのに，標本の関数である a，b を上手に選んで，与えられた正数 α に対し，

$$P(a<\theta<b)=\alpha$$

となるようにする。こうしておけば，「母数 θ が a と b との間におさまっている」という表現をしたとき，的中率が 100α パーセントということになる。このような推定の仕方を区間推定といい，区間(a，b)のことを，信頼度(信頼係数ともいう) 100α パーセントの信頼区間という。

たとえば，**分散**が σ^2（既知)の**正規母集団**からとられた大きさ n の**任意標本**（X_1，X_2，…，X_n）に基づいて，**母平均** μ の 95％信頼区間をつくるには，次のようにする。**標本平均** $\overline{X}$，正規分布 $N(\mu,\ \sigma^2/n)$ にしたがうので，

$$P\{-1.96<(\overline{X}-\mu)/\left(\frac{\delta}{\sqrt{n}}\right)<1.96\}=0.95$$

上の 1.96 という値は，**正規分布表**から読みとったものであり，もし 99％信頼区間をつくりたければ，これを 2.58 にすればよい(注．α は 0.95 とか 0.99 にとるのがふつうであるが，ときには 0.90 にすることもある)。

$$P\left(\overline{X}-1.96\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+1.96\frac{\sigma}{\sqrt{n}}\right)=0.95$$

よって，$\overline{X}$ の実現値 $\overline{x}$ を用いて，母平均 μ の 95％信頼区間として $\left(\overline{x}-1.96\frac{n}{\sigma},\ \overline{x}+1.96\frac{\sigma}{\sqrt{n}}\right)$ が求まる。

関連ページ
母数　254
分散　332
正規母集団　314
任意標本　54
標本平均　261
正規分布表　429

信頼区間	母分散が既知のとき			母分散が未知のとき		
	正規母集団	標本数が大		正規母集団	標本数が大	
		無限母集団	有限母集団		無限母集団	有限母集団
95% 信頼区間	$\bar{x}\pm 1.96\dfrac{\sigma}{\sqrt{n}}$		$\bar{x}\pm$ $1.96\sqrt{\dfrac{N-n}{N-1}}\cdot\dfrac{\sigma}{\sqrt{n}}$	$\bar{x}\pm$ $t_{n-1}(0.05)\cdot\dfrac{s}{\sqrt{n-1}}$	$\bar{x}\pm$ $1.96\dfrac{s}{\sqrt{n}}$	$\bar{x}\pm$ $1.96\sqrt{\dfrac{1}{n}-\dfrac{1}{N}}s$
99% 信頼区間	$\bar{x}\pm 2.58\dfrac{\sigma}{\sqrt{n}}$		$\bar{x}\pm$ $2.58\sqrt{\dfrac{N-n}{N-1}}\cdot\dfrac{\sigma}{\sqrt{n}}$	$\bar{x}\pm$ $t_{n-1}(0.01)\cdot\dfrac{s}{\sqrt{n-1}}$	$\bar{x}\pm$ $2.58\dfrac{s}{\sqrt{n}}$	$\bar{x}\pm$ $2.58\sqrt{\dfrac{1}{n}-\dfrac{1}{N}}s$

ただし，σ^2＝母分散，n＝標本の大きさ，$\bar{x}$＝標本平均，s^2＝標本分散，N＝母集団の大きさ

表1　母平均の信頼区間

表1は，母平均μの95％信頼区間，99％信頼区間を定めるための公式を列挙したものである。これの使い方について説明しておく。

■母平均の区間推定

(1)母分散が既知のとき　母集団分布が正規分布であると考えられるときには，前ページに述べたように，母平均μの95％信頼区間は，

$$\left(\bar{x}-1.96\frac{\sigma}{\sqrt{n}},\ \bar{x}+1.96\frac{\sigma}{\sqrt{n}}\right),$$

99％信頼区間は，

$$\left(\bar{x}-2.58\frac{\sigma}{\sqrt{n}},\ \bar{x}+2.58\frac{\sigma}{\sqrt{n}}\right)$$

と定めればよい。

母集団における抽出単位数のことを母集団の大きさというが，それが無限個のとき無限母集団，有限個のときには有限母集団という。

無限母集団から，相当大きな数の標本をとり出すのであれば，母平均の信頼区間をつくるのに，正規母集団からの標本による場合と同じように，信頼区間を定めてよい。しかし，有限母集団の場合には，$\sigma/\sqrt{n}$ のところを

$$\frac{N-\sqrt{n}}{N-1}\cdot\frac{\sigma}{\sqrt{n}}$$

のように修正しなくてはいけない。

(2)母分散が未知のとき　たとえば，正規母集団の場合を考えるにしても，母分散σ^2が未知であるために，母平均μの95％信頼区間を

$$\left(\bar{x}-1.96\frac{\sigma}{\sqrt{n}},\ \bar{x}+1.96\frac{\sigma}{\sqrt{n}}\right)$$

と定めるわけにはいかない。このようなときには，統計量

$$(\overline{X}-\mu)/(S/\sqrt{n-1})$$

が，自由度$(n-1)$のt分布(→314，431ページ)に従うことを利用して，表1に示したように，信頼度95％の信頼区間を

$$\left(\bar{x}-t_{n-1}(0.05)\cdot\frac{S}{\sqrt{n-1}},\ \bar{x}+t_{n-1}(0.05)\cdot\frac{S}{\sqrt{n-1}}\right)$$

とすればよい。ここで，$t_{n-1}(0.05)$は，自由度$(n-1)$のt分布での両側5％の点で，nがきわめて大きいときには，正規分布の両側5％の点である1.96と一致する。同様に，$t_{n-1}(0.01)$は両側1％の点で，正規分布のときの2.58に対応している。

信頼区間	標本数が大	
	無限母集団	有限母集団
95% 信頼区間	$\bar{p}\pm1.96\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$	$\bar{p}\pm1.96\sqrt{\left(\dfrac{1}{n}-\dfrac{1}{N}\right)\bar{p}(1-\bar{p})}$
99% 信頼区間	$\bar{p}\pm2.58\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$	$\bar{p}\pm2.58\sqrt{\left(\dfrac{1}{n}-\dfrac{1}{N}\right)\bar{p}(1-\bar{p})}$

ただし，$\bar{p}$＝標本比率，n＝標本の大きさ，N＝母集団の大きさ

表 2　母比率の信頼区間

■母比率の区間推定　たとえば，不良率pの製品の山から，n個の製品を抽出したところ，そのなかにR個の不良品がふくまれていたとする。もし抽出個数nが大きいならば，$p=R/n$ が，平均P，分散 $p(1-p)/n$ の正規分布で近似できる。したがって，確率変数

$$X=(\bar{p}-p)/\sqrt{p(1-p)/n}$$

の分布は標準正規分布 $N(0,\ 1)$ で近似されるので，母集団での比率(これを母比率という)pの 95 %信頼区間をつくるには，

$$P(-1.96<X<1.96)=0.95$$

つまり，

$$P(-1.96<\frac{p-p}{\sqrt{p(1-p)/n}}<1.96=0.95$$

に注目することにより，表 2 に示したように，

$$\left(\bar{p}-1.96\sqrt{\frac{\bar{p}(1-\bar{p})}{n}},\ \bar{p}+1.96\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}\right)$$

とすればよいことがわかる。

母比率pの 99 %信頼区間も同様に定めることができて，

信頼区間	母平均が既知のとき	母平均が未知のとき
95% 信頼区間	$\left(\dfrac{ns^2}{\chi^2_n(0.025)},\ \dfrac{ns^2}{\chi^2_n(0.975)}\right)$	$\left(\dfrac{ns^2}{\chi^2_{n-1}(0.025)},\ \dfrac{ns^2}{\chi^2_{n-1}(0.975)}\right)$
99% 信頼区間	$\left(\dfrac{ns^2}{\chi^2_n(0.005)},\ \dfrac{ns^2}{\chi^2_n(0.995)}\right)$	$\left(\dfrac{ns^2}{\chi^2_{n-1}(0.005)},\ \dfrac{ns^2}{\chi^2_{n-1}(0.995)}\right)$

ただし，n＝標本の大きさ，s^2＝標本分散，$\chi^2_\nu(\alpha)$＝自由点νのカイ 2 乗分布での片側100αパーセントの点

表 3　母分散の信頼区間(正規母集団)

$$\left(\bar{p}-2.58\sqrt{\frac{\bar{p}(1-\bar{p})}{n}},\ \bar{p}+2.58\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}\right)$$

になる。

■母分散の区間推定　正規母集団 $N(\mu,\ \sigma^2)$ からとられた大きさnの標本に基づいて，母分散 σ^2 の区間推定を行うには，次のようにする。表 3 に示したように，母平均μが既知の場合，σ^2 の 95 %信頼区間を

$$\left(\frac{ns^2}{\chi^2_n(0.025)},\ \frac{ns^2}{\chi^2_n(0.975)}\right)$$

と定める。$\chi^2_n(0.025)$および$\chi^2_n(0.975)$は，自由度nの χ^2 分布 (→ 432 ページ)での片側 2.5 %および 97.5 % の点である。たとえば，標本の大きさが 20のときには，χ^2 分布表により，

$$\chi^2_{20}(0.025)=34.17,\ \chi^2_{20}(0.975)=9.591$$

したがって，σ^2 の信頼区間の幅は，かなり大きいものになってしまうのがふつうである。もし，母平均が未知の場合には，

$$\chi^2_{19}(0.025)=32.85,$$
$$\chi^2_{19}(0.975)=8.907$$

となり，幅がいっそう広がってしまう。

■**標本数の決定**　標本を抽出して母数を推定するとき，標本の大きさ n をいくらにしたらよいかという問題が出てくる。n を大きくすれば，よい推定値が得られるが，時間や費用がかさんでくる。右に，ある企業で自社製品のよしあしについて，“よい”と答えてくれる顧客が，全体の何％くらいいるかを知りたいとして，標本の大きさを決める問題をあげてある。この場合，前ページ(母比率の区間推定)により，母比率 p の 95 ％信頼区間は(大標本を考えることにして)，

$$\left(\overline{p}-1.96\sqrt{\frac{\overline{p}(1-\overline{p})}{n}},\ \overline{p}+1.96\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\right)$$

になる。右例で**精度**といっているのは，$|p-\overline{p}|$ のことである。したがって，精度 10 ％以内ということは，

$$1.96\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}\leqq 0.10 \qquad (1)$$

となることを意味する。(1)を，n について解くと，

$$n\leqq\frac{1.96^2}{0.10^2}\cdot\overline{p}(1-\overline{p}) \qquad (2)$$

になるが，$\overline{p}$ がわからないので，n は一意に定まらない。このようなとき，

$$\overline{p}(1-\overline{p})$$

が，$\overline{p}=1/2$ のときに最大になることを用いて，

$$n\geqq\frac{1.96^2}{0.10^2}\times\frac{1}{2}\times\frac{1}{2}=96$$

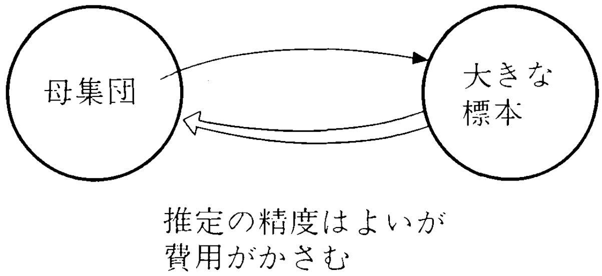


図 1　標本数 n を大きくするときの損失

> **(例)**　製品のよしあしについて，“よい”と答えてくれる客の比率 p の 95 ％，信頼区間；
> 精度 10 ％でよいならば，
> 標本数　$n=96$
> 精度 5 ％にしたいならば，↓4倍
> 標本数　$n=384$

としておけば，大きめに標本をとることになり，安全である。ただし，予備調査なり，過去の調査によって，$\overline{p}=0.70$ ぐらいだろうというような情報を入手している場合には，

$$n\geqq\frac{0.96^2}{0.10^2}\times 0.7\times 0.3=81$$

でよいことになる。

また，精度を 10 ％から 5 ％にしたいようなとき，ふつう「精度を 2 倍にする」といっているが，そのためには標本の大きさ n を 4 倍にしなければならないことが，式(1)からすぐわかる。

ここでは比率の区間推定について例示したが，母平均 μ についても同様で，精度 $|\mu-\overline{x}|$ をいくらにするかによって，標本数 n を決定することができる。　(牧野)

クラスター分析

個体の集まりである集団があり，よく似た個体を集めていくつかの群(クラスター)へ集団を分割する。似たものを集める過程で，個体と個，個体とクラスター，クラスターとクラスター間の類似度または距離を用いる。結果からクラスター内の個体の類似性とクラスター間の相違を把握して，対象集団の構造を分析する手法がクラスター分析法である。

★解説

クラスター分析は類似度または距離の入れ方によって結果が異なってくる。データの性格にあった適切な類似度あるいは距離を入れることが分析のポイントになる。

データ数をNとし，個体のそれぞれにp個の特性が測られており，

i 番目の個体の計測値を $(x_{1i},\ x_{2i},\ \cdots,\ x_{pi})$

j 番目の個体の計測値を $(x_{1j},\ x_{2j},\ \cdots,\ x_{pj})$

で表す。個体間の類似度としては，

$$ij\ 間相関係数=\frac{\sum_{k=1}^{p}(x_{ki}-\overline{x}_i)(x_{kj}-\overline{x}_j)}{\sqrt{\sum_{k=1}^{p}(x_{ki}-\overline{x}_i)^2}\sqrt{\sum_{k=1}^{p}(x_{kj}-\overline{x}_j)^2}}$$

個体間積和などがある。個体間の距離としては，

$$ij\ 間ユークリッド距離=\sum_{k=1}^{p}(x_{ki}-x_{kj})^2$$

や標準ユークリッド距離，マハラノビスの汎距離などがある。クラスター間の距離としては，平均距離，重心間距離などがある。

手法としては，樹形図などの階層的手法と非階層的手法とに分けられる。かなりの計算時間を要することもあり，広く利用されるようになったのは，コンピュータの発達してきた1960代以降である。

関連ページ

相関関係　222

マハラノビス距離　378

最近隣法では，クラスターhとクラスターkの距離を，

$$D_{hk}=\frac{1}{2}D_{hi}+\frac{1}{2}D_{hj}-\frac{1}{2}|D_{hi}-D_{hj}| \quad (1)$$

で定義する。ここでクラスターkは，n_i個の個体から成るグループiとn_j個の個体から成るグループjにより構成されている。D_{hi}は，クラスターhに一番近いグループiの個体までの距離であり，D_{hj}も同様である。式(1)は，

$$D_{hk}=\min(D_{hj},\ D_{hi})$$

と書きかえられる。

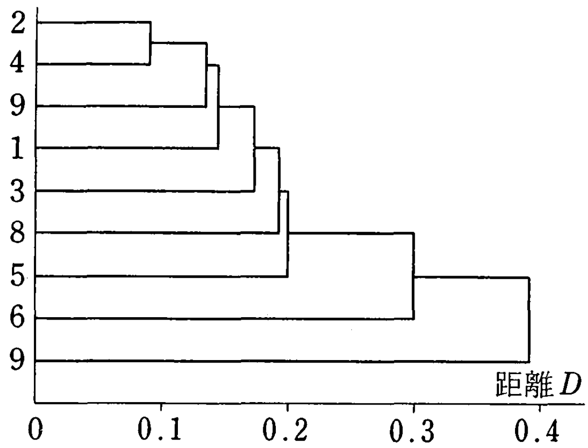
中学3年生137人の2学期の成績の相関行列は，335ページに示した。相関係数r_{ij}は，正で大きければ大きいほど科目x_iとx_jとの距離は近いことになる。そこで科目間の距離を，

$$1-r_{ij}$$

で与えてみよう。そのときの教科間の距離は，

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
国語 x_1									
社会 x_2	.16								
数学 x_3	.24	.22							
理科 x_4	.17	$\boxed{.09}$	.18						
音楽 x_5	.22	.22	.22	.20					
美術 x_6	.33	.30	.39	.34	.39				
体育 x_7	.43	.43	.45	.40	.45	.55			
技家 x_8	.19	.21	.29	.20	.21	.33	.39		
英語 x_9	.14	.16	.17	.13	.22	.38	.42	.42	

である。これに，最近隣法を適用して



クラスター分析を行う。

最小の距離はx_2とx_4の0.90である。そこで，x_2とx_4を合併して1つのクラスターと考える。$(x_2,\ x_4)$とx_1との距離は0.16と0.17であり，0.16をとる。$(x_2,\ x_4)$とx_3との距離は0.22と0.18であり，0.18をとる。x_5，x_6，…，x_9も同様にして，次の距離を得る。

	x_1	(x_2,x_4)	x_3	x_5	x_6	x_7	x_8
$(x_2,\ x_4)$	$\underline{.16}$						
x_3	.24	$\underline{.18}$					
x_5	.22	$\underline{.20}$	.22				
x_6	.33	$\underline{.30}$	.39	.39			
x_7	.43	$\underline{.40}$	.45	.45	.55		
x_8	.19	$\underline{.20}$	.29	.21	.33	.39	
x_9	.14	$\boxed{.13}$	.17	.22	.38	.42	.24

この距離表で最も小さいのは$(x_2,\ x_4)$とx_9の0.13である。最近隣法では，$(x_2,\ x_4,\ x_9)$を1つのクラスターとする。これを繰り返していくと，図1のような系統図ができる。　(杉山)

クリティカル・パス

PERTやCPMにおいて作った矢線図の上で，各仕事の所要時間を基に計算された最長路のことである。クリティカル・パス(critical path)の上にある仕事は，それが遅れたときはその遅れた時間だけ納期を遅らせてしまうことになる。したがって，それらの仕事の進行管理には特に注意を払わなければならない。この理由からクリティカル・パスを見つける作業がPERTの中心になる。

★解説

■パス　矢線図は1つのネットワークと見なされるから，ノードiからノードjに至る矢のつながりはパス(path)と称される。始点から終点に至るパスのうち，所要時間の最も長いパスが必ず1つは存在する。これがクリティカル・パスである。

■最早結合点時刻　各ノードで，そこから始まる仕事を開始できる最も早い時刻。すなわち，そこに矢じるしのはいってくるすべての仕事について，それらの仕事が全部終わる時刻である。

■最遅結合点時刻　各ノードで，(納期を守るためには)そこから始まる仕事を開始しなければならない最も遅い時刻。

■余裕時間　各ノードで，最遅結合点時刻から最早結合点時刻を引いた差。この時間がそのノードに与えられた余裕を表していることは，上の定義から容易に理解されよう。いいかえると，余裕時間内の遅れは，納期に影響を与えることがない。もし余裕時間が0であると，そのノードで終わる仕事も始まる仕事も遅れがないように気をつけなければならない。

クリティカル・パス上のノードは，すべて余裕時間が0となっている。したがって，このパス上の仕事の進行管理が大切になるのである。

関連ページ
PERT　292
CPM　142

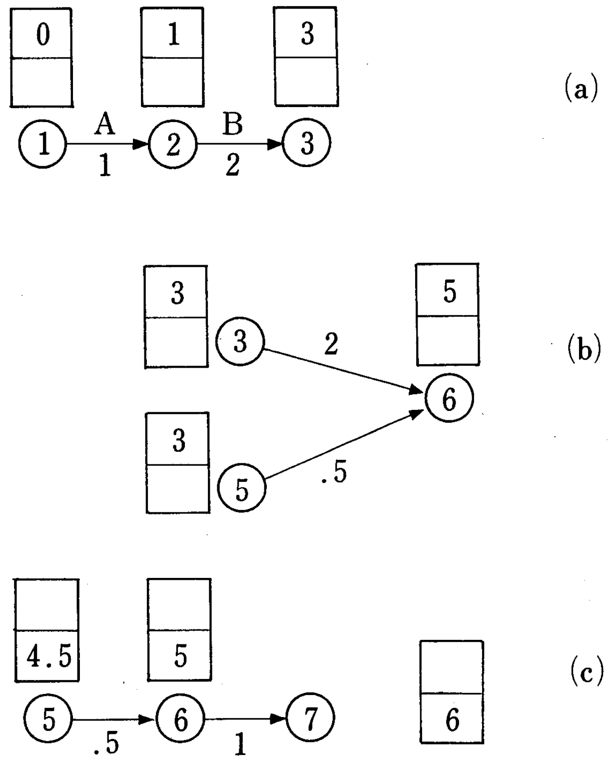


図1

293ページで扱った改修工事の例でクリティカル・パスを求めてみよう。仕事A，B，…，Gの所要時間は，それぞれ，1,2,1,1,0.5,2,1であるとする。単位は日でも時間でもよいのであるが，ここでは日としておく。

まず最早結合点時刻を計算する。そのために，図1に示すように，各ノードの傍らに四角の枠を2つずつ重ねて用意しておく。最早結合点時刻は，この枠の上段に記入する。

ノード1は改修工事のスタートを意味するから，ここには0と記入する。ノード2は仕事Aが終了すれば，次の仕事Bを始めることができる状態である。それ故，Aの所要時間1を0に加えた1を書き込む。同様に，この値に仕事Bの所要時間2を加えた

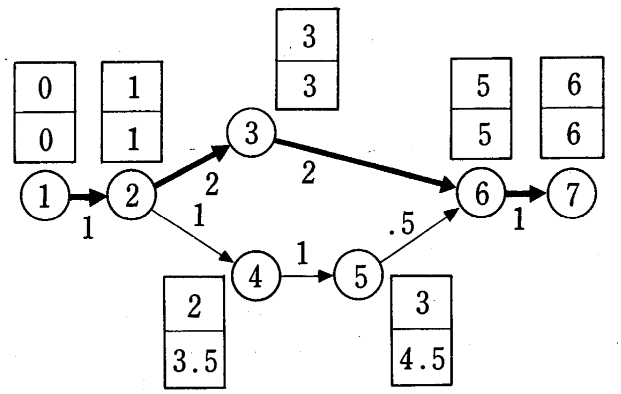


図2

値3が，ノード3の最早結合点時刻である(図1の(a))。ノード4，5の最早結合点時刻も同様に計算されるが，ノード6は少し違う。ここではE，Fの双方の仕事が終わらなければならない。それで3+2=5と3+.5=3.5の大きい方の値5がここでの最早結合点時刻となる。

このように，各ノードの最早結合点時刻は，そこにはいる矢じるしの前の各ノードの最早結合点時刻に各仕事の所要時間を加えて得た値のうち，最大のものとして定められる(図1の(b))。

このようにして最後のノード7の最早結合点時刻6が求まったら，それをそのまま下段に移して，ノード7の最遅結合点時刻とし，ここから前方のノードの方へ遡って最遅結合点時刻の計算を行う。図1の(c)に示すように，所要時間を引いて，前のノードの最遅結合点時刻が得られる。

図2は最終結果で，上・下の枠内の数の一致しているノードを結ぶと，太線のパスができる。これがクリティカル・パスである。（森村）

クロス集計(多重分割表)

n 個のアイテムのすべてのカテゴリーの組み合わせに対する頻度表を n 重クロスという。利用面としては，データ・チェック，再カテゴリー化の検討，アイテム(変数)間の独立性の検討，ベイズ診断に用いる経験確率の作成，判別成績の評価などに広く利用できる。

★解説

いま，n 個の観測対象がそれぞれ p 個のアイテムでカテゴリー値が与えられているものとする。

■1重(単純)クロス　アイテム X_i の各カテゴリーの度数分布を単純クロス表という。質的データの記述統計量的役割を果たす。解析に先だって，パンチミスや転記誤り等のミスデータのチェックに利用できる。次に度数分布を検討することによりカテゴリーの組み替え(再カテゴリー)に利用できる。たとえば，例数の少ないカテゴリーは，内容の似たカテゴリーと合併することが考えられる。

■2重クロス　アイテム X_i (n カテゴリー)と X_j (m カテゴリー)の $n \times m$ 個の組み合わせごとの度数分布を調べたものを2重クロス表，または $n \times m$ 分割表という。χ^2 検定により，X_i と X_j の独立性を検定できる。2×2 分割表，または $n \times m$ 分割表でカテゴリーに順序が考えられる場合には，属性相関係数を計算すればよい。

■3重クロス　アイテム X_i，X_j，X_k の各カテゴリーの組み合わせごとの度数分布を調べたものを，3重クロス表という。外的基準等の層別因子 X_i のカテゴリーごとに，X_j と X_k の2重クロス表を検討する場合などに用いる。

■n 重クロス　n が2以上のものを多重クロス表という。n が4以上の場合，分析が難しく注意を要する。

■その他の利用法　度数から経験確率を求めてベイズ診断を行ったり，尤度比方式による判別に利用できる。判別成績を 2×2 分割表にまとめたものを決定行列という。

関連ページ
質的データの解析　162
ベイズの定理を用いた判別　346
決定行列　104
2重クロス表の独立性と属性相関係数　278
尤度比方式による判別　390

項目	説　明	項目	説　明
x_3	年齢	$y_2{}^R$	耳鳴の有無
x_8	頭蓋骨骨折	$y_5{}^R$	耳鳴の強さ
x_{10}	意識障害(消失)	$y_9{}^R$	耳鳴の経過
x_{12}	鼻出血	$z_1{}^R$	平均聴力損失
$x_{13}{}^R$	右耳鼓膜異常	$z_4{}^R$	域値上検査
$y_1{}^R$	耳鳴発生時期	z_9	めまいの有無

表1　項目の一部(項目の右肩のRは右耳の意)

項目	カテゴリー	頻度	項目	カテゴリー	頻度
x_3	~29歳	24	$y_1{}^R$	~1日	15
	~39歳	32		~7日	18
	~49歳	32		~1ヵ月	38
	50歳~	16		1ヵ月~	33
x_{18}	無し	76	$y_2{}^R$	無し	44
	有り	28		有り	60
x_{10}	無し	36	$y_9{}^R$	消失・不変	52
	~1時間	40		軽快	36
	1時間~	28		増強	16
x_{12}	無し	83	$z_1{}^R$	~15dB	46
	有り	21		~30dB	42
$x_{13}{}^R$	無し	24		30dB~	16
	有り	80	$z_4{}^R$	(−)	43
$x_{13}{}^L$	無し	21		(±)	23
	有り	83		(+)~	38

表2　判別に用いる項目の度数分布表

■判別分析とクロス集計　頭部外傷による聴覚障害(難聴・耳鳴)の有無を外的基準とする判別式(数量化Ⅱ類)を作成する目的で，2024例を対象に，表1に示す29項目のデータを収集した。クロス集計を行い，異常データの修正と再カテゴリー化を行った後，すべての項目に欠測値のない104例(表2)を判別分析に用いることにした。

コード	内容	件数	コード	内容	件数
△	記入もれ	6	1	男	1557
0	ミスパンチ	1	2	女	60

表3　性別の単純クロス表

コード	内　容	件数	コード	内　容	件数
△	記入もれ	7	E	40 ~ 49	473
A	0 ~ 9	14	F	50 ~ 59	289
B	10 ~ 19	67	G	60 ~ 69	90
C	20 ~ 29	511	H	70 ~	6
D	30 ~ 39	566			

表4　年齢の分布(原票)

■異常データの発見と項目の削除　表3は性(X_4)の単純クロス表である。コード1は男を，2は女を示す。0は1のミスパンチであり，ブランク(△)は原票の記入もれであることが分かった。コード1と2の間のミスパンチは発見できないが，0や△のような明白な誤りは訂正できる。女性は全体の約3%しかないので，性別を説明変数に用いて判別分析を行うことは問題が多い。

■再カテゴリー化　表4は年齢の単純クロス表である。0歳から29歳までと，50歳以上を1つのカテゴリーとすることにより，カテゴリー数の減少とカテゴリーごとのデータ件数のバランスがとれる。このまま判別分析に用いた場合，結果に悪影響を及ぼす。このような順序尺度の再カテゴリー化は容易だが，普通は固有知識などの助けをかりて慎重に行わなければいけない。

(新村)

景気動向統計

個人や企業の経済活動の状況を“景気”といい，戦後の日本経済は好況と不況を繰り返しながら成長してきた。この好・不況の繰り返しのことを景気変動とか景気循環とよんでいるが，その状況をとらえる統計として企業家の景気判断を統計化した見通し調査と数多くの経済動向調査の結果をうまく総合した景気動向指数がある。景気動向指数は，景気に関する30の経済指標を総合したもので，28年以来，経済企画庁によって毎月作成されている。

★解説

(1)**「企業経営者見通し調査」**(経済企画庁)資本金1億円以上の企業約1,900社を対象に,向こう2四半期について前期に比べた見通しを,国内景気,国内·国外の需要動向,製品の価格,経常利益などについて聞いたもので,回答は「上昇」,「下降」,「不変」,あるいは,「強い」,「弱い」,「中間」もしくは,「増加」,「減少」,「保合い」というように分け，判断してもらったものをパーセントで表している。

(2)**「短期経済観測」**(日本銀行)　生産高,輸出額,設備投資額,資金ぐり，借入金利の水準，仕入価格，生産した商品の採算などについて四半期ごとの実績と当期の実績見込み及び来期の予測を調査したもので，結果が迅速にわかることから，政策判断の重要な指標になっている。調査は資本金10億円以上の530社を対象としている。

(3)**「労働経済動向調査」**(労働省)　生産動向とともに,残業,雇用，労働力の過剰と不足，新規採用の見通しなどの労働経済面について調査している。

(4)**「景気動向指数」**　種々の経済指数の中から，景気動向に先立って動く先行指標(12系列)，これと平行して動く一致指標(11系列)，およびやや遅れて動く遅行指標(7系列)に分けて集計しているもので，景気のすう勢と変換点を見ることができる。

関連ページ
企業経営統計　72

四半期	国内景気	業界景気	純益・経常利益
昭和			
58年1～3月	−16.0	−19.0	−15.0
4～6	−3.0	−9.0	−6.0
7～9	9.0	2.0	14.0
10～12	27.0	12.0	23.0
59年1～3月(見込み)	27.0	7.0	12.0
4～6 (見通し)	47.0	17.0	21.0
7～9 (見通し)	48.0	18.0	18.0

表1　企業経営者見通し(BSI)

月	先行指数	一致指数	遅行指数
58年2月	45.8	27.3	0.0
4	66.7	50.0	0.0
6	66.7	63.6	14.3
8	66.7	90.9	57.1
10	83.3	100.0	42.9
12	66.7	81.8	66.7

表2　景気動向指数(58年)

見通し調査は，企業経営者の経済の現況に対する考えや心理状態を示したものであるから，景気の先行きを知るうえで有益な判断資料を提供してくれる。

調査の結果得られる「上昇」，「下降」などの回答比率から次式によって算出されたBSI(ビジネス・サーベイ・インデックス)によって経済変動の先行きを判断することができる。

BSI＝「上昇」−「下降」

たとえば，59年7～9月期の国内景気の見通し(59年2月実施)によると，「上昇」とみたものが全体の48％，「あまり変らない」が51％，「下降」が0％であったから，BSIは，48−0＝48％となる。すべての経営者が「上昇」とみればBSIは100％，逆にすべて「下降」とみればマイナス100％，「あまり変わらない」とみれば0％となり，0を上回れば景気は上昇とみることができる。

景気動向指数はDI(ディフュージョン・インデックス)ともよばれており，それぞれの経済指標(ほとんどが季節変動を調整した数値)が，3カ月前に比べてプラスになったかマイナスになったかをチェックし，プラスの合計数を採用系列数で割ってパーセント表示したものをいう。0の場合は，0,5のウエートを掛けて総合する。

$$\mathrm{DI}=\frac{「+」の系列数\times 1+「0」の系列数\times 0.5}{全系列数}\times 100$$

である。

58年12月の景気動向指数(先行指数)によると，「＋」が8系列，「0」が0系列，「−」が4系列であったから，

$$\mathrm{DI}=(8\div 12)\times 100=66.7$$

という数値を得る。

したがって，DIが50％未満から50％を超えれば，景気は好況に転換したことを意味し，逆に50％以上から50％未満になれば，不況になったことになる。58年の月々のDI(表2)は，年間を通じて上昇しており，景気は上昇傾向にあると判断される。　(市野)

経済性工学

「経済的に有利な方策を探し，比較し，選択するための理論と技術の総合されたもの」である。別名として，「経済計算」，「採算計算」，「損得計算」などとよばれることもあるが，経済性工学という用語は，経済計算のための原理・原則，計算方法を指すにとどまらず，それらの活用技法をも含んだものとして通常は用いられる。

★解説

個人,企業を問わず,俗にいう「お金の計算」をしなければならない問題に直面する機会は数多く存在する。これらの問題は,その計算目的が基本的に異なる2つの問題に大別できる。その1つは,たとえば手作業と自動機のどちらが有利な製造法か,というような経済的に有利な方策を比較し判断する問題で,他の1つは,電気代を家主と下宿人でいくらずつ負担すべきかという経済的な成果や，発生したコストをいかに分配したり負担すべきかを決める問題である。

前者では「比較計算」が，後者では「割勘計算」が行われていることになるが，経済性工学は前者の比較計算を扱う。特に企業では，割勘計算が主役を演じる財務会計が利益を計算するために行われているが，比較計算で財務会計の計算法やデータを無批判に用いると誤りをおかすことになるので注意せねばならない。

経済性工学に関係する分野としては，管理会計，企業経済学，経済学での投資理論，エンジニアリング・エコノミー(EE)，インダストリアル・エンジニアリング(IE)などをあげることができる。これらの分野で得られた諸研究成果も経済性工学のなかに組み込まれている。

経済性工学の知識は，オペレーションズ・リサーチでは特にお金に関係しての，正しいデータのとり方，適切な目的関数や制約条件の設定の仕方，方策の比較法などの面で有効に活用されることになる。

関連ページ

経済性の比較の原則　98

インダストリアル・エンジニアリング　30

オペレーションズ・リサーチ　48

投資案の経済性指標　268

効率指標の使い方　116

	A	B
売り値	9万円/個	15万円/個
材料変動費	2万円/個	6万円/個
直接労務費	2万円/個	2万円/個
製造経費	4万円/個	8万円/個
利益	1万円/個	△1万円/個

表1　製品1個当たりの利益

	A 2個	B 2個
売り値	18万円	30万円
材料変動費	4万円	12万円
直接労務費	4万円	4万円
経費	12万円	12万円
利益	△ 2万円	2万円

表2　製品A, Bの比較

経済性工学を活用するには，計算のねらい・原則を正しく理解し，経済的な意志決定を助けるような分析をすることを心がけねばならない。正しいデータを使うこと，適切な判断指標で方策を比較すること，意思決定を助けるようにわかりやすく分析結果を表すことなどは特に大切である。これらについて，ごく簡単な例で説明する。

■正しいデータを使うこと　財務会計から得られるデータを無批判に使ってはいけない。つぎのような例を考えてみよう。ある工場で2種の製品AとBを製造している。両製品とも作業者の手間は製品1個当たり同じ工数をかけてつくっている。両製品の製造原価を計算し，1個当たりの利益を求めると表1のようになった。

ここで，直接労務費は工数に比例して配賦し，経費は材料変動費と直接労務費の合計からなる直接費に比例して配賦したものである。

表1の資料では，製品Aは黒字製品で製品Bは赤字製品となっている。そこで赤字製品の製品Bをやめて製品Aの製造販売に集中できるのならその方が経済的に有利になるだろうか。ちなみにBをやめAを2個つくった場合と，AをやめBを2個つくった場合を比べてみたのが表2である。

表2からわかるように，このケースでは赤字製品の方が有利な製品であったのである。配賦計算(割勘計算)に基づいた製品1個当たりの製品原価をむやみに判断資料に使ってはいけない。正しいデータは，**「経済性の比較の原則」**(→98ページ）にしたがって測定，整理することからはじめて得られることになる。

■判断指標の正しい使い方　2つ以上の方策の経済性を比較検討するのに，生産性，効率，利益率というような「率」の指標を使うことがよくある。率の指標も使い方の原則を理解して用いないと誤りをおかすことになる。

特定の製品の製造工場としてA，B，Cが候補にのぼっている。おのおのの工場では立地条件や資源の利用条件が異なるため，表3に示すように，設備投資からなる設備費と利益が異なった数字になった。

工場	設備費	設備費差引前利益	利益率
A	5,000万円	12,000万円	140%
B	10,000万円	20,000万円	100%
C	15,000万円	24,000万円	60%

表 3

工場Aを例にとると，利益は7000万円になる。利益率を計算すると，

$$(12000-5000)\div 5000\times 100=140\%$$

となる。

表3の結果では工場Aが最も利益率が高くなるが，利益率のいちばん低い工場Cでつくるのが最大の利益をもたらしてくれる。

この例のように，利益率で相互に排反的な案の優劣を判断してはいけないことがわかる。付加価値生産性，スペース生産性，売上高利益率などいろいろな率の指標はその正しい使い方を知って利用しなければならない。

■意思決定を助ける分析　経済性の分析は，意思決定をする人の判断を助けるようなものでなければならず，分析結果もわかりやすく表示する工夫が必要となる。

将来の見通しが不確実であると，意思決定は困難になる。特に，1回限りの出来事である投資案の分析では，確率のモデルが必ずしも頼りになるとはいえない。こんなとき，不確実な数値を動かしてみて案の有利さが変わる様子を調べる方法が有効になることがよくある。簡単な例を考えてみよう。

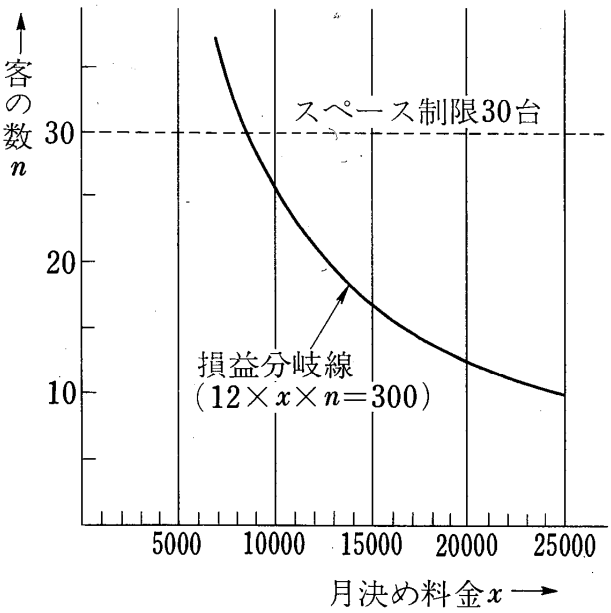


図１　損益分岐線

空き地の利用の仕方として2つある。車30台分のスペースをもつ月決め駐車場をつくると，必要な資金を借り，その返済と金利および維持費を加えて年間100万円になる。一方，この土地を他人に貸すと，年間200万円の収入になる。したがって，駐車場の収入が年間300万円以上あれば有利になることはわかる。

ところが，月決め料金と客の数を確実に予測できない。このようなとき，収入が300万円になるような月決め料金と客の数の組み合わせを，図1のような境界線(損益分岐線)で描いてみると判断の助けになる。料金が1万円なら24人以上でないとだめ，2万円なら…というようにして図を読み，必要なら月決め料金と客の数の関係のしかたをさらに調査をするなどし，八方にらんで決めることになる。

■**問題を適切にとらえること**　これまで述べたように，正しいデータや判断指標を使うこと，分かりやすく分析結果を表すことなどは意思決定を助ける分析のために必要なことであるが，経済性工学を活用するために，さらに重要な考慮すべき側面がある。

それは，解くべき**問題**を適切に受けとめて，その内容を明確にとらえることである。問題を明確にとらえるための手がかりは，経済性の分析対象となる「比較の対象（方策)」を明らかにすることである（→経済性の比較の原則)。比較の対象を明らかにする際に，それをとりまく周辺のいくつかの要因について適切な考慮を払うことにより，問題は明らかになってくる。それらの主な要因について，ここで簡単に説明しておこう。

(1)**計算の主体を明らかにすること**　分かりやすく言えば，誰の財布の中身を大きくすることをねらっているかということである。たとえば，AさんとBさんが帰宅する際，図2に示すようにタクシーで相乗りすると別々に乗るより安くつく状況を考えてみよう。A，Bさんは，ともにC社の社員だとする。このとき，相乗りするといくら得になるかをみると，計算主体がC社，すなわちタクシー代をC社が支払うのなら，明らかに1000円の得になる。ところが，計算主体がA，Bさんの各人の場合は割勘計算の問題も

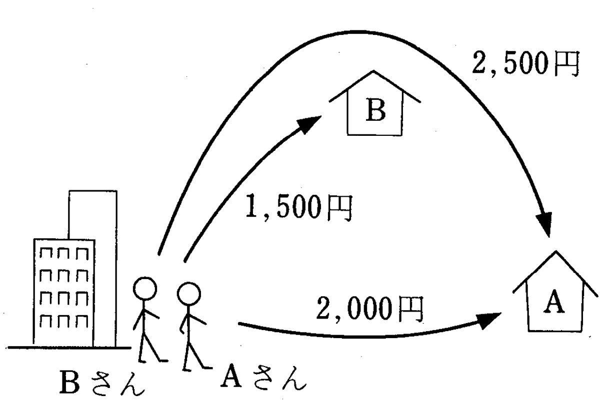


図2　タクシーの相乗り

処理しないと解答は得られなくなる。Aさんにとってはタクシー代の負担額が2000円以下，Bさんにとっては1500円以下ならば相乗りが有利ということしかわからないのである。

(2)**比較の対象（方策）の役割，目的を明らかにすること**　一般的にいえば方策は手段であり，その目的にかなった方策を比較しなければならない。たとえば，乗用車とトラックとどちらが有利か，と問題をとらえても意味がない。目的は何か。ドライブすることか引っ越しをすることか，で両者の有利さは異なってくるのである。

(3)**考察の時間的・空間的範囲を明らかにすること**　考察の時間的範囲については短期でみるか長期でみるかということであり，空間的範囲については一部の領域（たとえば製品の種類，工程，職場，組織などについて）に分析を限定するかという点に関することである。考察の範囲をせまくとりすぎたり，ひろくとりすぎると適切な分析ができなくなるので注意しなければならない。（中村）

経済性の比較の原則

方策の経済性を比較，検討するときに必ず従わねばならない基本的な原則で，次の2つがある。

第1原則：“比較の対象”を明確にする。

第2原則：比較の対象間の“違い”を，収入と支出の面から明らかにする。

★解説

▩第1原則：比較の対象の明確化　何と何を比較するのかという比較の対象を明確にすることである。たとえば，交通事故で骨折することを，えらい災難だと嘆く場合もあろうが，命が助かって幸運だと思う場合もあるだろう。交通事故に「遭った場合」と「遭わない場合」を比較するのか，事故に遭って「命を失った場合」と「骨折した場合」を比較するのかで，“交通事故で骨折すること”の評価は正反対なものになってしまう。

ひとつの案を検討することは，比較の第1原則を適用してみると，その案を「実施しない場合」と「実施する場合」の2つの分かれ道を比較することを意味しているのである。

▩第2原則：比較の対象間の違いを収入と支出の面からつかむ　比較の対象の間で収入，支出の違いを生みだす要素をつかみ，それによって収入，支出にどれだけの違いが生じるかを適切に推定することである。

たとえば，ひとつの注文の有利さを検討する際，注文を受けない場合と比べ，受けることにより，売上高および，材料，設備，工数…といった諸要素の投入，消費量に違いが生じるのか，生じるとしたら収入や支出でいくらの増加あるいは減少をもたらすのか，と問いかけて，その違いの内容を金額で明らかにすることである。

関連ページ

機会損失　70

方策のキャッシュ・フロー　358

収　　益	30万円
材　料　費	10万円
人　件　費	6万円
経　　費	4万円
利　　益	10万円

表1　注文Aの原価と利益

比較の原則のもつ意味，役割を簡単な例で説明してみよう。

■注文の有利さ　いま注文Aの引き合いがある。この注文を受けると，人の工数が40Hかかる。原価を試算すると，表1のように20万円になった。

この注文から得られる収益が30万円だとすると，利益は10万円ということになる。この利益額は注文の有利さを表しているといえるだろうか。

比較の原則を適用してみよう。まず第1原則についてみると，比較の対象は注文Aを「受けない場合」と「受ける場合」となる。つぎにこの注文を受けることにより，どの費用要素がどれだけ変化するかを問いかけてみるのが第2原則である。材料費は増加するか。答えは，増えて10万円の支出増になる。人件費は増加するか。答えは，もし給料でまかなっているのなら支出増にならない。経費は？　経費の中身は減価償却費が主なもので実際の支出が増加するのではない。そして収益をみると，収入が30万円増加することになる。以上まとめると，収入増が30万円，支出増が10万円となり，結局この注文の利益(の増加分)は20万円になる。

要　素	変化の仕方	変化量
収　益	増加	30万円
材料費	増加	10万円
人件費	変化せず	—
経　費	変化せず	—
利　益	増加	20万円

表2　注文Aの有利さ

比較の原則を適用した結果は，表2にまとめてある。ここで注意すべきことは，おのおのの要素についての変化の仕方は，種々の状況によって違ってくることである。たとえば，人件費についてみると，注文を受けることにより残業しなければならない場合は，残業賃金の支出増になる。また，もし材料が過剰在庫で処分値2万円で処分するはずのものを使うのなら，変化の仕方は処分収入の減少となり，その値は2万円になる。このように，おのおのの要素の変化の仕方を分析者が状況によって判断しないと，正しいデータに基づいた計算ができなくなるのである。

損得の計算では，機会費用，埋没費用といった用語がよく使われる。これらの用語の意味も，比較の原則から導かれることになる(→70, 364ページ)。経済性工学での損得の正しい意味，正しいデータのとり方，正しい費用計算の仕方を身につけ活用するには，ぜひとも経済性の比較の原則を理解しておかねばならない。　(中村)

ゲームの解法

ゲームを解くとは最適戦略とゲームの値を求めることである(→217ページ)。ゼロ和2人ゲームのような行列形ゲームにあって，利得行列に鞍点があればそれは最適純戦略を与えるからゲームは解けるが，そうでないときはめんどうになる。ゼロ和2人ゲームで，一方の純戦略が2つの場合は図を使って簡単に解けるが，一般には線形計画法に定式化しなおして解く。展開形ゲームでも，標準化という方法で行列形ゲームに直してから解けば，線形計画問題に帰着される。

★解説

▩図的解法　ゼロ和2人ゲームで，プレーヤー1が2つの純戦略を持ち，プレーヤー2はn個の純戦略を持っているものとする。具体的な例は右ページで示すが，そこの図1のように，直ちにn本の直線が引け，それらの最小値を結んで太線のような折れ線がかける。この図の横軸はプレーヤー1が戦略1をとる確率を表している。そして，この折れ線はプレーヤー1が最悪の場合に得る利得であるから，それを最大にする点Aの横座標が最適戦略を与え，縦座標がゲームの値を示している。

▩線形計画問題への帰着　図1の直線は，プレーヤー2がとる純戦略のおのおのに対応して，プレーヤー1の利得が，その混合戦略によってどのように変化するかを示している。プレーヤー1のとりうる純戦略がm個あり，彼の混合戦略が$(p_1,\ p_2,\ \cdots,\ p_m)$であるならば，図1の直線は$m$次元の1次式に拡張され，プレーヤー2が戦略$j$をとるとき，プレーヤー1の期待利得は，

$$a_{1j}p_1+a_{2j}p_2+\cdots+a_{mj}p_m \qquad (j=1,\ \cdots,\ n)$$

で表される。これらj個の直線がすべてvを越えているという制約の下で，そのvを最大化する線形計画問題は図1による解法と本質的に同じものである。

関連ページ
純戦略　216
ゲームの値　217
線形計画法　218
混合戦略　216

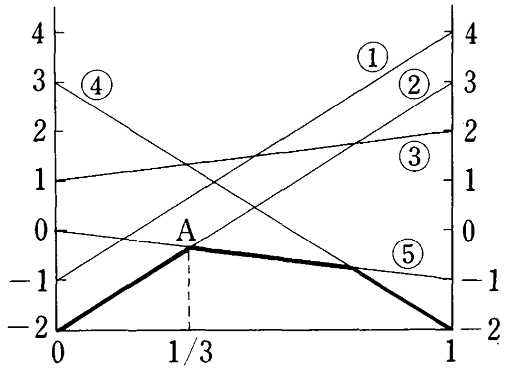


図 1

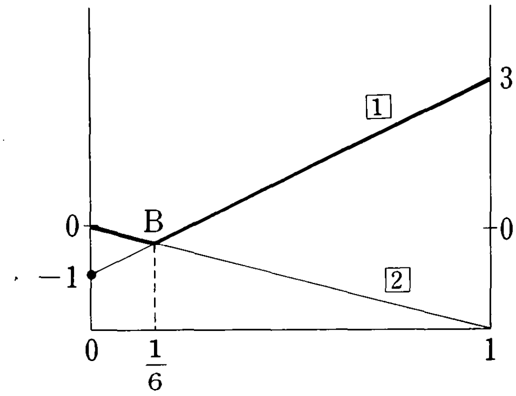


図 2

■例題(図的解法)　利得行列が次のような 2×5 行列で与えられるゲームを考える。

$$\begin{bmatrix} 4 & 3 & 2 & -2 & -1 \\ -1 & -2 & 1 & 3 & 0 \end{bmatrix}$$

プレーヤー 2 が第 1 の手をとったと仮定すると，プレーヤー 1 が第 1 の手を確率 p (したがって第 2 の手を確率 $1-p$)でとるときの期待利得は，

$$4p+(-1)(1-p)=5p-1$$

となり，図 1 の①という直線で表される。プレーヤー 2 のとる純戦略ごとにこの期待利得は別々の直線で与えられ，図 1 には 5 本の直線がかかれる。これらの直線をかくには，左側(横座標 0)の軸上に第 2 の手をとったときの利得(①の場合は −2)，右側(横座標 1)の軸上に第 1 の手をとったときの利得(①の場合は 4)に当たる点をとってその両者を結べばよい。

5 本の直線が引けたら，その中で一番低い値に当たるところを太線で示す。そして，太線の中で最も高い位置の点を求めればよい。図 1 では A がその点であり，A の座標は (1/3, −1/3) である。横座標はプレーヤー 1 が第 1 の手をとる確率であるから，彼の最適混合戦略は (1/3, 2/3) ある。また縦座標はそのときの期待利得であるから，ゲームの値もこれに一致し，−1/3 となる。

一方，点 A は②と⑤と 2 本の直線の交点であって，他の直線とは関係がない。このことは，プレーヤー 2 にとって，②と⑤以外の手には確率 0 が与えられることを意味する。プレーヤー 2 が②を確率 q でとるときの相手の期待利得を図 1 と同様に表すと，図 2 となる。図中の$\boxed{1}$，$\boxed{2}$はプレーヤー 1 のとる手に対応する。プレーヤー 2 にとっては，プレーヤー 1 の利得をできるだけ下げたいので，太線の最低点，つまり B 点が彼にとっての最適戦略を与える。図 2 の場合は，$q=1/6$ であるから，プレーヤー 2 の最適戦略は

$$(0,\ 1/6,\ 0,\ 0,\ 5/6)$$

となり，B 点の縦座標は −1/3 となっていて，ゲームの値が −1/3 であることがこれで確かめられた。

(森村)

ゲーム理論

ポーカーのようなゲームは，プレーヤーが定められたルールのもとで競い合う。社会でも人々は，あるルールに基づいて社会活動を行う。それで，社会行動の合理性，公平感などを定式化する上で，ゲームの定式化が役立つ。ゲーム理論は本来のゲームからだんだん離れ，人間の行動特に経済活動の数学理論として大きく発展している。

★解説

歴史的には数学者フォン・ノイマンと経済学者モルゲンシュテルンとの共著『ゲーム理論と経済行動』の発行された1944年が出発点である。戦後**OR**の発展の中で，意思決定の理論的基礎として研究が進められ，70年代にはいって応用面で著しい進歩をとげた。と同時に理論面でもさまざまな概念が導入され，内容も深められている。1972年には専門の国際的論文誌が創刊された。

ゲーム理論でいうゲームとは1組のルールのことであり，それを規定する要因として，プレーヤー数，戦略，継続時間，利得，協力の可否などがあげられる。

最も簡単なゲームは**ゼロ和2人ゲーム**である。プレーヤーは2人で，それぞれが複数の「手」の中から自分にとって最適と思われる手を選び，それでゲームは終わる。その結果，一方のプレーヤーは他方の払う利得を受け取る。つまり，一方が支払うものをマイナスの利得と見れば，両者の利得の和は常にゼロに等しいので**ゼロ和**という。このとき，2人が互いに協力することはありえないが，プレーヤーが3人以上となったり，ゼロ和でなかったりすると協力した方がお互いに得ということもある。それで**協力ゲーム**，**非協力ゲーム**の別が生じる。

ゲームの表現には，ゲームの構造を**木**を使って表現する方法や戦略を中心にする方法などがある。右ページにその簡単な例を示す。

関連ページ

オペレーションズリサーチ　48

ゼロ和2人ゲーム　216

協力ゲーム　80

非協力ゲーム　310

■ゼロ和2人ゲームの利得行列　ゼロ和2人ゲームでは一方が他方から支払いを得るのであるから，一方の利得のみを記せば，他方はそれに負号をつけたものになるので，改めて記す必要はない。プレーヤーを区別するため，それぞれをプレーヤー1，プレーヤー2と称すると，プレーヤー1の利得のみを記せばよい。

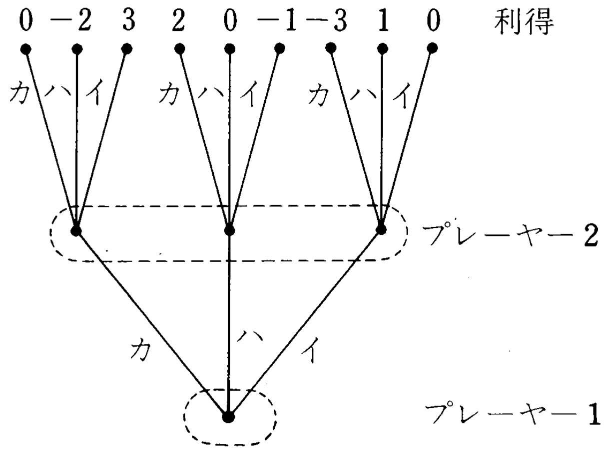
いま，2人のプレーヤーがジャンケンをする場合を考えよう。2人ともカミ，ハサミ，イシの3つの手のうちのどれかをとることができる。どの手で勝とうとも，勝ったとき相手から10円をもらうのであれば，勝ったときの利得は10，負けたときの利得は-10とおけばよい。このとき，プレーヤー1の利得は右の行列で表される。両者の手が同じであれば「あいこ」なので利得は0，プレーヤー1がカミのとき相手がハサミなら負けて利得は-10円となり，イシなら勝って10円の利得を得ることが1行目の数で表されている。他の行についても同様である。

	手＼手	プレーヤー2 カ	ハ	イ
プレーヤー1	カ	0	-10	10
	ハ	10	0	-10
	イ	-10	10	0

もし，カミで勝てば30円，ハサミなら20円，イシなら10円と定められているならば，その利得表は右のようになる。

	カ	ハ	イ
カ	0	-20	30
ハ	20	0	-10
イ	-30	10	0

■ゲームの木　プレーヤーのとる手に応じて，ゲームの推移する様子を，いわゆる樹形図の形に表すこともある。下の図は，ジャンケンの例を表したもので，「根」にあたるノードは，ゲームの開始前の状態を示している。そこから出る3本の枝のおのおのは，プレーヤー1がカミ，ハサミ，イシのいずれかの手をとることを示している。



プレーヤー2も同様に3つの中の1つの手をとる。もしプレーヤー1がカミという手をとっていたとき，プレーヤー2がハサミをとれば，ゲームは，最上段の左端から2番目のノードに至って終了する。そのとき，プレーヤー1は-20円の利得を得る(図では-2として記してある)ことを図は示している。この例では，2人のプレーヤーは同時に手の選択をしており，プレーヤー2はプレーヤー1が何を出したかを知った後に行動するわけではない。それで彼らは2段目のノードではどの点にいるのかを知らない。このことを点線の囲みで示している。〔30〕

（森村）

決定行列(デシジョン・マトリックス)

決定行列とは，2×2 分割表を用いて判別結果をまとめたものである。判別成績の評価は，決定行列からいわゆる誤分類確率を計算して用いることが多いが，ROC 曲線を用いた評価のほうが好ましい。

★解説

■決定行列　図1の上段は，判別軸上の2群の分布の模式図である。判別境界点 d の決定により，$x<d$ の領域は G_1 群に，$d<x$ の領域は G_2 群に判別される($x=d$ はいずれかに入れる)。これを 2×2 の分割表にまとめたものを決定行列という。

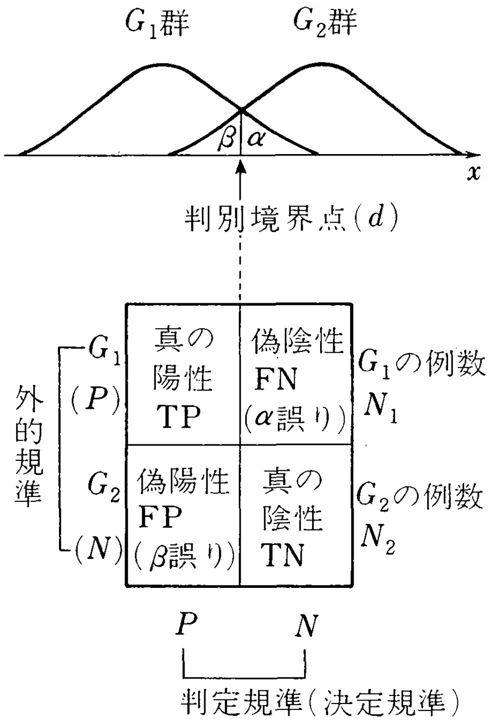


図1　決定行列と分布

G_1 群で正しく判別されたものを TP(True Positive)とよび，誤分類されたものを FN (False Negative) または α 誤り(第1種の過誤)とよぶ。同様にして G_2 群に対して，TN(True Negative)，FP(False Positive，β 誤り，第2種の過誤)が求まる。

TP と FN を G_1 の観測数 N_1 で除したものを，TPR (TP 比，True Positive Rate) と FNR という。同様に FPR と TNR が決められる。FNR と FPR は一種の誤分類確率だが，通常，式(1)を内部標本の誤分類確率という。

$$\mathrm{P}=\frac{\mathrm{FP}+\mathrm{FN}}{N_1+N_2} \tag{1}$$

関連ページ

クロス集計　90

ベイズの定理を用いた判別　346

FUNCAT による判別　320

数量化Ⅱ類　190

ROC 曲線　14

■**例題**　いま外科手術によって，がん細胞の有無が確認されている100例について，手術前のX線フィルムの読影による医師の判断(1：がん確診，2：がん疑，3：良性確認しえず，4：良性確診)を表1に示す。

G_1はがん，G_2は正常を表し，外科手術で判定されたものを外的基準とする。D_1はがんと判定された悪性(X線による判定が1と2)を，D_2は良性(同じく3と4)を表している。

	D_1	D_2	計
G_1	16	4	20
G_2	20	60	80

表1

この表を見ると，TPは16例で，TPRは80%，α誤りは4例で20%，β誤りは20例で25%となっていることが分かる。また，TNは60例で，TNRは75%である。

誤分類確率(式(1))は24%(=24/100)とαとβの間にくる。(1−P)で表される判別的中率76%を用いることもある。

医学では，検査の結果，疾病を表すと考えられる所見を＋，正常所見を−と表すことが多い。このため，疾病や興味の強い疾病の方を陽性群といい，他方を陰性群と考える。

■**感度と特異度**　医学分野では，伝統的に次式で表される感度(sensitivity)と特異度(specifity)を用いることが多い。

$$\begin{aligned}\text{感　度} &= TP/(TP+FN)\\ \text{特異度} &= TN/(TN+FP)\end{aligned} \tag{2}$$

ただし，

$$\alpha=1-\text{感度},\ \beta=1-\text{特異度}$$

表1から，感度は80%でこの検査ががん症例の診断にかなり良好であることが分かる。

■**決定行列の問題点**　図2は，68例のがん症例と33例の正常症例のX線写真の情報をもとに，医師診断とベイズ診断による判別成績を決定行列で表示した。

		医師診断 悪性	医師診断 良性	ベイズ診断 悪性	ベイズ診断 良性
疾病分類	がん	63	5	61	7
	正常	10	23	7	26

図2　2つの決定行列

上の2つの決定行列は，どちらの方が成績がよいといえるだろうか。医師診断の誤分類数が15，ベイズ診断が14例だから，ベイズ診断がよいといってよいだろうか。FNではベイズ診断が悪く，FPでは医師診断の方が悪い。もし，ベイズ診断の診断基準を変えてFNが5例になったとき，FPが仮に20例となったとすれば，この場合には医師診断のほうがよいということになる。

■**ROC曲線**　上で述べた決定行列の不明確さをさける目的で，判別境界点を動かして複数個の決定行列を作ればよい。これを1枚の図に表示したものがROC曲線であり，応用性に優れている。　(新村)

決定の木と PDPC

将来どのように展開するか不確実な過程において，時々の節目節目に何らかの決定を下す必要がある場合，決定の影響如何を視覚的に見やすくして決定を下す助けにする手法の代表的なもの。決定の木は，決定の回数が比較的少なく，かつその影響が定量的に測りやすいときに向いている。

PDPC (Process Decision Program Chart) は過程決定計画図ともいわれ，望ましい結果を得るために，どのような局面でどういう決定をしていくべきかを積極的に探し，現実をそれに近づけるべく努力をするのに用いられる。

★解説

■決定の木　分類のために自然に用いられている樹形図のノードを，決定選択点と確率選択点の2者に分けて記述し，決定選択点における決定をどのようにすれば結果から見て望ましいかを考察するために利用する。これをふつう決定の木(decision tree)という。決定選択点では，意志決定者の意思決定の如何により，その後の事象の展開が定まり，確率選択点では，彼の意思とはかかわりなく，何かほかの要因によって，そこで分岐する道のいずれかが実現するものと考えられている。考え方も図示方法も単純なので使いやすいが，ノードの数が増えるとわずらわしくなる。

■PDPC　交渉や競争の状況下では，不確実さの要因は主として他者の決定と環境変化である。このようなとき，確率を既知とした決定の木では対処しきれない。いわゆる堂々めぐりやデッドロックといった現象も起こりうるので，ネットワークも木にならない。PDPC は，そういった状況を現在わかっている限り記述して決定の助けとし，状況が変われば，その時点で再考するという動的な手法である。

関連ページ
発想的 OR　294
ネットワーク　284

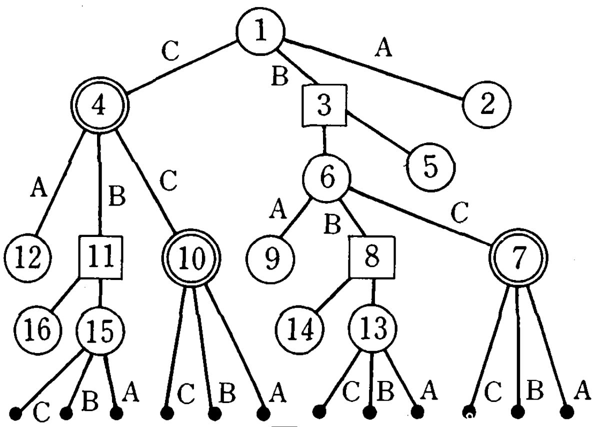


図1

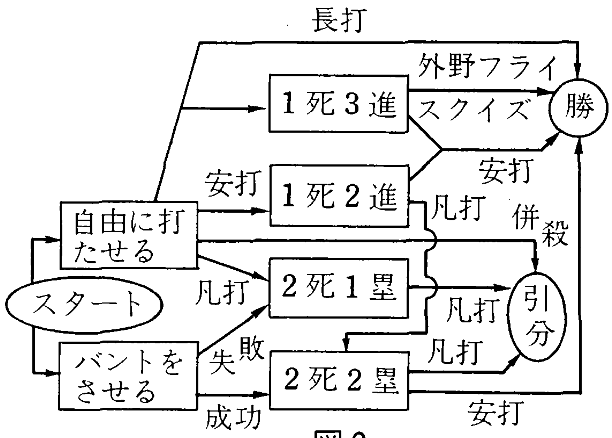


図2

■決定の木の例　いま，秘書選びの問題というニックネームのつけられた問題を考えよう。ある重役が彼の秘書を選ぶのに3回まで面接をすることとし，候補者を順次呼んでその都度採用か否かを決定する。従来の経験から，秘書候補者はA，B，Cの3段階に大体分けられ，2：5：3の割合で面接に訪れると仮定される。また，彼はそれぞれのグレードの秘書に3，2，1の点を与えている。できるだけ高い点の秘書を得るためにはどのような決定を下すべきか，これが問題である。

図1の○印のノードは確率選択点，□印のノードは決定選択点を示している。①は始めの状態で，第1の候補者がA，B，Cのいずれかによって分岐する。Aが来れば採用して以後の面接は行わないから，ノード②から先へは進まない。しかしノード③では，更に面接を続けるか，やめてBの人の採用に踏み切るかの選択がある。面接を続けることにすればノード⑥に移って，第2の候補者がA，B，Cのいずれかによって，⑨，⑧，⑦のいずれかにいく。以下同様に進む。ノード⑦においては，必ず次の候補者と面接するので，決定選択点は省略したが，そのことを明示するため，図中では2重丸にしてある。仮定した比率の下では，1度の面接の得点の期待値は，

$$0.2\times3+0.5\times2+0.3\times1=1.9$$

であるから，ノード⑦や⑬での期待値は1.9で，ノード⑭の2より小さい。それゆえ，ノード⑧の選択は⑭を選ぶべきである。つまり，2人目がBのときはその人を採用するのがよい。ノード⑥の期待値は上式の1を1.9に代えて，2.17という結果を得るから，③では⑥を選択し，第2の候補者と面接すべきであろう。この解法は動的計画法(→270ページ)だが，決定の木はそれを視覚化している。

■PDPCの例　ごく簡単な図にするため，野球で同点を迎えた9回裏1死後四球で出塁したところからの動きを図2に例示する。ここに書かなかった事態が生じたときは，その状況を踏まえて，また書き直す。参考文献〔22〕

（森村）

検査と抜取検査

検査とは，JIS では「品物をなんらかの方法で試験した結果を，品質判定基準と比較して，個々の品物の良品・不良品の判定を下し，またはロット判定基準と比較して，ロットの合格・不合格の判定を下すこと」と定めている。つまり，検査は品物の良否またはロットの合否を定めることを意味する。

抜取検査とは，JIS では「検査ロットから，あらかじめ定められた抜取検査方式にしたがって，サンプルを抜き取って試験し，その結果をロット判定基準と比較して，そのロットの合格・不合格を判定する検査」としている。したがって，抜取検査は検査の一種といえる。

★解説

生産工場では品質を保証する 1 つの目的として検査を行っているが，検査の目的によって，

①受け入れ　②購入　③工程間　④最終　⑤出荷などの検査がある。①は外部より部品や素材を受け入れるときに，⑤は製品を工場より出荷するときに用いるなど，それぞれの目的にかなった名前がつけられている。

検査は，その機能が，次の 3 種に分けられる。

(1)**予防**　悪い製品が顧客や次工程に流れるのを予防する。

(2)**管理**　検査の結果によって工場の問題点を見出し，これを改善するという管理に役立てる。

(3)**情報**　検査データを整備し，解析して必要な情報を必要な個所へ供給する。

検査には**全数検査**と**抜取検査**がある。ロットの不良率が高く，検査コストが低いときには全数検査を，検査コストが高いときや，破壊検査のときなどは抜取検査を行う。

関連ページ

品質管理　318

■**検査と品質管理**　検査を主体とした品質管理は必ず失敗するという。品質管理は，あくまでも品質を工程の中で作り込むことを目的としているからである。このことは，いかに取り締まりをきびしくしても，社会，体制が安定していなければ，犯罪は減少しないことと同様である。

■**検査の組織**　検査部門は製造部門と独立したほうがよい，とする考え方が昔はあった。つまり，これは2権分立である。製造部門と検査部門が同一の部門にあると，品質の悪い製品に対して，検査部門は顧客の立場に立って検査を行うことができなくなり，つい製造部門のことを考えて，検査が甘くなるためである。しかし，この両者があまりにも対立的な立場にあるより，協調的に品質を改善することが望ましいので，最近は2権分立の考え方はあまり強調されていない。

■**抜取検査の原理**　抜取検査を行うには，次の手順が必要である(図1)。

(1)　ロットを明示する。このロットを構成する製品の数をロットの大きさという。これをNとする。

(2)　ロットより抽出する標本(品質管理では試料ともいう)の大きさnを定める。

(3)　ロットの合格・不合格の判定基準を定める。

この標本が不良個数や欠点数のときには計数抜取検査，長さや目方のようにx_iの値ではかられるときには計量抜取検査という。

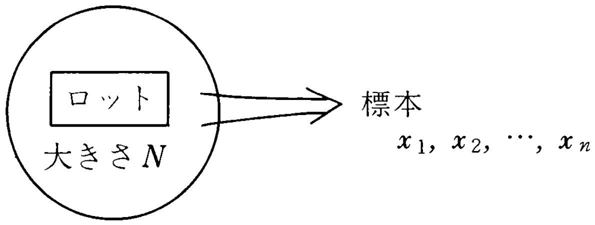


図1　ロットと標本

計数抜取検査のとき，よいロットの不良率をp_0，悪いロットの不良率をp_1とするとき，大きさnの標本中の不良個数kが

$k \leqq c$ のとき　　ロット合格

$k \geqq c+1$ のとき　ロット不合格

とすれば，cは合格判定個数となる。不良個数は二項分布にしたがうので，よいロットが不合格となる(これを**生産者危険**という)確率 $\overline{\mathrm{L}}(p_0)$ は，

$$\overline{\mathrm{L}}(p_0)=\sum_{k=c+1}^{n}\binom{n}{k}p_0{}^k(1-p_0)^{n-k}$$

となり，悪いロットの合格となる(これを**消費者危険**という)確率 $\mathrm{L}(p_1)$ は，

$$\mathrm{L}(p_1)=\sum_{k=0}^{c}\binom{n}{k}p_1{}^k(1-p_1)^{n-k}$$

となる。ここで，

$\overline{\mathrm{L}}(p_0)=\alpha$，$\mathrm{L}(p_1)=\beta$

($\alpha=0.05$，$\beta=0.10$ とすることが多い)

となるようなnとcを定めると，これは一意に定まる。このような抜取検査を**規準型抜取検査**という。抜取検査には，このほかに，選別型，調整型などがある。〔4〕〔5〕〔41〕〔48〕　　(真壁)

構造模型

実験により得られるデータの構造を，要因効果や誤差に分解して式で表したもの。平方和を意味のある要因別に分解するときに，構造模型は重要な役割を果たす。

★解説

ある化学反応工程で収率向上のため，**因子**として反応温度のみをとりあげ，最適条件を求めるため水準として 80°(A_1)，100°(A_2)，120°(A_3)の 3 水準を選び，各水準での繰り返しは 3 とし，計 3×3=9 回の実験を**無作為化**して行った。この場合に得られるデータは，i を因子 A の水準を表す添字とし，j を第 i 水準における実験番号を示す添字とすれば，x_{ij} と表すことができる。このときデータ x_{ij} は，μ_i を A_i 水準での真の収率，ε_{ij} を実験誤差として，

$$x_{ij}=\mu_i+\varepsilon_{ij} \qquad (i,\ j=1,\ 2,\ 3)$$

と考えることができる。さらに μ_i を

$$\mu_i=\mu.+(\mu_i-\mu.)=\mu+\alpha_i$$

ここで $\mu.=(\mu_1+\mu_2+\mu_3)/3$，$\mu=\mu.$，$\alpha_i=\mu_i-\mu.$

と分解する。μ は μ_1，μ_2，μ_3 の平均であるから一般平均，α_i は μ_i の μ からのずれであるから A_i 水準の効果とよばれ，定義より $\alpha_1+\alpha_2+\alpha_3=0$ となる。これより，x_{ij} を(一般平均)+(A_i 水準の効果=要因効果)+誤差 と考えると，これは，

$$x_{ij}=\mu+\alpha_i+\varepsilon_{ij}$$

と書くことができる。ここで誤差 ε_{ij} は，実験順序を無作為化しているので，偶然変動であるとみなしてよいから，互いに独立に平均 0，分散 σ^2 の**正規分布**に従う変量であると考えられる。

データの構造模型は，これに基づいて以後の検定，推定が展開されるため，非常に重要なものである。

関連ページ
因子 20
無作為化 384
要因効果 394
正規分布 200

実験順序	1	2	3	4	5	6	7	8	9
水準	A_1	A_2	A_2	A_3	A_1	A_3	A_1	A_3	A_2

表1　完全無作為化法による実験

日(ブロック)	1			2			3		
実験順序	1	2	3	1	2	3	1	2	3
水準	A_1	A_3	A_2	A_2	A_1	A_3	A_2	A_3	A_1

表2　乱塊法による実験

■乱塊法のデータの構造模型　前ページの例では，因子として反応温度をとりあげ，水準 A_1，A_2，A_3 を選び，繰り返し数3の計9回の実験を表1のように無作為化して行った。ところが，この実験は1日3回しか行えず，また収率は日による環境条件の違いなどの影響をうけると思われたので，日をブロック因子としてとりあげ，各ブロック内で A_1，A_2，A_3 の実験順序を無作為化して実験を表2のように行った。このときには，日による効果 β_j (ブロック効果)を誤差から分離できるので，データ x_{ij} の構造模型は，

$$x_{ij}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$$

となる。

■多因子要因実験のデータの構造模型

同様の実験で，反応温度のほかに反応時間 (B) も因子としてとりあげ，その水準を5分 (B_1)，10分 (B_2)，15分 (B_3) として，各水準組み合わせで2回繰り返し，計 $3\times3\times2=18$ 回の実験を無作為化して行った。このとき，データは x_{ijk} と表すことができる。i は反応温度の水準，j は反応時間の水準，k は繰り返し数を示す。2因子要因実験では，主効果のほかに2因子交互作用が考えられるから，x_{ijk} の構造模型は，

$$x_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk}$$

となる。

3因子実験の場合にも全く同様にして，データ x_{ijkl} の構造模型は，

$$x_{ijkl}=\mu+\alpha_i+\beta_j+\gamma_k+(\alpha\beta)_{ij}+(\alpha\gamma)_{ik}+(\beta\gamma)_{jk}+(\alpha\beta\gamma)_{ijk}+\varepsilon_{ijkl}$$

となる。

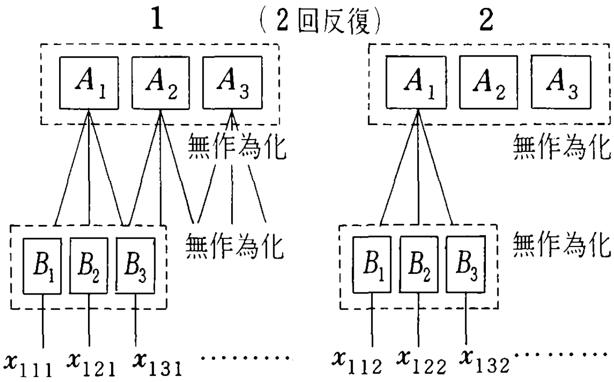


図1　1段分割による実験

■分割法によるデータの構造模型　前の2因子実験で反応温度の水準変更が難しい場合には，まず A の実験順序の無作為化を，つぎに B の各水準の実験順序の無作為化を行って，このような実験を2回反復することがよく行われる。これは，図1のように無作為化を2回に分けて行うので1段分割実験とよばれる。このときには，各無作為化ごとに共通の誤差が伴うので，構造模型は，

$$x_{ijk}=\mu+\gamma_k+\alpha_i+e_{(1)ik}+\beta_j+(\alpha\beta)_{ij}+e_{(2)ijk}$$

となる。γ_k は反復のブロック効果。

(宮村)

行動科学的意思決定論

OR は現状の意思決定の変革を目指すという意味で，本来規範論ないしは方策論である。しかし，現実の意思決定メカニズム，すなわち企業ないしは組織の解剖学的・生理学的知識をふまえた方策論でなければならない。行動科学的意思決定論（ないしは組織論）は，この観点から現実を分析した実態論である。

★解説

伝統的には，意思決定論は**経済人モデル**（economic man）によってきた。多くの OR モデルも例外ではない。経済人モデルでは，意思決定者はすべての代替案とすべての不確実性の範囲（環境条件に対する知識），代替案実行によって生ずる結果に対する知識，結果の良し悪しを評価する尺度などを完全に知りつくしており，意思決定者は目的関数を最大にする代替案を選択するとする**最適性基準**（optimality principle）によっている。

サイモン（H. A. Simon）は，経済人モデルに基づいた組織の意思決定やコントロールメカニズムは非現実的であり，現実に対する説明力もないとして，**管理人モデル**（admicistrative man）を提唱した。管理人モデルに基づく組織的意思決定は，その後サイアート（R. M. Cyert）とマーチ（J. G. March）によって発展された。その解説は次ページにゆずるが，これは非合理的な意思決定者を想定しているのではない。完全な知識をもった完全な合理性（economic rationality）を発揮できないとしても，限られた範囲ではあるが合理的な行動をとる人間（administrative rationality）をモデル化したものである。最適性基準に代わって，**満足性基準**（satisfactory principle）を用いる。

関連ページ

OR の実施理論
44

■合理的意思決定　意思決定とは，**事実前提**に基づいて行為代替案に対する結果を推定し，これを**価値前提**に照らして選好順位が最大となる代替案を選択することである。合理的意思決定論の情報的前提は，①**評価力**，②**計算力**，③**予測力**，④**洞察力**が完全であるという，**全知的合理性**(omniscient rationality)を前提としている。

■限定された合理性　現実の意思決定者が上記のような全知全能の合理性を持ちえないのは言うまでもなく，より現実的な**限られた合理性**(bounded rationality；初めてサイモンによって提唱)に基づく意思決定論の1つの姿がサイアートとマーチによって発展された行動科学的意思決定論である。

■行動的合理性(behavioral rationality)による意思決定

(1)　**葛藤の準解決(Quasi - resolution of Conflict)**　組織はさまざまな動機を持つ参加者からなり，利害に基づき組織内に**提携体**(coalition)が形成され，各提携体間に存在する葛藤について完全な合意が得られなくても，組織は存続している。すなわち葛藤の部分解決によりまがりなりにもバランスしながら組織は存続していく。これを**葛藤の準解決**という。

①多次元目標の希求水準と満足化原理

多次元目標を一次元化する能力を欠く(評価力不足)とき，現実には各次元ごとに目標として希求水準を定めておき，それを越えれば良しとする**満足化基準**(satisficing criterion)を用いる。

②局所的合理性(local rationality)

組織全体目標の追求ではなく，組織各部がそれぞれ目標を抱き，各次元ごとに希求水準を定めて行動する。これが**局所的合理性**である。

③許容水準意思決定ルール　これなら我慢できるという水準を設け，提携体間の交渉を通じてバランスを実現する。**組織余裕**(organizational slack)を設け，多少の目的の不一致が存在してもそれを吸収させる。

④目標への逐次注意ルール(sequential attention rules to goals)　組織は一度に多くのことを思考する能力(洞察力)を持たないため，一時点では1つの目標にのみ注意を集中しながら意思決定していく。

(2)　**不確実性の回避(uncertainty avoidance)**

①フィードバック反応意思決定ルール

生産計画をするとき，将来需要の予測を行って，それに基づき計画を立てるとされているが，実際には現在の在庫状況を見て，その過不足をフィードバックしながら生産計画を調整していく。

②談合された環境(negotiated environment)　環境の不確実性をそのまま受け取って反応するのではなく，不確実性を減らす方向に，さまざまな

談合や調整を行っていく。標準作業手続きの作成も，他者の行動を予測しやすくする不確実性削減のための一方法と考えられる。

(3) 問題本位の探索（problemistic search）

① **動機付けの強い探索**（motivational search） 問題が解決されず，欲求不満が残っている間は非常に動機付けが強く，満足できそうな案が見つかった途端，動機は消滅し，探索は打ち切られ，選択が行われる。

② **単純思考的探索**（simple minded search） 探索はありとあらゆる方向に向けられることはなく，問題の徴候の見られる周辺部とか，現在用いている解決策の周辺といった比較的単純な因果関係を想定して探索が行われることが多い。

③ **バイアスのある探索**（biased search） 過去の経験や身近に利用できる知識に頼るといったバイアスのある探索が行われる。

(4) 組織学習（organizational learning） 個人だけでなく，組織も不断に学習しながら，状況に反応し，よりよい意思決定へと進んでいく。

① **目標適応**（goal adaptation） 目標の次元は一定ではなく，状況に応じて変化していく。たとえば，当初は技術的実現が最大の目標であったものが，少し軌道に乗りはじめると，次は売上高という目標次元が出現してくるなどである。

② **期待適応**（expectation adaptation） たとえば，談合協定の相手が変化していくなどであり，当初競争相手と談合していたものが，それが一応片付くと，次に，政府機関と関係を持つように努めるなどである。すなわち，環境のある部分に注目を与え，他の部分には与えないことも学習する。**注目ルールにおける適応**（adaptation in atention rules）ともいう。

③ **探索適応**（search adaptation） これは探索方向に関する適応をいい，前述のバイアスを削減する方向に方向転換していくとか，探索の視野を拡大していくなどである。たとえば生産部門が費用削減に専念したとしても，生産部門の範囲だけで考えていたのではその成果に限度があり，生産技術部門の協力のもとに探索範囲の拡大を図るなどである。

以上，現実の組織の意思決定過程の調査の中から，主要概念を抽出し，それらの関係を規定し，図表化したのが次ページの図1である。ORのような分析手法が，このような意思決定のメカニズムの実態を無視した形で提案されるのでは，なかなかうまくいかないであろうことは明らかである。現実の意思決定過程はよいところもあれば，不十分なところもあるわけだから，不十分なところをどこまで，どうやれば変革可能かを，それぞれの置かれた状態に照らして考察しなければならない。（中野）

葛藤の準解決	不確実性の回避	問題本位の探索	組織学習
多次元目標の希求水準と満足化原理 局所的合理性	フィードバック反応意思決定ルール	動機付けの強い探索 単純思考的探索	目標適応 期待適応
許容水準意思決定ルール 目標への逐次注意ルール	談合された環境	バイアスのある探索	探索適応

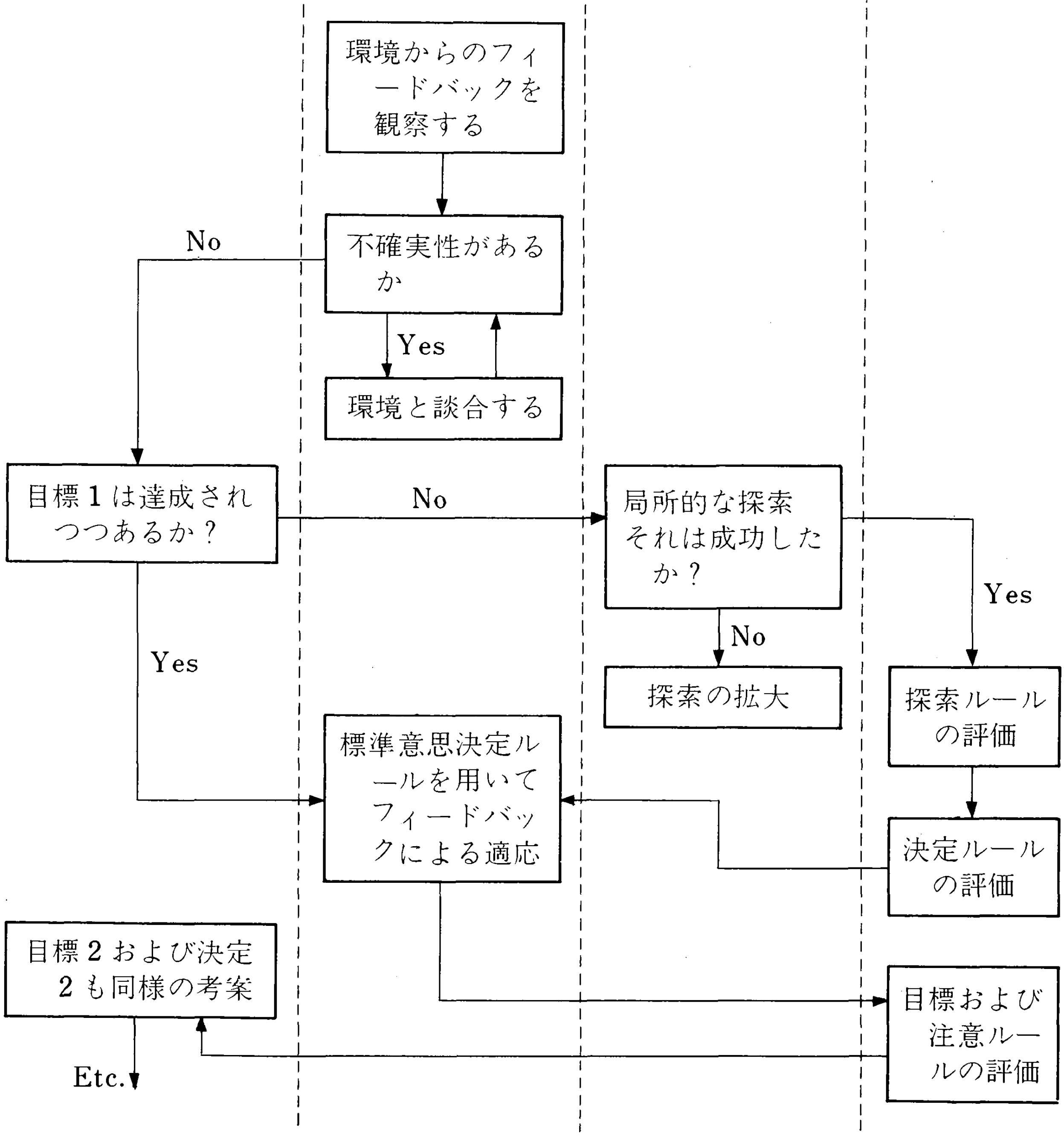


図1　組織的意思決定過程の一般的構造

効率指標の使い方

方策の経済性を示す指標の1つに効率指標がある。効率指標とは，方策から得られる「効果」と，方策に使用される制約された資源の「投入量」との比率である。

$$効率指標=\frac{効\quad果}{投入量}$$

資金，人員，スペース，原材料……といった特定の資源を使って効果が得られる方策が複数個あり，手持ちの資源が不足していて，すべての方策に投入できない場合に，どの方策に資源を優先して投入すべきかを検討するのに役立つのが効率指標である。

★解説

検討の対象に「複数の案(方策)の集まり」がリストされているとき，この案の集まりの基本的なタイプとして次の2つをあげることができる。

(1)**独立案**　リストされている案のなかから同時に複数の案を自由に組み合わせて採れるとき，リストされた案の集まりを独立的諸案，略して独立案という。

(2)**排反案**　リストされている案のなかから同時に2つ以上の案を採ることができないとき，リストされた案の集まりを排反的諸案，略して排反案という。排反案とは,いわゆる代替案のことである。

効率指標は上記の独立案での選択に使える指標で，排反案には直接使うことはできないので注意しなければならない。

効率指標の使い方は，独立案のおのおのの案について効率の値を求め，大きい順に，制約されている資源の量を使い切るところまで選択していくというやり方である。こうして選ばれた案から得られる効果の和を最大にすることをねらっている。

関連ページ

経済性の比較の原則　98

利回り法と差額投資　406

効率指標は実用的で便利な指標であるが，その使い方を誤ると真価を発揮しない。効率指標が使える条件とその使い方を整理して示しておく。

■効率指標が使える条件

(1) 検討の対象が独立案のタイプになっていること。排反案に直接使ってはいけない。

(2) おのおのの案に，効果と資源の必要投入量の値がきまっていて，かつこれらの値は，案について加法性がなりたっていること。加法性がなりたっているということは，複数の案の組み合わせをとったとき，その組み合わせから生じる効果と投入量がおのおのの案の和になっていることを意味する。

(3) 効率の分母になる資源はその使用可能量に制約があるもので，1種類であること。

■効率指標が適用できる問題状況

(1) 手持ちの資源量に制約があり，それをフルに活用して，全体で最大の効果が得られるような案の組み合わせを求めたいときに使える。

(2) 上記とは逆に，効果の総量に目標値が与えられていて，その目標を最少の資源投入量で達成するような案の組み合わせを求めたいときにも使える。

■効率指標の使い方

おのおのの案の効率の値を計算し，その値の大きい順に，資源の制約を満たすまで(あるいは目標の効果値を達成するまで)採用していく。

案	材料消費量〔kg〕	効　果〔万円〕	効　率〔万円/kg〕
A	10	30	3.0
B	10	55	5.5
C	15	100	6.7
D	15	120	8.0
E	20	130	6.5
F	20	125	6.3
G	30	140	4.7
H	40	160	4.0

表1　設計変更の独立案

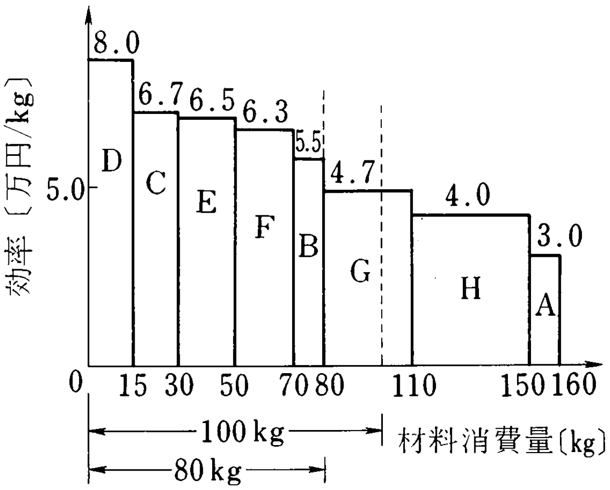


図1　独立案の選択

■効率指標による選択例

表1は，受注製品の設計変更による改善案のなかから，特定の材料を使う案をまとめてぬき出してきたものである。効果は，その案を実施することから得られる費用の削減額を表している。

これらの案に使う材料が80 kgしか手に入らないとき，すべての案を採用できなくなる。そこで，効率を計算して高い順にとっていくと，D，C，E，F，Bの各案を採るのがよいことになる(図1参照)。　　　　(中村)

効率指標の内外製区分への応用

内外製区分の問題は，特定の注文や仕事を社内で行うか外注に出すかを比較することである。社内の手持ち能力が手余り状態か手不足状態(→ 242 ページ)かによって，比較の計算の仕方が異なるので注意しなければならない。とくに手不足状態のもとでは，効率指標(→ 116 ページ)を有効に活用することができる。

★解説

内外製区分の計算を行う際には，内作のコストを正しくつかむことがまず大切である。内作のコストは，社外で仕事をした場合と比べて，社内で行うとどの種のコストがいくら増加するか，という見方で決めなければいけない。たとえば，人件費についてみると，もし定時作業時間内でこなせるなら内作コストとして考慮してはいけない。内作のために設備を買ったり，借りたりしない場合は設備費もかからないことになる。原価計算資料の製造コストを，内作コストとしてそのまま用いるのは危険である。

内外製区分の検討対象が複数ある場合は，手余り状態か手不足状態かによって計算の仕方が変わってくる。手余り状態の場合，すなわち検討の対象になっている仕事をすべて社内で処理できる場合は，おのおのの対象について個別に内作コストと外作コストを比較すればよいことになる。

手不足状態の場合，すなわち実施すべきすべての仕事を社内でこなせない場合は，社内の人員，設備などの能力を有効に使えるような内外製区分をしなければならず，**効率指標**が役立つのである。とくに社内で仕事する方が安くなることがわかっていても，能力不足でできないときなどは，損失を最小にする内外製区分を効率指標を用いて決めることができる。

関連ページ
経済性の比較の原則　98
手余り状態と手不足状態　242
効率指標の使い方　116

仕事	外作費(万円)	内作費(万円)	内作効果(万円)	工数(H)	内作効率(万円/H)
A	120	80	40	100	0.4
B	180	130	50	100	0.5
C	115	95	20	200	0.1
D	210	150	60	300	0.2
E	290	200	90	200	0.45
F	195	160	35	100	0.35
G	420	300	120	400	0.3

表1　内外作の検討

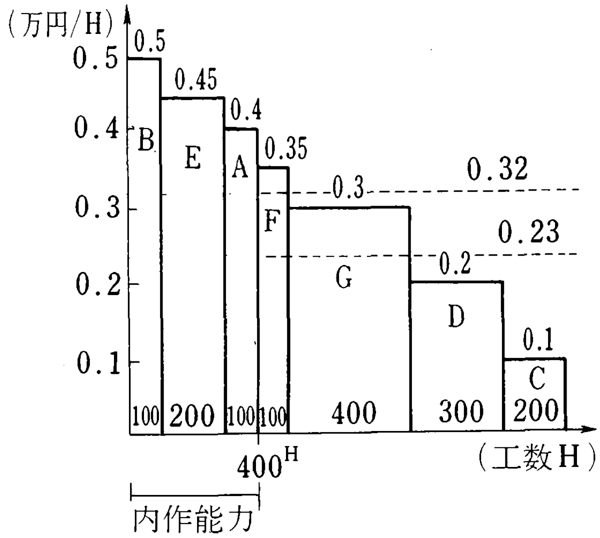


図1　内作の仕事の優先順位

■内作の順位づけの例　ある工場では，今月の納期でこなさねばならない仕事のうち，どれを内作し，どれを外作に出すかを決めなければならない。この工場の1カ月間の稼動可能工数は2000工数あるが，種々の事情で社内でつくることが決まっている仕事が1600工数分ある。したがって，余裕工数は400工数しかない。

一方，外作の方が有利だとわかっていたり，政策上外作に回す仕事もある。それらを除くと内作のほうが有利になる仕事は，表1に示す通りA～Gの7つである。これらすべてを内作すると1400工数になり，どれかの仕事を外作に出さねばならない。

表1の内作効果は，外作費から内作費を引いた値で，内作による利益である。

このような場合，内作効果(利益)を内作工数で割った**工数当たりの内作効率**の指標が役立つことになる。その値は，表1の右側に求めてある。

つぎにそれぞれの仕事を，内作効率の大きい順に並べると図1のようになる。

内作により得られる利益を最大にするには，内作効率の大きい順に，内作の手持ち工数の400Hまで内作の仕事として取っていけばよい。B，E，Aを内作にし，F，G，D，Cを外作へだすのがよいことになる。

仮にこの例で，残業のコストとして工数当たり0.23万円かければ，内作の能力工数を十分大きくできるとしよう。このときは，残業の工数当たりコスト(賃率)を図1のように記入して判断すればよい。GとFは工数当たり0.23万円かけることにより，内作による利益が得られるので，内作した方が得になる。残業コストが0.32万円/Hのときは，Fだけ内作に追加するのがよくなる。　(中村)

国民経済計算

国民経済計算(National Accounts)とは，国民経済の循環と構造を体系的に記録した経済統計で，①一国経済の中での生産・分配・支出の状況を明らかにした国民所得勘定，②産業間のモノとサービスの取引の状況を記録した産業連関表，③取引に伴う資金の流れを明らかにした資金循環表(マネー・フロー表)，④外国との取引の状況を記録した国際収支表，⑤一国経済の生産と国民生活の基盤である資産の大きさを示した国民貸借対照表，の5つからなっている。これら5つの勘定体系は相互に関連しており，経済活動をフローとストック，モノとカネの面から体系的にとらえられるシステムになっている。

★解説

わが国の国民経済計算は，1968年に国連から提示された「A System of National Accounts」に基づいて作成されている。それまでは，1953年の国連の基準による国民所得勘定が中心であった。その内容は，生産の成果は経営者と労働者に分配され，経営者はそれによって投資し，労働者は生産によって得た賃金によって消費を行い，これら投資や消費は，生産活動の起動力として作用する，という国民経済の循環の実態を明らかにする体系であった。その後，経済活動の諸側面を明らかにするための研究が進められた。とりわけ，1930年ごろレオンチェフ(W. W. Leontief)によって開発された産業間の投入・産出構造を示す産業連関表や，1952年にコープランド(M. Copeland)によって開拓された金融の流れを記録する資金循環表が各国で作成されるようになった。

今日では，これらの統計は経済政策の立案や経済分析に有用であることが広く認められている。

関連ページ
日本の統計　282
産業連関表　122
資金循環表　122

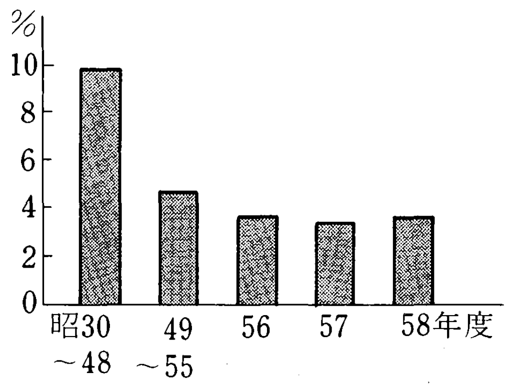


図1　経済成長率(年率)の推移

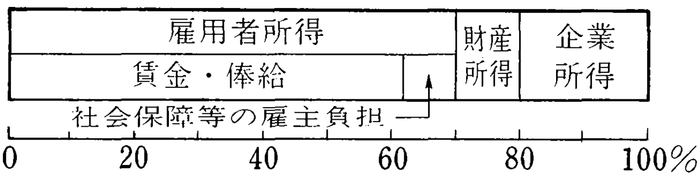


図2　所得の分配(57年度)

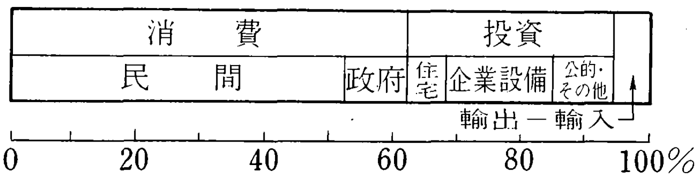


図3　所得の支出(57年度)

(1)**国民所得勘定**　この統計は，一定期間内の生産活動によって生み出された財貨・サービスを，生産・分配・支出の3つの面から把握したもので，経済統計の中で最も重要なものの1つである。

生産国民所得とは，生産の側面からとらえた国民所得のことで，一定期間内に，あらゆる産業部門で生産された生産額から生産に必要な原材料や電力などの中間生産物を差し引いた付加価値額の合計である。これを国民総生産という。その時々の時価で計算した国民総生産を名目国民総生産，特定年次の価格で評価し直した国民総生産を実質国民総生産とよんでいる。価格の換算率をデフレーター(deflater)という。

実質国民総生産は，昭和30年度から48年度の間に4.97倍(年率9.9％)に拡大したが，石油危機(48年秋)以降9年間に当たる48～57年度には，1.42倍(年率4.4％)に低下しており，石油危機以後，日本経済は低成長経済に転型したといわれている。

分配国民所得(National Income)は，生産活動の成果を賃金や企業利潤などにどのように分配されたかを示したもので，雇用者に支払われた賃金・退職金などの総額を雇用者所得といい，この所得は昭和57年度には，149兆円に達し，分配国民所得212兆円の70％を占めて，つぎにのべる民間消費の主な源泉となっている。

支出国民所得(Gross National Expenditure；国民総支出)は，消費，投資あるいは輸出などの最終需要(これに対し，生産に必要な原材料需要を中間需要という)を集計したもので，経済を動かす起動力と考えられている。

すなわち，経済政策としては，道路，橋などを作るための公共投資を増加させたり，金利(公定歩合)を引き下げることにより，企業家の投資意欲を刺激するなどによって投資を増大させ，これをテコにして生産活動を活発化させ，賃金や企業利潤を増やすことができる。

昭和57年度の国民総支出は，267兆円にのぼり，このうち，消費支出が185兆円(うち民間が158兆円，政府が

買った産業 / 売った産業	中間需要			最終需要			輸入(控除)	生産額
	農・林・水産業	鉱・建・製造業	その他の産業	消費	投資	輸出		
農・林・水産業	2	13	1	4	−	−	− 4	16
鉱・建・製造業	4	141	41	37	77	28	−30	298
その他の産業	2	48	56	135	−	7	− 5	243
雇用者所得	1	50	80					
営業余剰など	7	46	65					
生産額	16	298	243					

(単位：兆円)

表1　産業連関表の概形(昭和55年)

27兆円)で，69.1％を占め，投資が80兆円(うち民間が55兆円，公的投資が24兆円)で29.9％，輸出マイナス輸入が3兆円(輸出44兆円，輸入41兆円)で1.0％となっている。

一定期間内に生産された財貨とサービスの総価値額は，生産・分配・支出のいずれの面からとらえても一致することから，これを三面等価の原則とよんでいる。

(2)**産業連関表**　この表は，国民所得勘定の中の生産勘定を産業別に細分化したもので，各産業の生産に要した原材料の費用と所得を縦列に，生産物の販売先を横列とする行列形式にまとめたもので，投入産出表あるいはレオンチェフ表とよばれている。

産業連関表は，わが国では昭和26年にはじめて作られ，昭和30年以降は5年ごとに作成されており，55年表は，行政管理庁を中心に11の省庁の共同作業によって作成された。この表は，産業相互間の取引の状況がひと目でわかるとともに，これを補充する雇用表などによって，産業構造，雇用構造，分配構造，支出構造がとらえられており，これら諸側面の分析や予測に使われている。

たとえば，自動車の輸出が100億円増加した場合，それに見合うエンジン，タイヤの需要を通して鉄などの原材料生産を刺激し，これら産業の生産と雇用水準を高めるなどの効果を計量的にとらえることができる。

(3)**資金循環表**　この表はマネー・フロー表ともいわれ，経済取引に伴う金融の流れを，会社・個人・銀行・外国などの経済主体区分と，現金・貯蓄性預金・保険などの取引区分別にとらえられたものである。

(4)**国際収支表**　この表は外国との取引を記録したもので，外国との貿易，観光収入と支出，海外投資や外国からの国内投資の状況がわかるようになっている。

(5)**国民貸借対照表**　この表は生産や国民生活の基盤である土地，機械，住宅などの有形固定資産や貯金・株式などの金融資産について，国全体および種々の制度部門(民間企業，政府，銀行

（単位　ドル）

国　名	1人当たり国民所得	国　名	1人当たり国民所得	国　名	1人当たり国民所得
スイス	13,755	デンマーク	9,832	ニュージーランド	7,293
スウェーデン[1)]	12,998	西ドイツ	9,723	イタリア	5,509
ノルウェー	11,533	ベルギー	8,921	ベネズエラ	4,455
アメリカ	11,347	オランダ	8,820	アイルランド	4,262
フランス[1)]	10,824	フィンランド	8,692	ギリシャ	3,587
アイスランド	10,796	イギリス	7,875		
カナダ	10,065	日　本	7,729		
オーストラリア	9,895	オーストリア	7,691		

資料出所　経済企画庁「海外経済動向指標」（注）1981年，ただし 1) は1980年

表1　1人当たり国民所得

などの金融機関,個人家計)別に資産の存在高を表示したものである。昭和56年末の総資産は2,795兆円で，有形固定資産633兆円，株式を含む金融資産が1,330兆円となっており，総資産は前年末に比べ265兆円，4.7％増加している。

国民経済計算によって，日本経済がどのように発展し，現在どのような状況になっているか理解することができる。なお，国民所得統計は，経済運営に欠くことのできない資料なので，四半期ごとに推計され，各四半期の最終月の2カ月後に公表されている。これを四半期別国民所得統計という。これらの結果については，物価変動を調整した実質値のデータや，季節的要因によるデータのぶれを調整した季節修正値も計算・公表されている。とりわけ，実質の季節調整値の対前期比は，国民経済規模の拡大テンポを示すものであり，これを4倍した上昇率は，この傾向が1年間続くと仮定したときの経済成長率に相当するものであるから，国民経済規模の成長すう勢の指標(瞬間風速)として使われている。

また，人口1人当たりの国民所得は，その国の経済的豊かさの指標として利用されている。昭和56年現在の状況は上表に示すように，日本は，1人当たり7,729ドルで，イタリア(5,509ドル)の水準を上回っているが，アメリカ(11,347ドル)と比べると32％下回っている。（市野）

コンピュータ・ソフトウェア

ソフトウェアは，ハードウェアと並んでコンピュータシステムの車の両輪である。ハードウェアのコストパフォーマンスが年々向上するのに反して，ソフフトウェア開発の生産性の低さが問題化してきている。

★解説

■プログラムの作成過程 図1は汎用プログラミング言語(たとえばFORTRAN)で作成されたプログラムがコンピュータで実行され結果を得るまでの過程を示す。第1段階では，ソースプログラムがコンパイラまたはアセンブラで翻訳され，機械語で書かれたオブジェクトプログラムになる。

オブジェクトプログラムはソースプログラムの直訳であり，そのまま実行するには不完全である。そこで，連結編集プログラムによって，オブジェクトプログラムを実行可能な完全な機械語にする作業を連結編集という。連結編集後のプログラムをロードモジュールという。必要なデーターを入力すれば実行して，その結果を出力する。

プログラム作成のミスはエラーとよばれ，翻訳時に検出されたものはコンパイル時エラーとよばれる。同様に，連結編集時エラー，実行時エラーがあり，エラーをなくす作業をデバッキングという。

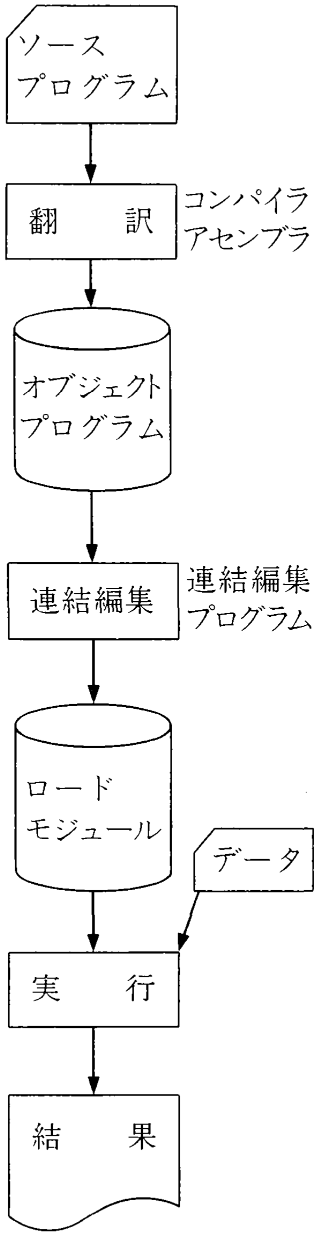


図1　プログラムの作成過程

関連ページ

コンピュータ・ハードウェア　126
汎用プログラミング言語　308
シミュレーション言語　168
汎用統計パッケージ　306
SAS　136

- プログラミング言語
  - 汎用プログラミング言語
    - 機械向き言語
      - 機械語
      - アセンブラ言語
    - 問題向き言語（高水準言語）
      - 手続き向き言語（FORTRAN, COBOL, BASIC, PL／1, C, Ada, PASCAL）
      - 非手続き向き言語（SAS, FOCUS, APL, RPG）
  - 特殊問題向き言語
    - シミュレーション言語（GPSS, DYNAMO, SIMSCRIPT）
    - リスト処理用言語（LISP, PROLOG）
    - 数式処理言語（FORMAC）
    - 数値制御用言語（APT）
    - プロセス制御用言語
    - 数理計画言語（LINDO, VINO, GINO, SAS/OR）
    - DBMS言語（TOTAL, ADABAS, IMS）

図 2　プログラミング言語の分類

■**言語処理プログラム**　アセンブラ言語で書かれたソースプログラムを，機械語のオブジェクトプログラムに翻訳する言語処理プログラムを**アセンブラ**という。**コンパイラ**は，手続き向き言語で書かれたソースプログラムを機械語に翻訳する。**プリ・コンパイラ**は，特定問題を記述するコマンドやパラメータから手続き向き言語のプログラムを生成する。

ジェネレータは，処理内容を記述したパラメータ(コマンド)をあらかじめ準備されているルーチン群の中に挿入することにより完全なプログラムを作りだす。非手続き向き言語の多くはこの形式をとる。一般に簡単に習得できるので，エンド・ユーザー向け言語ともいわれる。今後ますます多くの言語が開発されるものと期待される。ソースプログラムの命令を1個ずつ翻訳実行する言語処理プログラムを**インタプリタ**という。BASICやAPLがこの方式をとる。プログラムの作成や結果を会話的に得られるので便利である。

特殊問題向き言語は，特定問題向けのアプリケーションプログラムであるが，不特定多数のユーザーが利用できることを目的としている。　（新村）

コンピュータ・ハードウェア

ハードウェア(hardware)は，ソフトウェアとともにコンピュータ(計算機)システムの車の両輪である。主な分類としては，デジタル(digital)，アナログ(analog)，ハイブリッド(hybrid)の3タイプのコンピュータがある。

★解説

■アナログ計算機　歯車の回転角や電圧などの物理量を用いて表現された数をアナログ量という。アナログ計算機は，データを電圧の変化として入力し，加減乗除・積分・関数発生を行う回路を用いて算術演算を行う。出力はブラウン管上の波形として得られたり，制御信号として直接に電気信号のまま他の機械に入力されたりする。

アナログ計算機は実時間で動作するので，オンライン・シミュレーションや生産工程の制御に用いられる。しかし，ディジタル計算機と比較して，記憶容量が少ないこと，汎用性に乏しいこと，精度的に劣るなどの欠点をもつ。

■ディジタル計算機　数や文字を離散的な信号を用いて表すことをディジタル(計数型)表示という。ディジタル計算機は，データを0と1で表される2進法表示を用いてコード化し，算術演算・論理演算を高精度で行う。2進数表示は，論理回路や記憶装置を作るのが容易なため広く用いられている。

右ページにハードウェア構成を示す。中央処理装置(CPU)に用いられる主な論理素子によって，第1世代(真空管)，第2世代(トランジスター)，第3世代(IC)，第3.5世代(LSI)と歴史的に分類されている。

■ハイブリッド計算機　ディジタル計算機とアナログ計算機をAD変換器(Analog-to-Digital convertor)とDA変換器で結合し，両計算機の利点を取り入れた計算機。

関連ページ
コンピュータ・ソフトウェア
124
シミュレーション
166

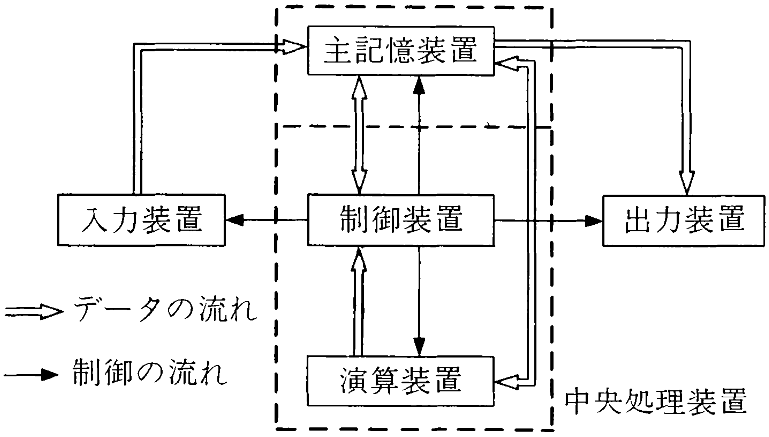


図1　ディジタル計算機のハードウェア構成

■ハードウェアの構成　われわれが認識できる形のデータやプログラムをコンピュータが認識できる機械コード(電気記号)に変換する装置を**入力装置**という。電気信号に変換されたデータとプログラムは，主記憶装置に格納される。この主記憶装置は即時アクセス記憶装置ともいわれ，(プログラム)制御装置によって直接に番地指定できる。

主記憶装置内のプログラムは**制御装置**に読み込まれ，逐時的に解読される。その結果，制御装置から制御命令が主記憶装置と**演算装置**に出され，主記憶上のデータが演算装置に読み込まれて演算が行われる。演算結果は，ふたたび主記憶装置に格納される。

制御命令はこのほか，入力装置と**出力装置**にも出される。出力装置に出された命令に基づき，コンピュータ内部に機械コードで格納されているデータが出力される。入力装置としてはカードリーダーが，出力装置としてはラインプリンターが代表的である。しかし，最近は人間がコンピュータと対話して即時的に結果を得るリアルタイム処理が普及し，入出力のいずれにも使える端末が増えてきている。グラフィックディスプレイ端末は，文字や数字のほか，グラフが出力できる入出力端末である。

一方，磁気テープ装置(MT)や磁気ディスク装置は，入出力装置として用いられるが，主記憶装置の容量不足をおぎなうことに比重があり，補助記憶装置とよばれる。補助記憶装置上のデータは，制御装置から出された制御命令により主記憶装置に格納される。また，制御命令に従い主記憶装置のデータを補助記憶装置に出力できる。

■ノイマン型計算機　1946年にペンシルバニア大学のモークリとエッカートは真空管を用いた世界初の電子計算機を開発した。プログラムごとに配線をしなおす方式を採用していた。

これに対し，フォン・ノイマンが提案した「プログラム記憶方式」が第1世代以降今日までの計算機の主流をなすディジタル計算機の基礎を作った。そこでノイマン型計算機ともよばれる。特徴は，プログラムをデータと同じく主記憶装置に格納し，制御装置で逐次的に解釈・実行を行う。

この方式以外の計算機を非ノイマン型とよび，第5世代以降の中心になると考えられている。　(新村)

在庫モデル

物の流れを時々堰止めると，一時的に溜まる。これが在庫である。原材料や部品にしても製品にしても，それらの在庫は生産の一様化のための緩衝装置として働く。しかし，在庫切れも困るが過剰在庫は無駄を招く。適正な管理方式を見出すため，需要の構造，調達の状況，費用のかかり方などに応じ，在庫のでき方を表すのが在庫モデルである。

★解説

■在庫変動の原因　中間製品であれば後工程の生産水準，最終製品であれば販売量が**需要**と考えられるが，その変動及び，部品の購入，完成品の受け入れといった流入の変動が在庫変動の原因である。納入に際しては，納入指示(**発注**という)から納入までの時間(これを**調達期間**とよぶ)も変動の要因となる。したがって，需要変動を的確に予測し，調達期間を減少させることができれば，在庫は常に適正在庫でありつづける。そのためには，需要や在庫の状況を直ちに在庫管理に反映できる態勢，特に情報システムを整備することが大切である。

■在庫管理方式　ふつうの在庫モデルでは，「いつ，いくら」発注するかを定めるようになっている。その定め方によっていくつかの在庫管理方式(たとえば発注点方式や定期発注方式)ができるが，どの品目はどの管理方式によるのがよいかなどを定めるため **ABC分析**が利用されたりする。

■ MRP　主として OR の分野で研究されてきた在庫管理方式は，あまりにも理想化した仮定のため，現実の工場での部品在庫には当てはめにくい。部品の需要は最終製品の需要に応じて，それらがいつ必要になるかを考慮し，電算機を用いて，常に各部品の必要量，在庫量，既発注量を把握して，必要な督促の時期を失しないようにする在庫管理方式が最近多くの成果を上げている。これを **MRP** とよぶ。

関連ページ

ABC分析　300

発注点方式　296
定期発注方式　244
MRP　38

■**基本モデル**　最も基本的な在庫モデルでは，需要が既知で時間に関し一様とする。1期間当たりの需要量が r，在庫量1単位1期間の保管費が h，1回の発注量は x で，発注費は K としよう。品切れは許されないものとし，調達期間も0であるとすると，在庫が0になったとき発注すればよい。このとき，在庫変動は図1の実線のようになる。1期間の発注回数は r/x となるし，延べ在庫量は図1の影の部分の面積で表され，三角形の面積の公式から $x/2$ である。結局，在庫変動費は，

$$C(x)=Kr/x+hx/2$$

で与えられるから，これを x で微分して0と置き，$C(x)$ を最小にする x^* を求めると，$x^*=\sqrt{2Kr/h}$ となる。x^* を**経済発注量**，略してEOQという。

■**発注点**　調達期間が0でないときは在庫量が0になる前に発注しておかなければならない。図1に示したような調達期間であれば，鎖線の位置まで在庫が下がったときEOQだけ発注すれば，在庫量がちょうど0になったとき，発注済みの品物の納入が行われ,在庫切れを生じない。鎖線の位置の在庫量を発注点とよぶ。

調達期間や需要が確率的に変動する場合には**安全余裕**を置かなければならない。それらについては296ページ参照。

■**在庫費用**　上のモデルでは発注費と保管費だけを考えた。ふつう，それらは発注量の関数として図2のようになる。したがって，総在庫費が最小になる点が1つ存在する。なお，在庫費の内容は次のようなものと考えられている。

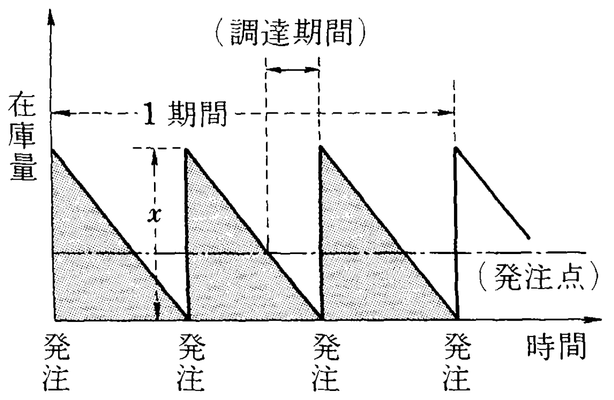


図1　確定的在庫モデル

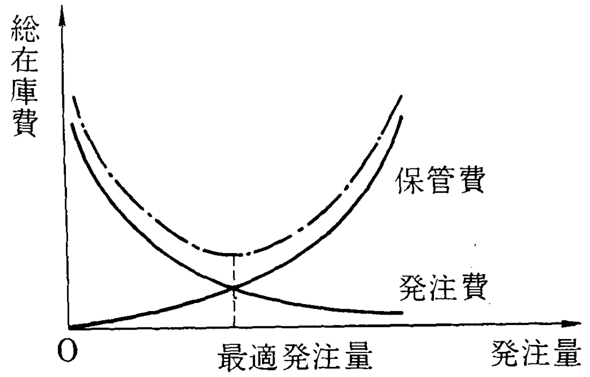


図2　在庫費用のバランス

発注費：事務用品，通信，運搬，受入検査の費用及び人件費など。

保管費：腐敗,変質,破損,紛失,目減り,漏洩などによる減耗費,建物設備,地代などの物件費,倉庫人件費,場内運搬費及び保管物品費の金利など。

段取費：生産ラインで段取りの費用。

陳腐化損失：流行の推移や技術改新のため在庫中に減少する経済的価値。

品切れ損失：品切れによる信用失墜，納入遅れによる延滞賠償金，品切れ品補充のための手配費用など。

その他：売損じのときの機会損失など。

（森村）

最大流問題

ネットワークを構成する各アークに容量が与えられているとき，その容量以下の物を各アークに流して，全体として最も大量の物を流す流し方を求める問題をいう。流通網の上での配送計画を作る際に利用される。その他，CPMの解法の基礎としてなどの応用もある。

★解説

■ラベリング法　最大流問題の解法にはいくつかが知られているが，ラベリング法とよばれる算法が名高い。これは，ソースとよばれるノードからシンクとよばれるノードに，各アークの容量以下の流れを流し，もっと流せるアークを探し出しては追加するという方法を繰り返して，遂には最大流量に到達しようという方法で，各ノードにラベルをつけながら余裕のあるアークを探していくためこの名がある。

各アークに下方の容量制限があるものとし，それがすべて負ならば，流量0という流れが条件を満たすことが明らかであるから，そこから出発する。また，もしそれが正であれば，条件に合った流れが存在するかどうかを調べ，とにかく初期の流れを1つ見つけ出さなければならない。

具体的なラベルの見つけ方は右の図1を例にして右ページで解説する。図1のアークはすべて一方向への方向がついており，アーク上に書かれた数値が，その容量を示している。

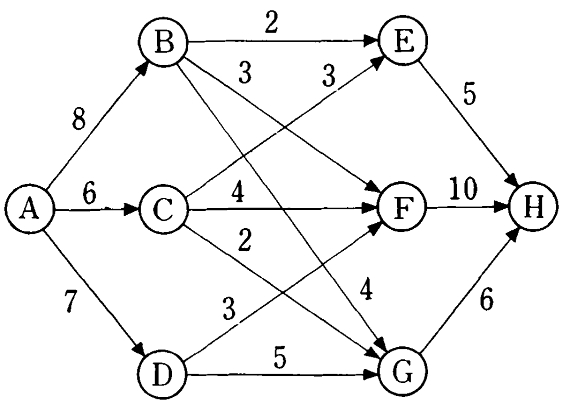


図1

ノードAがソースであり，ここからすべての流れは流れ出してシンクであるノードHに至る。すべてのアークに0の流れが自明な流れになっている。

関連ページ

ネットワークモデル　284
CPM　142

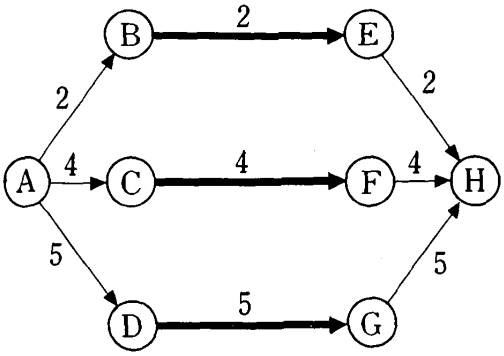


図 2

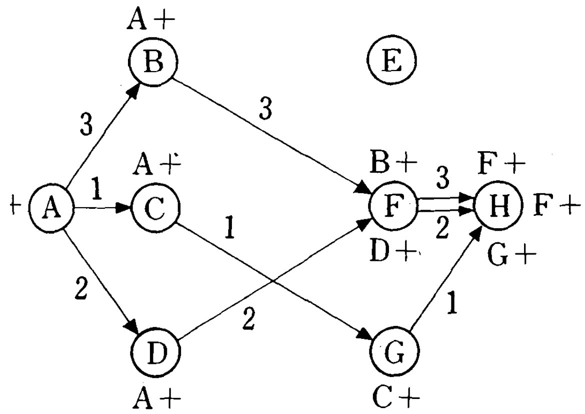


図 3

手順①　ソースからシンクに至るルートごとに，ルート上のどこかのアークの容量で制限されるまで流す(図2)。太線のアークで限度いっぱい流れている。

手順②　ソースからのアークで容量に余裕があれば，始点に＋をつけ，アークの終点に**A＋**と記す。ルート上次のアークに対しても同様，余裕のあるアークにはその終点に＋とそのノード名を書く。そして，シンクまで＋の記号がつけば，**ブレークスルー**したといい，そのルート上に流せる最大量を流す。図3にはそのようなルートだけ記した。図2に加えると図4のような流れが得られる。

手順③　図4で手順2のような方法

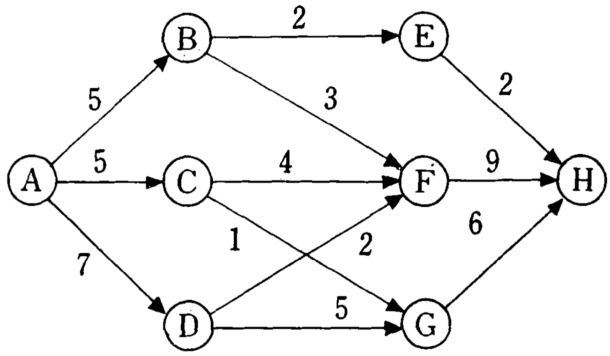


図 4

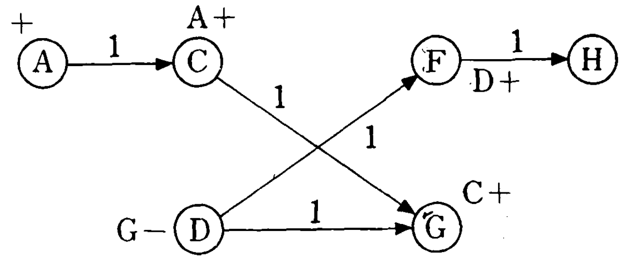


図 5

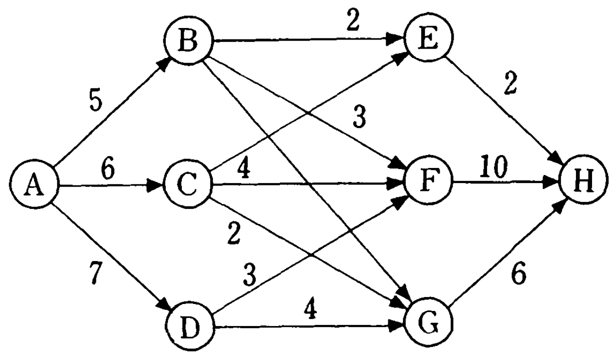


図 6

で見つかるルートはないが，**A→C→G**のルートでは**D→G**の流れを減らすことによりブレークスルーする。つまり，容量いっぱいのアークの終点に＋がついたときは，その始点に－をつけ，そこから手順2の方法でルートを探す。こうして図5のルートが得られた。

手順④　以上の方法でブレークスルーするルートがなくなれば最大流を得る。結果は図6のとおり。　(森村)

財務シミュレーション

外部環境条件の変化や企業のとる方策の違いが，企業全体にどのような影響を与えるかを検討するために，企業の諸活動をモデル化した企業モデルを用いて行うシミュレーションの1分野である。特に，そうした影響を損益計算書や貸借対照表，資金繰表などの財務諸表の上での変化として，財務的側面から把握しようと意図して行われる場合に財務シミュレーションとよばれる。

★解説

財務**シミュレーション**で使われるモデルは，対象とする企業の業種や規模の違いに加えて，シミュレートする期間の違いや，対象とする企業活動の重点の置き方の違いにより多様であるが，大別して次の3タイプがある。

(1)**予算編成モデル**　予算編成の手順にしたがい，各種予算上の諸数値の算出過程をそのまま記述することにより，企業の諸活動をモデル化するもの。

(2)**財務モデル**　企業の状態を貸借対照表の諸残高勘定により，また企業の活動を製造原価明細書，損益計算書，資金繰表などの諸取引勘定としてとらえ，各勘定の値を推定式や勘定間の定義式などで記述することにより，企業の諸活動をモデル化するもの。

(3)**経営計画モデル**　生産や販売，調達など企業の諸活動を，人・物・金などの経営諸資源の投入による製品やサービスの産出過程として，その実体的側面からモデル化する。さらに財務数値に置き換える会計手続きを付加することにより財務諸表を出力できるようにしたもの。このタイプのモデルの利用は，企業全体の活動の計画や決定の検討にあるため，経営計画シミュレーションとか，企業シミュレーションとよばれることが多い。

関連ページ
シミュレーション
166

■発展過程　財務シミュレーションは計算機の利用を前提としているため，計算機やその利用技術の発達にともなってモデルの形態やモデル化の形式，シミュレーションの方法もさまざまに多様化してきている。

マテジッヒ(R. Mattessich)が1964年に，計算機による予算編成モデルを発表して以来，多くのモデルが開発され，シミュレーションが行われてきている。

当初は，マテジッヒ流の予算編成モデルや年度利益計画モデルが中心であったが，計算機の大型化と経営情報システム(MIS)のブームにともない，MISの予測情報・未来情報を作成する機能を果たすものとして大規模な経営計画モデルが作られた。他方，金融機関や調査機関を中心に，貸付先の財務状態の予測を行う汎用の財務モデルや財務政策の検討用の財務モデルが作られてきた。

1980年代に入り，パソコンが普及するにしたがって，簡単なモデルでしかもさまざまに想定できる情況に対応して容易にシミュレートできるような操作性の高いモデルが要求されるようになり，簡易言語を用いたモデル化とシミュレーションが行われだしている。

■意思決定支援システムとの関係　計算機の対話形の利用技術の発達にともない，完成したモデルを使って検討情報を得るという考え方から，モデルそのものを作り出す過程から計画や政策の決定過程までを支援する1つのシステムを作り上げ，その支援過程で適宜モデルの作成や修正を行いながらシミュレーションも行うようにする，といった考え方が生じてきている。

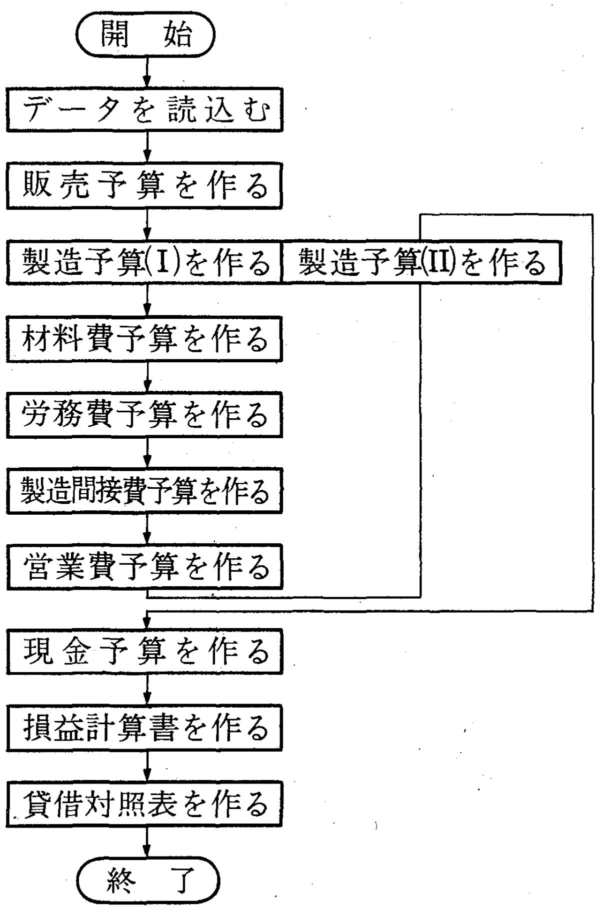


図1　マテジッヒのモデルの流れ図

シミュレーションの方法も，「もしこうすると(こうなると)どうなるか」といった実行のさせ方だけでなく，感度分析やレンジ分析，リスク分析といった分析手法に則した方法や，「こういう目標を達成するためにはどういう方策をとるのがよいのか」といった，方策探索用の数理計画法(→186ページ)をモデルに組み込んでシミュレートする方法も行われている。　(福川)

最尤法

点推定により母数の推定を行うに際して，どのような方法を用いたら，よい推定量を得ることができるであろうか。これに答えてくれる1つの方法が最尤法(maximum likelihood method)である。これは，最も尤もらしい方法で推定する，ということである。ただし，最尤法による推定量は，標本数が相当大きければ，近似的に有効性・十分性・一致性をもつが，不偏性は保証されない。

★解説

最尤法の考え　たとえば貨幣をn回投げたところ，ちょうどk回表が出たとする。このことから，1回の試行で表の出る確率pを推定するのに，次のように考える。すなわち，表がk回出るということは，pの値がいくらであると考えたときに，もっとも起こりやすいか，という観点に立って，**母数**pの**推定値**pを求める。いまの場合，$p=k/n$となることがすぐわかる。これが，最尤法の原理である。

より一般的には，最尤法による推定の仕方は，次の手順による。いま，**母集団**の**確率密度**を$f(x;\theta)$とする。ただし，θは推定しようとしている母数である。この母集団から大きさnの標本をとったとき，標本値が$(x_1, x_2, \cdots, x_n)$であったとする。このとき，次のような関数をつくる。

$$L(\theta)=f(x_1;\theta)\cdot f(x_2;\theta)\cdots f(x_n;\theta)$$

このような関数$L(\theta)$のことを**尤度関数**(likelihood function)という。この$L(\theta)$を最大にする$\theta=\hat{\theta}$が存在するとき，$\hat{\theta}$をθの最尤推定量という。最尤推定量は通常，**尤度方程式**とよばれている方程式

$$\frac{\partial \log L(\theta)}{\partial \theta}=0$$

の解として求められる。

関連ページ

点推定　254
母数　254
母集団　260
確率密度　410
標本値　260

■**最尤法による計算**　右の例は，母集団分布が正規分布(→200 ページ)であるとして，それからとられた大きさ n の任意標本(→54 ページ)に基づいて，母平均 μ，母分散 σ^2 を推定するにはどうしたらよいかを示している。すなわち，最尤法により，母平均 μ を推定するには，標本平均 $\overline{x}$ を推定値とすればよいし，母分散 σ^2 を推定するには，標本分散 s^2 を推定値とすればよい。そこで，$\hat{\mu}=x$ でよいわけを，前ページの手順に従って示してみる。この場合，母数 θ として母平均 μ を考えているので，密度関数 $f(x;\theta)$ を，$f(x;\mu)$ と明記することにして，尤度関数をつくると，

$$L(\mu)=\left\{\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x_1-\mu)^2}{2\sigma^2}\right)\right\}$$
$$\left\{\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x_2-\mu)^2}{2\sigma^2}\right)\right\}$$
$$\cdots\left\{\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x_n-\mu)^2}{2\sigma^2}\right)\right\}$$

上式両辺の対数をとり，μ について微分して 0 とおくことにより，尤度方程式

$$\frac{1}{\sigma^2}\{(x_1-\mu)+(x_2-\mu)+\cdots$$
$$+(x_n-\mu)\}=0$$

が得られる。この解 μ を $\hat{\mu}$ とすればよいので，

$$\hat{\mu}=\frac{1}{n}(x_1+x_2+\cdots+x_n)=\overline{x}$$

つまり，$\overline{x}$ が母平均 μ の最尤推定値である。

(例)　正規母集団からの大きさ μ の標本，

$$(x_1,\ x_2,\ \cdots,\ x_n)$$

$$\Rightarrow\left\{\begin{array}{l}\text{母平均 }\mu\text{ の最尤推定値;}\\ \qquad\hat{\mu}=\overline{x}\ (\text{標本平均})\\ \text{母分散 }\sigma^2\text{ の最尤推定値;}\\ \qquad\sigma^2=s^2\ (\text{標本分散})\end{array}\right.$$

■**投銭実験の場合**　前ページで述べた投銭実験においては，$\hat{p}=k/n$ が p の最尤推定値になることを示そう。そのために，各回の試行結果に対応して，確率変数 $X_i(i=1,\ 2,\ \cdots,\ n)$ を考え，表が出れば 1，裏ならば 0 を与えるものとする。そうすると，確率変数 X_1 は，1 と 0 の値をそれぞれ p, $q=1-p$ でとることになる。このような確率変数 X_i のしたがう分布のことを 0−1 分布とよんでいるが，その確率密度は，

$$f(x)=p^x q^{1-x}\qquad (x=0,\ 1)$$

と書ける。そこで，0−1 分布の母数 p の最尤推定値を求めようというのであるが，この場合の尤度関数は，

$$L(p)=(p^{x_1}q^{1-x_1})\cdots(p^{x_n}\cdot q^{1-x_n})$$
$$=p^{x_1+\cdots+x_n}\cdot q^{n-(x_1+\cdots+x_n)}$$
$$=p^k\cdot(1-p)^{m-k}$$

よって，尤度方程式は，

$$\frac{k}{p}-\frac{n-k}{1-p}=0$$

これを満足する p が最尤推定値 p で，$\hat{p}=k/n$ になる。　　(牧野)

SAS(Statistical Analysis System)

SASは代表的な汎用統計解析システムであるとともに，意思決定支援システム(DSS)または，第4世代言語と考えられる。簡単なSAS言語は，DATAステップとPROCステップに大別でき，SASスーパーバイザーの下で一元管理されている。その設計思想は「all in one system」，すなわちSASのみですべてのデータ解析の要求を満たすことを目標としている。

★解説

■ All in one system　1つのシステム(SAS)で，データ解析を希望する不特定多数のユーザーのさまざまな処理を，簡単な言語で実行できることを目的としている。このため，①単に統計処理にとどまらず，データの並べ換え(ソート)などの一般の汎用情報処理機能までも提供し，ユーザーの便に供すること，②ユーザーの所有するデータをデータベース(DB)に1元管理し，解析に便利な形に変換できること，③これらの異なった種々の要求を統一的に簡単に記述できる言語の体系であること，を実現している。

■ SAS言語　従来の汎用プログラミング言語と異なり，習得の容易な自由記述形式のSAS文から構成されている。DATAステップは，SASデータベース(DB)の運用・管理・検索・レポートライティング・モンテカルロシミュレーションなどを行う。PROCステップは，統計処理のみならず，時系列解析・シミュレーション・グラフィック出力・OR手法・作表・各種ユーティリティが簡単な命令で実行できる。この他，MATRIXはユーザーが簡単にプログラミングできる異色の高級言語であり，SASの提供しないプロセジャーの開発がフォートラン言語による開発に比べ格段に速く行える。

両ステップとも，簡単な場合には2～3個のSAS文で実行できる。

関連ページ
汎用統計パッケージ　306
統計パッケージのファイル管理　264
統計シミュレーション　256
汎用プログラミング言語　308
モンテカルロ法　388
シミュレーション言語　168

■SAS言語　SAS言語は，SASスーパーバイザーのもと，DATAステップとPROCステップに大別できる。SAS言語は予約語で始まり，セミコロン(;)で終わるSAS文からなる。SAS文は自由記述である。すなわち，カードやTSS端末のどこから始めてもよい。1枚のカードに複数のSAS文や，1つのSAS文を複数のカードに書いてもよい。

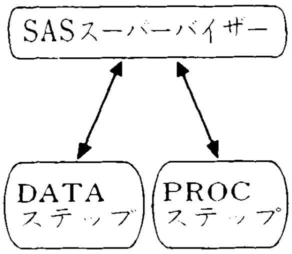


図1　SAS言語

DATAステップは，外部ファイルを入力し，SAS固有のファイル(SASデータセット)としてSASデータベースに登録する。また，任意のデータセットの変数を削除したり，値の変換を行ったり；個別データの操作が行える。PL/1に似たSASプログラミング文でプログラミングでき，レポートライティングなどが行える。データセットを外部ファイルとして出力することもできる。

PROCステップは，DATAステップで作られたデータセットを用いて処理を行う。たとえば，1000人の社員の50項目の人事ファイルをデータセットAとして，この50項目の平均や合計や最大・最小値などの基礎統計量は，次のPROCステップで計算できる。

```
PROC  MEANS  DATA=A ;
VAR  _NUMERIC_ ;
```

■MATRIXプロセジャー　汎用統計パッケージが提供しない処理を行いたい場合，普通は別途フォートランなどでプログラムを作成しなければならない。しかし，行列を処理(演算)の単位とする言語があれば，比較的容易にプログラミングできる。SASのMATRIXプロセジャーが提供するMATRIX言語は，汎用統計パッケージのプロセジャーと個別ユーザーが作成するフォートラン・プログラムなどとのギャップをうめるユニークな言語である。

すなわち，通常のプログラム言語は，スカラー量が処理の単位である。このため，ステップ数が増えてみにくくなり，人間の思考と対応しにくい。MATRIX言語は，処理単位が行列であるため，行列AとBを足して行列Cにすることや，重回帰分析の**デザイン行列**をX，目的変数値のベクトルをY，係数ベクトルをBとして次の各1ステップで実行できる。

```
C=A+B ;
B=INV(X'*X)*X'*Y ;
```

このため，プログラム自体がアルゴリズムのドキュメントになる。しかも生産性の向上が計れる。統計処理は，データ行列から派生した何らかの行列の操作と考えられるから，統計教育上からも有効な道具と考えられる。また，スカラーやベクトルも一種の行列であり，ユーザーが次元を与えなくても

SASが自動的に認知してくれる。

■データ解析の流れ

データ解析は，次の過程に分かれる。

①データの収集・編集

②解析

③解析結果の表示

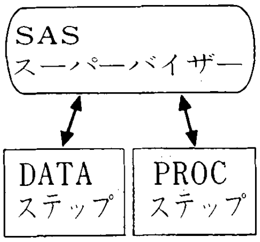


図2　SASの2元管理

①は広い意味でのデータ管理機能であり，個別データに対する操作やデータファイルに対する操作，さらに検索機能が考えられる。

③の解析結果の表示は②と結びつくが，解析結果の一部をシステムの標準ファイルに出力し次の解析(グラフィック，レポートライティングを含む)に利用する場合は，データ管理に組み入れられる。すなわち，図2にみるようにデータ管理を行うDATAステップとプロセジャーの実行をつかさどるPROCステップに機能分化し，スーパーバイザーがこの2元管理を行う。

両ステップを統一的に実行するSAS言語は，強力なスーパーバイザー機能により自由記述でユーザー負担の少ない言語である。

■プログラムの歴史　プログラム言語は，マシン能力の増大とともに3世代，すなわち，①機械語，②アセンブラー，③フォートラン，コボル，PL/1，などと変遷してきた。マシン負荷の増大と引き換えに，ユーザーの負担を軽減

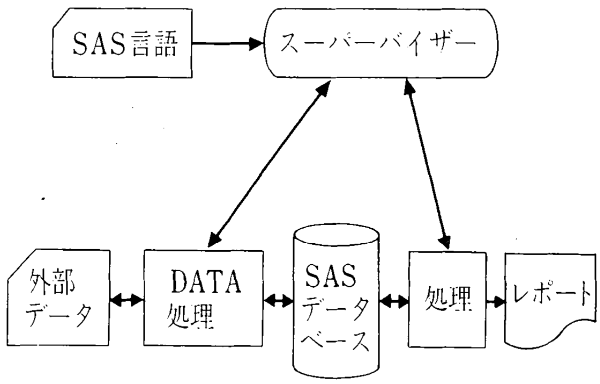


図3　SASの流れ

するための歴史といえる。SASは，フォートランやPL/1などで開発されているが，さらに簡単な言語でユーザーが利用できることから，第4世代言語とか超言語などとよばれる簡易言語である。簡単な処理を行う場合は，半日程度で習得できる。また，DSS(意思決定支援システム)の入り口に位置している。

■スーパーバイザーの役割　SASスーパーバイザーは，データベースの運用管理を行うDATAステップとプロセジャーの実行を行うPROCステップの管理を行う(図2)。各ステップごとに文法チェックを行い，管理情報を作成するとともに，正しければ所定の処理を行わせる。終了後，コントロールは再びスーパーバイザーにもどり，図3のような連続処理が継続される。DB上のデータは，後日再度処理に利用できる。

またスーパーバイザーのparsing generatorの機能を用いれば，ユーザーが作成したプログラムをSASのプロセジャーとして登録できる。すな

わち，システムの拡張性に富んでいる設計思想といえる。

〔SAS本体の代表的プロセジャー〕

MATRIX　マトリックス言語によるプログラミング。

記述統計量　基礎統計量，クロス集計，相関分析。

多変量解析　正準相関，主成分分析，因子分析。

回帰・分散分析　重回帰分析，分散分析，共分散分析，説明変数のすべての組み合わせの回帰分析，変数選択法，非線型回帰推定，ノンパラ，t 検定。

クラスタリング　階層型，大規模データ用，変数のクラスタリング。

質的データの分析　FUNCAT，プロビット分析。

スコアリング　順位の計算，係数行列とデータ行列の積，データの正規化。

判別分析　(正準)判別分析，最近隣判別分析，変数選択。

レポートライティング　各種チャート図，散布図，簡易プリント。

ユーティリティとその他　SASデータセットまたはSASデータライブラリーの履歴情報の出力·削除·編集・OSファイルのコピー，ソート，BMDP・SPSSなどとのインタフェイス，区分データセットの操作とコピー，未使用エリアの解放，テープのコピーとラベル情報の出力，変換テーブルの登録。

〔**増補版**〕SASユーザーが作成したプロセジャー。

MDS，Coxの生命表分析，多重ロジスティックモデル，遺伝子データの解析，LAV解析，箱ひげ図，時系列と横断面モデルの回帰，リッジ回帰，入力データのチェック，ガットマンの尺度解析，マルチプルテストの項目分析。

〔**SAS/ETS**〕　時系列解析とシミュレーション。

予測　センサス局法，ARIMAモデル，定常時系列解析，スペクトル解析，自己回帰型の誤差項をもつ回帰分析，多項回帰式または指数平滑法と自己回帰の組み合わせ。

シミュレーション　線型方程式系の係数の推定とシミュレーション，非線型システムのモデル化・係数の推定・シミュレーション。

レポートライティング　表形式の簡易言語，ローン計算(mortgage表)

〔**SAS/GRAPH**〕　各種グラフィック端末への出力。

各種チャート図，等高線図，各種地図と図形上への統計量の出力，折れ線グラフと散布図，SASの出力結果のグラフ出力，テキストラインの出力，3次元図作成。

〔**SAS/OR**〕　LP，PERT，割り当て問題，CPM法，ネットワークフロー，ガンチャート，輸送問題。

（新村）

残差分散

重回帰分析においてデータ数をN，用いる説明変数の個数をpで表す。実測値と回帰による推定値との差の平方和を，$N-p-1$ で割ったのが残差分散である。残差分散は回帰式の当てはまりのよさをみるのに有用である。

★解説

実測値yと回帰による推定値$\hat{y}$との差の平方和

$$(y_1-\hat{y})^2+(y_2-\hat{y}_2)^2+\cdots+(y_n-\hat{y}_n)^2$$

が残差平方和であり，これを $N-p-1$ で割ったのが残差分散である。残差分散が小さければ小さいほど，推定の精度はそれだけよいことになる。これは重回帰モデル

$$y_j=\beta_0+\beta_1 x_{1j}+\cdots+\beta_p x_{pj}+\varepsilon_j \quad (j=1,\ 2,\ \cdots,\ N)$$

において，残差 ε_j の分散の推定値である。

残差 ε_1，ε_2，…，ε_N については，次のことを仮定している。

(1) ε_j と ε_k とは**無相関**である($j\neq k$)。

(2) ε_j の**期待値**は 0 である。

(3) ε_j の分散は等しく一定である。

条件 1 が満たされているかどうかは，データから計算される

$$e_j=y_j-\hat{y}_j$$

の列 e_1，e_2，…，e_N を調べて吟味することになる。必要条件の 1 つとして，**ダービン・ワトソン統計量**がよく知られている。

偏回帰係数βの信頼区間や，仮説検定などを問題にするときは，条件(4) ε_j は正規分布に従う，を入れる。たとえば，求めた重回帰式はyの推測に役立たないという仮説の検定に ***F* 検定**を行うが，F検定になる根拠はこの条件(4)によるのである。

関連ページ

無相関　222
期待値　58
ダービン・ワトソン比　224
F検定　315

日本のドルベース輸出価格指数yの工業製品卸売物価指数x_1と円の対ドルレート指数x_2による重回帰式

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\varepsilon$$

を考える。用いるデータは1972年1期から1980年2期までの四半期データである。

年	期	y	x_1	x_2
72	1	61.8	65.9	96.8
72	2	63.1	66.2	97.7
72	3	64.5	66.7	95.6
72	4	65.7	68.3	98.6
73	1	69.5	71.4	105.2
73	2	75.4	74.0	112.0
73	3	81.2	77.6	112.0
73	4	85.4	84.3	108.0
74	1	91.7	93.8	102.1
74	2	101.9	98.1	106.0
74	3	102.1	99.7	99.7
74	4	102.4	100.5	98.9
75	1	103.2	100.1	101.2
75	2	102.1	99.6	101.5
75	3	99.7	99.8	99.6
75	4	98.1	100.4	97.8
76	1	96.3	102.2	98.2
76	2	97.1	103.8	99.2
76	3	99.8	105.5	102.0
76	4	103.2	106.2	101.1
77	1	103.2	106.4	103.9
77	2	106.7	106.5	107.8
77	3	109.5	106.3	111.5
77	4	114.4	105.8	120.1
78	1	118.2	105.3	124.9
78	2	132.7	105.3	134.4
78	3	140.1	103.9	153.9
78	4	142.3	103.4	155.8
79	1	141.8	105.1	147.3
79	2	140.3	108.4	136.4
79	3	144.7	112.8	135.6
79	4	139.7	116.6	124.4
80	1	143.2	122.5	121.5
80	2	147.3	127.8	127.3

重回帰式は,

$$\hat{y}=-84.269+1.0518x_1+0.77105x_2$$

である。$x_{11}=65.9$, $x_{21}=96.8$ を代入して,

$$\hat{y}_1=-84.269+1.0518\times 65.9+0.77105\times 96.8=59.68$$

を得る。以下同様にして,

$$\hat{y}_2=60.69,\ \cdots\cdots,\ \hat{y}_{34}=148.30$$

を得る。残差平方和は,

$$(61.8-59.68)^2+\cdots+(147.3-148.30)^2=478.3$$

になる。これを $N-p-1=31$ で割ることにより,残差分散3.928^2を得る。

残差εは正規分布に従っているとする。帰無仮説：yのx_1とx_2による重回帰式はyの推測に役立たない,という仮説検定を考える。そのために,回帰による分散

$$\{(\hat{y}_1-\overline{y})^2+\cdots+(\hat{y}_N-\overline{y})^2\}/(p-1)$$
$$=\{(59.68-\overline{y})^2+\cdots+(148.30-\overline{y})^2\}/2$$
$$=11.086$$

より,

$$F=\frac{\text{回帰分散}}{\text{残差分散}}=718.5$$

を得る。これが自由度$(p-1,\ N-p-1)$のF分布に従うことにより検定を行う。いまの場合,自由度(2, 31)の1％点は5.4ほどであり,718.5は5.4に比べ著しく大きな値であり,帰無仮説を棄却することになる。(杉山)

CPM

Critical Path Method の頭文字を並べたことばで，シー・ピー・エムとよんでいる。PERT と同様な手法であるが，それとは独立に開発された。矢線図を作り，クリティカル・パスを求めるという点では，全く変わりがない。ただ，費用の最適化をはかるという点で PERT と異なっている。最近は PERT に含めてよばれることも多いが，PERT 系手法の総称として用いていた文献もある。

★解説

矢線図を書くとき，293 ページでは，仕事に A，B，… のような名前をつけてよんでいたが，ノードに番号をつけているのであるから，矢線図ができ上がってからは，(1, 2) のように，その仕事の始めのノードと終わりのノードとを対(つい)にしてよんでもよい。理論的な扱いをするときにはこの方が便利であるので，このよび方に従っている文献も多い。

▩所要時間と費用 292 ページでは，見積もりが困難であったり，他の要因が加わったりするために，仕事の所要時間が一意に定まらないことがあると注意した。一方，実際には，残業をしたり，人手を増やしたり，あるいは高度の機械力を投入したりすることで，所要時間を早めることのできる場合もある。ふつう，そのときは費用が余分にかかる。しかし，納期を短縮したかったり，予定通りに作業が進んでいなかったりした場合には，余分な費用を払っても作業を早める必要が生ずる。CPM では，右ページの図に示すように所要時間と費用の間の関係が与えられると仮定する。

すなわち，仕事 (i, j) に要する費用は，**標準時間** D_{ij} のときは M_{ij}，**特急時間** d_{ij} のときは m_{ij} であって，その間の所要時間に対しては線形に変化すると考える。特急時間よりも早めることは不可能であると仮定している。

関連ページ
PERT 292

292ページで取り扱った例のクリティカル・パスは89ページで求めてある。この例について，**CPM**の考えを説明しよう。

この例ではクリティカル・パスに要する全時間は6日であった。つまり，この改修工事の工期は6日というわけである。左ページで述べたD_{ij}などの値は次表のようであったとすると，総費用は14万円となる。

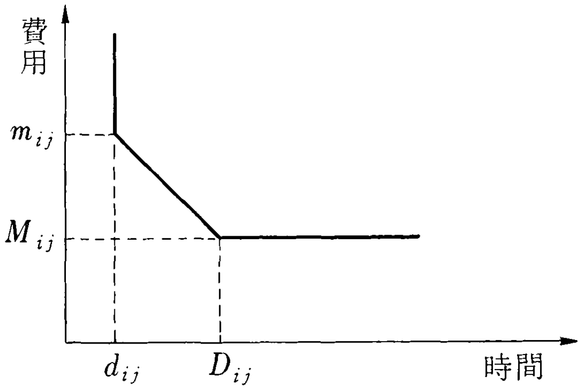
(i, j)	D_{ij}	M_{ij}	d_{ij}	m_{ij}
(1, 2)	1	1	0.5	3
(2, 3)	2	4	1	7
(3, 6)	2	3	1	5
(6, 7)	1	1	0.5	3
(2, 4)	1	2	0.7	4
(4, 5)	1	2	0.5	4
(5, 6)	0.5	1	0.3	2

右上の図で斜めになっている部分の勾配が低いほど，単位時間を短縮するための費用は小さいから，クリティカル・パス上の仕事について，それらを求めてみると，順次

$$\frac{3-1}{1-0.5}=4,\ \frac{7-4}{2-1}=3,\ 2,\ 4$$

となる。したがって(3, 6)の仕事に余分な費用を払って所要時間を縮めるのがよい。この仕事は1日短縮が可能で，その限りにおいては，クリティカル・パスに変更を与えないから，仕事(3, 6)に5万円を払うことにより，工期を1日短縮して，総計5日とすることができる。総費用は16万円となる。

2→4→5→6というパスは総所要時間2.5日で，短縮したクリティカル・パスの所要時間3日より小さいのでクリティカル・パスは1→2→3→6→7で変わっていない。そこで，更に短縮をはかるには，仕事(2, 3)を早めることになる。この場合，仕事(2, 3)自身は1日短縮可能であるが，1日縮めるとクリティカル・パスが1→2→4→5→6→7へ移ってしまうから，0.5日だけ短縮する。つまり，1.5万円を余分に払って，仕事(2, 3)の所要時間を1.5日とするのである。



仕事(2, 4)以下の費用勾配はそれぞれ，6.67, 4, 5と計算される。もし仕事(2, 3)を更に縮めようとすると，仕事(4, 5)も同時に縮めないと，工期短縮に結びつかないので，この方法をとるときの費用勾配は3+4=7となるはずであろう。かくして，仕事(1, 2)または(6, 7)を短縮するのが得策ということになる。

このように順次短縮が可能であるから，必要な工期に合わせて仕事を急がせることを考えればよい。

以上の計算を系統的に行うには，最大流問題の解法を利用した計算方法がいくつか考えられている。（森村）

資金の時間的価値

同じ額の収入や支出でも発生する時期が異なると，利子の影響により経済的価値が相違することになる。収入や支出の発生時期のズレを調整して得られた金額のことを資金の時間的価値という。資金の時間的価値の表し方としては，現在価値(現価)，終価，年金価値(年価)が通常よく用いられる。これらの価を求めるには，資金の時間的価値の換算係数(→ 146 ページ)が使われる。

★解説

現在のP円は，期(年)当たりの利子率iのもとでは，1期(年)後には$P\times(1+i)$円になり，n期(年)後では$P\times(1+i)^n$円となる。このことから，次の資金の時間的価値を求めるための基本となる等価式が得られる。

現在のP円 = n期後の$P\times(1+i)^n$円

ここでの利子率iは，資金の調達に伴う利子率と資金の標準的な運用に伴う利子率(利益率)を意味し，「資本の利率」あるいは「計算利率(caluculating rate)」とよばれている。

一定幅の期間(投資の効果が及ぶ期間，計算上考慮する期間など)の各時点に分布する資金(収入や支出)の時間的価値の表し方としては，次に示す3つの方法が一般に広く用いられている。

(1) **現在価値**(present value)　現在における価値。略して“現価”ともいい，記号Pで表す。

(2) **終価**(final value)　考慮する期間の最終時点での価値。記号Sで表す。

(3) **年金価値**(adjusted mean，または annual worth)　考慮する全期間の毎期末における均等額で表された価値。略して“年価”ともいい，記号Mを用いる(→ 146, 287 ページ)。

関連ページ

方策のキャッシュ・フロー　358
投資案の経済性指標　268
年価法の役立ち　286

収入や支出の発生する時期(タイミング)の違いが経済性を左右することを単純な数値例でみてみよう。

■収入が遅れる損失　現在100万円の収入が得られるはずのものが，1年遅れてしまった。いくらの損失になるだろうか。たとえば，売り値100万円の製品を，販売できずに1年間在庫してしまうことによる損失などは，この種のものになる。

ここでの損失を求めるために，まず比較の対象を明らかにする(→98ページ)。

A：現在の100万円の収入

B：1年後の100万円の収入

AとBでは，収入の発生時期が異なるので直接比較できない。そこで，両者とも1年後の価値で比較してみよう。

A：1年後の$100 \times (1+i)$万円の収入

B：1年後の100万円の収入

したがって，100万円の収入が1年遅れることによる(BのAと比べての)損失は，$i=10\%$ のもとでは，

$$100 \times (1+0.1) - 100 = 100 \times 0.1 = 10 \text{（万円）}$$

になる。

このように考えると，一般に1円の収入がn年遅れることによるn年後での損失額は，年当たりの計算利率iのもとで，

$$(1+i)^n - 1 \text{（円）}$$

となることがわかる。

■一括払いと分割払いの比較　一括払いで現在50万円のものを分割払いにするとまず5万円支払い，残りは5年にわたり，各年度末に10万円ずつ支払うことになる。$i=6\%$ のとき，どちらが安くつくだろうか。

一括払いのために用意した50万円を分割払いに充当してみよう。現在は5万円を支払い，残りの45万円を6％で1年間運用(たとえば預金)して元利合計のなかから10万円を支払い，残りを次の1年間運用する…，としたときの各年度末の手持ち資金額を計算すると次のようになる。

1年後　$45 \times 1.06 - 10 = 37.7$

2年後　$37.7 \times 1.06 - 10 = 29.96$

3年後　$29.96 \times 1.06 - 10 = 21.76$

4年後　$21.76 \times 1.06 - 10 = 13.07$

5年後　$13.07 \times 1.06 - 10 = 3.85$

この結果，5年後に3.85万円手もとに残るので，分割払いが有利になることがわかる。

ここで，「一括払い」の終価S_Aと，「分割払い」の終価S_Bを計算してみよう。

$$S_A = 50 \times (1+0.06)^5 = 66.91$$

$$S_B = 5 \times (1+0.06)^5 + 10 \times (1+0.06)^4 + 10 \times (1+0.06)^3 + 10 \times (1+0.06)^2 + 10 \times (1+0.06) + 10 = 63.06$$

この計算からも，分割払いのほうが3.85万円安いことがわかる。　(中村)

資金の時間的価値の換算係数

資金の時間的価値の現価，終価，年価(→144ページ)の相互を関係づける換算計算のための係数で，つぎの6種類からなる。

(1) 終　価　係　数　　現価(P)から終価(S)への換算係数。

(2) 現　価　係　数　　終価(S)から現価(P)への換算係数。

(3) 資本回収係数　　現価(P)から年価(M)への換算係数。

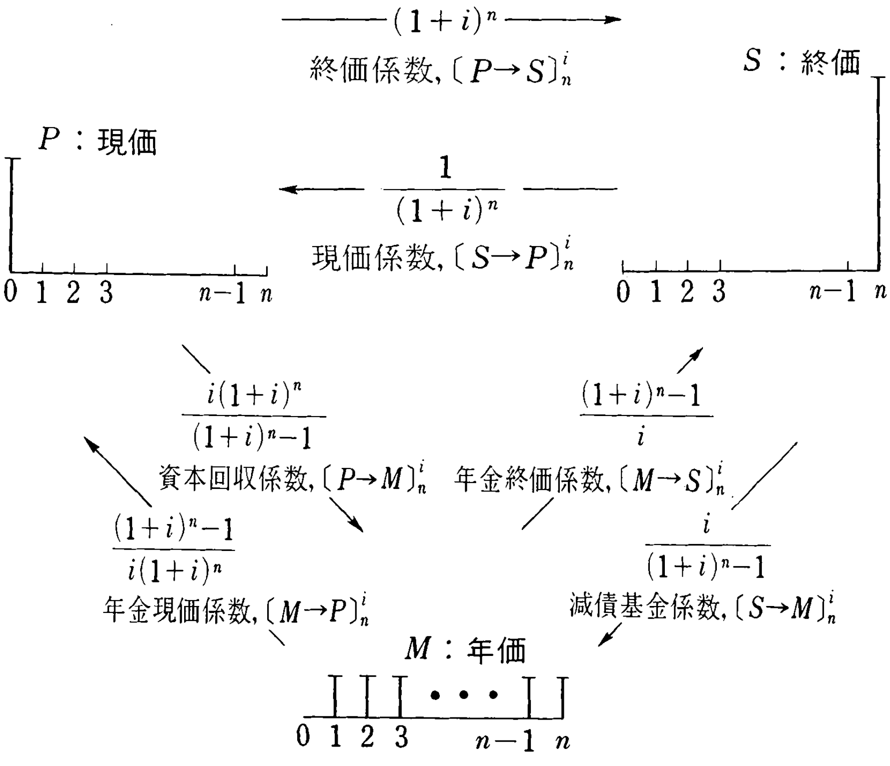
(4) 年金現価係数　　年価(M)から現価(P)への換算係数。

(5) 減債基金係数　　終価(S)から年価(M)への換算係数。

(6) 年金終価係数　　年価(M)から終価(S)への換算係数。

★解説

6種類の換算係数の働き，式，略記号を下図に示す。なお，これらの換算係数の値を知るための数表が用意されている(→438ページ)。



関連ページ

現在価値(現価)　144

終価　144

年金価値(年価)　144

換算係数の使い方を簡単な例で示す。

(1)**終価係数**(final worth factor)

100万円を年利6％の複利で3年間預金すると，3年後の元利合計(S)は次のようになる。

$$S=100\times[P\rightarrow S]_{3}^{6\%}$$
$$=100\times1.191=119.1 \text{（万円）}$$

ある資材のトン当たりの価格は500万円であるが，毎年8％ずつ価格が上昇すると，5年後の価格は次のようになる。

$$(\text{5年後の価格})=500\times[P\rightarrow S]_{5}^{8\%}$$
$$=500\times1.4693$$
$$=734.65 \text{（万円）}$$

(2)**現価係数**(present worth factor)

10年後に1000万円の資金を得るために，年利8％で運用するとき現在準備すべき金額(P)を求める。

$$P=1000\times[S\rightarrow P]_{10}^{8\%}$$
$$=1000\times0.4632=463.2 \text{（万円）}$$

(3)**資本回収係数**(capital recovery factor)

2000万円の設備投資案があり，この設備の使用年数が10年のとき年平均(年末での)の設備費(M)はいくらにつくか。$i=10\%$ とする。

$$M=2000\times[P\rightarrow M]_{10}^{10\%}$$
$$=2000\times0.163=325.6 \text{（万円）}$$

(4)**年金現価係数**(uniform series present worth factor)

工程を機械化すると，毎年末に400万円の経費が節減される。この機械化に必要な初期投資額がいくらまでなら引き合うかを計算する。機械の寿命は8年で，$i=12\%$ とする。

$$M=400\times[M\rightarrow P]_{8}^{12\%}$$
$$=400\times4.968=1987.2 \text{（万円）}$$

(5)**減債基金係数**(sinking fund factor)

7年後に1000万円の社債を償還するために，毎年末に一定額を積み立てていきたい(これを減債基金という)。積立預金の利子率が8％のときの積立額は，次のようにして求められる。

$$M=1000\times[S\rightarrow M]_{7}^{8\%}$$
$$=1000\times0.1121=112.1 \text{（万円）}$$

(6)**年金終価係数**(uniform series final worth factor)

5年後に1億円の設備投資を計画している。そのために5年間にわたり毎年末に1000万円の資金を準備できる。5年後に不足額を借り入れするとすれば，いくら借り入れしなければならないか。$i=10\%$ とする。

5年後の自己資金総額(S)を求めると，次のようになる。

$$S=1000\times[M\rightarrow S]_{5}^{10\%}$$
$$=1000\times6.1051=6105.1 \text{（万円）}$$

したがって，5年後に借り入れしなければならない金額(B)は，次のようになる。

$$B=10000-6105.1=3894.9 \text{（万円）}$$

（中村）

時系列データへ曲線の当てはめ

長期傾向変動を知るために，時間 t の関数 $f(t)$ を当てはめることを行う。当てはめる関数としては，直線，２次曲線，指数曲線，ロジスティック曲線などいろいろと考えられる。当てはめる曲線が理論的に決まる場合は問題ないが，そうでないときは時系列データをグラフにかき，どのような曲線が当てはまるかを検討することになる。

★解説

時系列データに当てはめる曲線の型を決めたあとは，回帰分析の場合と同じく**最小２乗法**により計算する。

(1)　直線　$f(t)=a+bt$

(2)　２次曲線　$f(t)=a+bt+ct^2$

(3)　指数曲線　$f(t)=ab^2$

(4)　ロジスティック曲線　$f(t)=\dfrac{K}{1+ab^t}$

関連ページ
最小２乗法

移動平均法　18
指数平滑法　158
予測　398

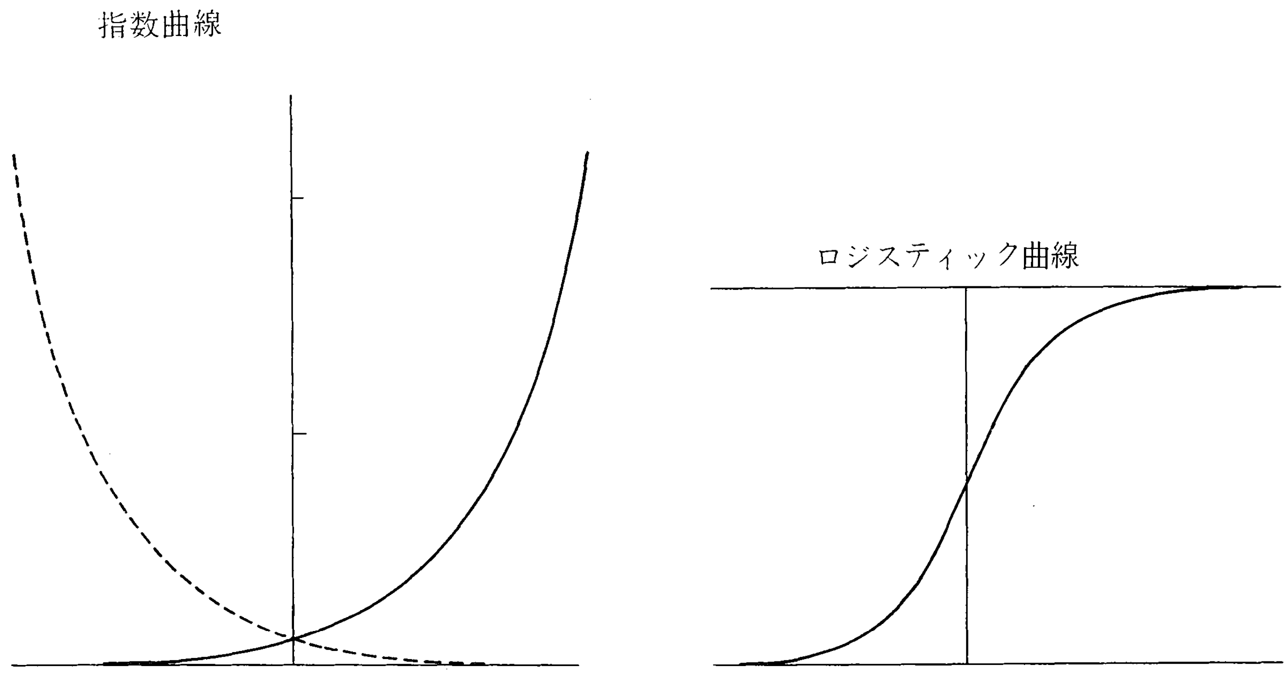


図１

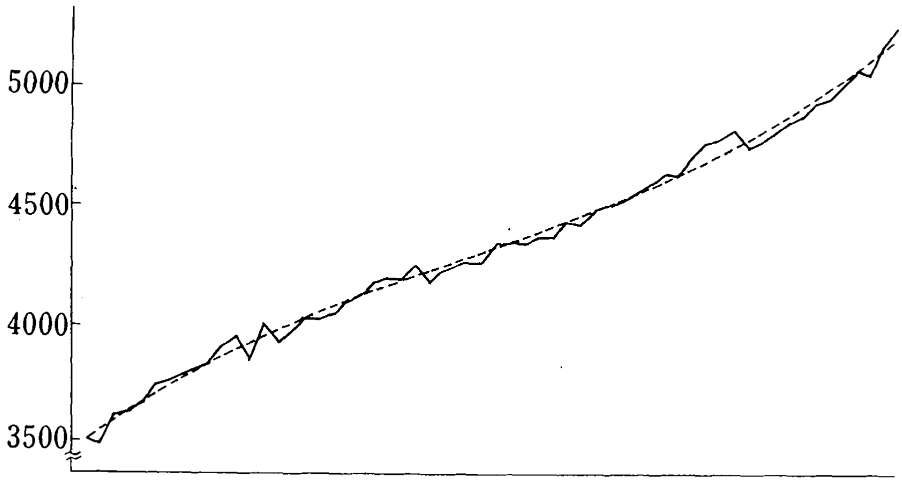


図 2

「移動平均法」のところで用いた百貨店販売額の季節調整済販売額は次のようになる。

月	1975年	1976年	1977年	1978年	1979年
1	3499.	3836.	4243.	4411.	4720.
2	3474.	4001.	4170.	4461.	4751.
3	3610.	3918.	4225.	4499.	4790.
4	3621.	3970.	4254.	4514.	4840.
5	3647.	4025.	4268.	4542.	4868.
6	3740.	4018.	4260.	4584.	4919.
7	3744.	4042.	4336.	4620.	4930.
8	3785.	4097.	4346.	4608.	4991.
9	3816.	4110.	4340.	4686.	5049.
10	3826.	4173.	4365.	4753.	5041.
11	3905.	4192.	4356.	4765.	5164.
12	3947.	4189.	4428.	4800.	5235.

これを図示したのが図 1 の実線である。この系列に 3 次曲線

$$f(t)=a+bt+ct^2+dt^3$$

を最小 2 乗法により当てはめると,

$$f(t)=3513.3+35.305t-0.45581t^2+0.0053406t^3$$

を得る。たとえば, 1975 年 7 月の推定値は,

$$f(7)=3513.3+35.305\times7-0.45581\times7^2+0.0053406\times7^3=3740$$

になる。このときの残差 e_7 は,

$$e_7=3760-3740=20$$

である。$f(1)$, $f(2)$, …, $f(60)$ を計算して図示したのが, 図 1 の点線である。残差系列 e_1, e_2, …, e_{60} にダービン・ワトソン比を計算してみると 2.028 であり, ほぼ 2.0 である。つまり, 残差系列の中にダービン・ワトソン比が見ているような規則性はないことがわかる。

指数曲線 $f(t)=ab^t$ は, 両辺に対数をとり,

$$\log f(t)=\log a+(\log b)t$$

として $\log a=A$, $\log b=B$ とおき,

$$\log f(t)=A+Bt$$

としてから最小 2 乗法を当てはめることになる。指数曲線は, 一定の増加率 b で急速に増加あるいは減少していくという特徴をもっている。また,

$$f(t)=a(1+r)^t$$

は複利計算の t 年後の元利合計である。

(杉山)

自己回帰モデル

時系列データにおいて，時刻 t での値 y_t を，それ以前の過去の値 y_{t-1}，y_{t-2}，… を用いて予測するモデル

$$y_t=\beta_1 y_{t-1}+\beta_2 y_{t-2}+\cdots+\beta_k y_{t-k}+\varepsilon_t$$

を自己回帰モデルという。ここで，ε_t は時刻 t における不規則な変動を表している。

★解説

関連ページ

y_t の期待値は時間に無関係な一定値

$$E(y_t)=\mu \quad (t=0,\ 1,\ 2,\ \cdots)$$

をとり，任意の 2 点での共分散が時間差 t_1-t_2 だけによって定まる時系列 y_t を考えよう。このとき，自己回帰モデルは，

$$y_t-\mu=\beta_1(y_{t-1}-\mu)+\cdots+\beta_{t-k}(y_{t-k}-\mu)+\varepsilon_t$$

とかける。平均 μ が既知であれば，$y_{t-i}-\mu$ を改めて y_{t-i} とおけば，要約の所で記した式になる。これを k 階の自己回帰モデルという。

一般には，平均 μ は時刻 t に依存するので $\mu(t)$ と表し，まず初めにこの平均関数 $\mu(t)$ を推定することになる。$\mu(t)$ の推定法としては，移動平均による方法，多項式による方法などがある。

移動平均法　18

上式の自己回帰モデルにおいて，階数 k はわからない場合が多い。k を決める方法として，ここでは赤池弘次博士によって提案された次の統計量を記す。可能性のある k の範囲〔M，K〕を定めておき，区間

$$M\leqq k\leqq L$$

の間にあるすべての k について，

$$(FPE)_k=\frac{\mu+k+1}{n(n-k-1)}\sum_{t=1}^{n}(y_t-\sum_{i=1}^{k}\hat{\beta}_i(y_{t-i})^2$$

を計算し，$(FPE)_k$ が最小となる k を採用する。

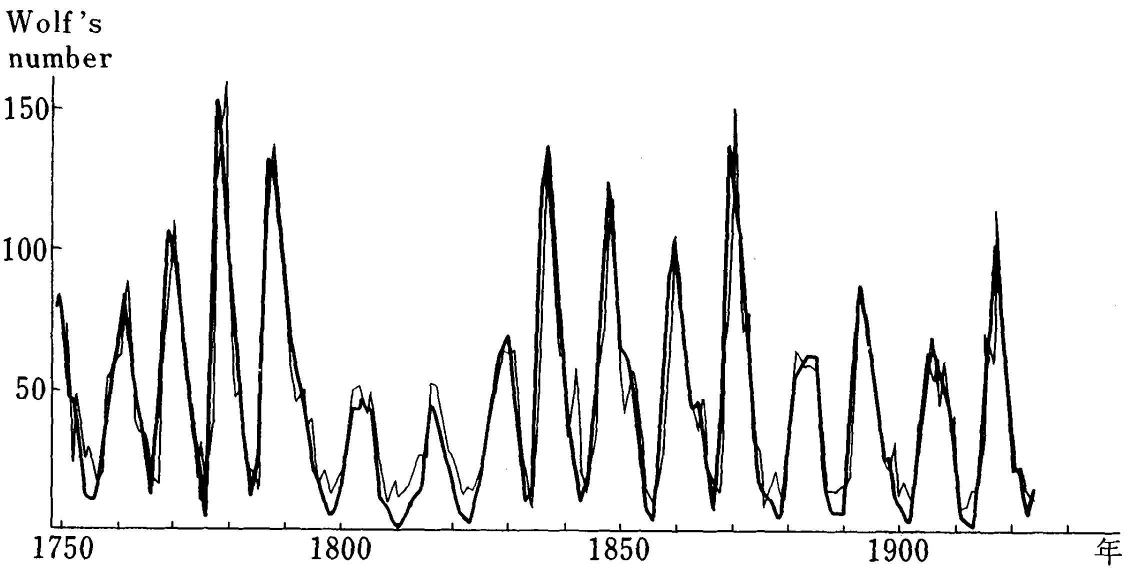


図1　ウォルフの太陽黒点数とその予測値

図 1 の太線は，1749年から1924年までの太陽の黒点の活動を 1 年間にならしたウォルフの黒点数を図示したものである。約 11 年周期での繰り返しが見られ，規則性がありそうである。自己相関係数は，

$r_1=0.811$

$r_2=0.434$

$r_3=0.032$

$r_4=-0.264$

$r_5=-0.404$

である。2 階の自己回帰式を計算すると，

$$\hat{y}_t=1.3425y_{t-1}-0.65504y_{t-2}+13.854$$

である。これを図示したのが図 1 の細線である。

自己回帰モデルの次数 k はいくつにとるのがよいのか，回帰係数の大きさからさぐって見よう。次数 k を決めて，データ ($N=176$) から計算した係数 b_k（最後の項の係数）は次のようになる。

k	b_k	$\sqrt{N}\,b_k$
1	−0.811180	−10.761
2	0.655040	8.690
3	0.101043	1.340
4	−0.013531	−0.180
5	0.050001	0.663

$\sqrt{N}\,b_k$ は $\beta_k=\beta_{k+1}=\cdots\cdots=0$ のとき，大きな N に対して近似的に標準正規分布に従うことが知られている。両側 10 % は 1.64 であり，$k=3, 4, 5$ のとき，この値は小さく，上記の仮説はこの数値からは棄却されない。

実際には，$(FPE)_k$ により次数を決定するのがよいであろう。この方法で信頼できる結果を得るには，藤井光昭（〔45〕参照）は，データ数 N は少なくとも 100 以上，できれば 200 を超えていることが望ましく，500 あるいは 1000 と大きくなれば安定してくることを確かめている。（杉山）

自己相関係数

時系列データ y_1, y_2, …, y_T において，k だけ離れた観測値の組 (y_{k+1}, y_1), (y_{k+2}, y_2), …, (y_T, y_{T-k}) を考える。このとき，y_t と y_{t-k} の相関係数を k 次の自己相関係数という。自己相関係数は時系列の構造を見るのに役立つ。

★解説

k 次の自己相関係数 r_k ($k=1, 2, \cdots$) は，

$$r_k=\frac{\sum_{t=k+1}^{n}(y_t-\overline{y})(y_{t-k}-\overline{y})}{\sum_{t=1}^{n}(y_t-\overline{y})^2}$$

により計算される。q 以上離れた観測値の間に相関がないとき，いいかえれば母相関係数

$$\rho_{q+1}=\rho_{q+2}=\rho_{q+3}=\cdots=0$$

のとき，q より大きい k に対して，

$$r_k \text{ の分散は } \frac{1}{n}\left(1+2\sum_{k=1}^{q}\rho_k^2\right)$$

で近似される。特別な場合として $q=0$，すなわち系列間に相関がない場合は，$k>0$ に対し，

$$r_k \text{ の分散は } \frac{1}{n}$$

で近似される。r_k は，母相関係数 ρ_k の推定量である。

自己相関係数の絶対値が 1 に近いときは，時刻 t の値 y_t は 1 つ手前の値 y_{t-1} により，かなりの確かさで予測できる。このように，自己相関係数は時系列の予測に関して重要な役割を果たす指標といえる。

関連ページ

相関係数　222

年　月		y_t	年　月		y_t
77	1	105.4	78	5	95.2
77	2	106.7	78	6	95.2
77	3	107.8	78	7	92.9
77	4	106.3	78	8	92.5
77	5	105.3	78	9	92.5
77	6	102.8	78	10	93.4
77	7	101.3	78	11	93.7
77	8	99.8	78	12	94.9
77	9	99.7	79	1	97.4
77	10	99.8	79	2	104.7
77	11	97.7	79	3	108.6
77	12	97.0	79	4	112.5
78	1	97.5	79	5	114.5
78	2	97.5	79	6	116.5
78	3	96.4			
78	4	95.0			

表１　非鉄金属卸売物価指数

表１の数値は，1977年１月より1979年６月までの非鉄金属卸売物価指数である。平均 $\overline{y}$ は 100.683 である。

１次の自己相関係数を計算するために次のようにする。

　　x 列　　y 列　　　x 列　y 列

(77年1月, 77年2月) = (105.4, 106.7)

(77年2月, 77年3月) = (106.7, 107,8)

(77年3月, 77年4月) = (107.8, 106.3)

……　……　……　……

(79年5月, 79年6月) = (114.5, 116.5)

この１月ずらしたデータを対にして，x と y の相関係数を計算する。いま，

$$S_x^2=(105,4-\overline{y})^2+\cdots+(114.5-\overline{y})^2$$

$$S_y^2=(106.7-\overline{y})^2+\cdots+(116.5-\overline{y})^2$$

$$S_{xy}=(105.4-\overline{y})(106.7-\overline{y})+\cdots+(114.5-\overline{y})(116.5-\overline{y})$$

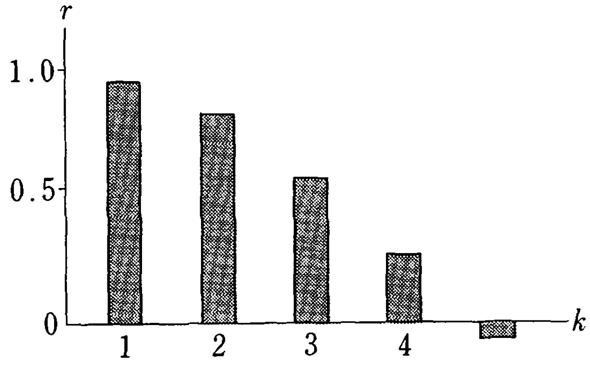


図１

を計算し，これより１次の自己相関係数は，

$$r_1=\frac{S_{xy}}{S_xS_y}=0.951$$

を得る。

２次の自己相関係数は，２カ月ずらした数値を対にしたもの

(77年1月, 77年3月) = (105.4, 107.8)

(77年2月, 77年4月) = (106.7, 106.3)

(77年3月, 77年5月) = (107.8, 105.3)

……　……　……　……

(79年4月, 79年6月) = (112.5, 116.5)

より同様にして計算すると，

$$r_2=0.814$$

を得る。k 次の自己相関係数は k カ月ずらした数値を対にしたものから計算すると，

$$r_3=0.575$$

$$r_4=0.264$$

$$r_5=-0.070$$

を得る。系列の時点を k ずらした値どうしの相関係数は，時系列の構造を推察するのに役立つ。r_1，r_2，r_3，…を図示したものをコレログラム（図１）という。　　（杉山）

指数

2つの統計数値(または2つの統計数値群)の大小を比較する目的で計算された比例数を指数(index)といい，経済現象を解析する手法として使われている。指数は，構造指数と時系列指数に大別され，前者は，地域別賃金水準について東京都を基準(＝100)として指数化し，格差として利用する手法であり，後者は，物価の上昇傾向を特定年(たとえば昭和55年)を基準として，昭和58年は109.7のように統計量の時間的変化を明示する目的で使用されている。

★解説

経済統計では，物価指数，生産指数，輸送活動指数，第3次産業活動指数，稼働率指数，賃金指数，雇用指数，消費水準指数，貿易価格指数，貿易数量指数など40数種類にのぼる指数が主として関係の統計を作成している官庁などによって作成・公表されている。このほか，景気が上向いているか下降しているかを示す**景気動向指数**や，月々の統計数値がすう勢よりどの程度高いか(低いか)を示す季節指数のような特殊な指数もある。経済指数は季節的要因から，月によって上下するため，すう勢をとらえることが難しい。このため，原指数を季節指数で除した季節調整済指数を計算し，すう勢判断に役立てている。また経済分析者の便に供するため，いくつかの指数をうまく組み合わせて，ある特定面の実態を明示する試みもなされている。賃金指数(名目)を消費者物価指数で割った指数(実質賃金指数)を作成することによって賃金の実質的購買力の動向を明らかにするなどである。

指数を理解し，誤りなく活用するには，①指数が明らかにしている経済的側面，②基準時，③採用品目，④総合の方法(総合算式)，⑤総合ウエート付け，の5つについての知識を得ておく必要がある。

関連ページ
主な経済指数　425
景気動向統計　92

指数の作成方法を，消費者物価指数を例として述べると以下のとおりである。

■ねらい　消費者物価指数(Consumer Price Index，略称CPI)は，消費者世帯が日常生活において購入する種々の商品やサービスの価格(あるいは料金)の動きを総合的にとらえて指数化したもので，消費生活に関係する物価動向をとらえることをねらいとしている。

■基準時　指数は昭和21年8月以降，毎月計算されており，昭和30年以降は，5年ごとに基準時の改正が行われ，現在は55年基準となっている。

■範囲　農林漁家世帯及び単身者世帯を除く全世帯が消費する財とサービスのうち，これを代表する品目であり，かつ税金あるいは有価証券，土地購入などの財産購入を除く商品やサービスに限られている。指数を作るには，家計調査(総務庁統計局)結果から得られる消費者の購入品目を代表する512の品目が家計における支出金額が大きく，また類似品目の値動きを代表し，広い地域で実際に購入できるという基準の下に選ばれている。

■作り方　小売物価統計調査(総務庁統計局)によって得られた価格を基礎に，同一品目についての基準時との比例数を計算し，この比例数をつぎにのべるウエートによって加重平均して求めている。ウエートは，家計調査の昭和55年1年間に家計が購入した個々の品目の家計支出金額とその全体に対する割合を基礎に1万分比で決められている。たとえば，総合感冒薬の「新ルルゴールド(30錠入)」は他の総合感冒薬を代表するものとして，総合感冒薬全体のウエート7を与えるという方法である。

ウエートによる総合には，基準時加重相対法算式(ラスパイレス型)が使われており，食料，住居，光熱，被服，雑費などの中間分類項目ごとに結果が算出されている。

■指数の意味　消費者物価指数は，ラスパイレス指数(Laspeyres index)によっているから，その経済的意味は，「基準時ウエートと同じ品目と数量のものを比較時の価格で購入した場合の支出金額を指数化したもの」ということができる。実際には消費者は価格が上昇した品目の購入量を差し控える傾向があるから，この意味での物価動向を示していないといえる。この欠点を補うためには，比較時加重相対法算式(パーシェ型，Paasche)があるが，比較時の品目別価格だけでなく，購入数量をも把握する必要があって調査上の負担となるため，ほとんどの総合指数が，ラスパイレス指数によって作成されている。

なお，生産指数の場合は，生産額ウエートと付加価値額ウエートの2種類がある。　(市野)

指数分布

計算機や空調などの大型システムの故障発生などのように，事象が時間的にランダムに生起する場合，相続く事象間の時間の長さ T の分布は指数分布(exponential distribution)となる。

★解説

このTの分布の密度関数と特性値は次のとおりである。

$f(x)=\lambda e^{-\lambda x}$ ($x\geqq 0$ のとき)：$=0$ ($x<0$ のとき)

平均　$E(T)=1/\lambda$,　　分散　$\mathrm{Var}(T)=1/\lambda^2$

変動係数＝標準偏差／平均＝1.0

信頼度関数　$R(t)=\mathrm{P}\{T\geqq t\}=\int_t^\infty f(x)dx=e^{-\lambda t}$

故障率＝$f(t)/\mathrm{R}(t)=\lambda$

これをパラメータ λ の **(負の) 指数分布** とよぶ。図1に見るように，T の値はほとんどの場合小さめだが，ときたま大きな値も生じることがあり，バラツキが大きい。

指数分布はポアソン分布ときわめて関係が深い。つまり，事象が時間的にランダムに生起する場合，ある時間区間内に生ずる事象の個数はポアソン分布に従うが，その生起時間間隔は指数分布に従っている。

指数分布のいちばんの特徴は，メモリーレスということである。前の事象が生起してから時間 t 待っても次の事象が生じなかったとき，次の瞬間 Δ の間に事象が生ずる確率は，

$$\mathrm{P}\{T\leqq t+\Delta\,|\,T>t\}=\frac{f(t)\Delta}{R(t)}=\frac{\lambda e^{-\lambda t}\cdot\Delta}{e^{-\lambda t}}=\lambda\Delta$$

となり，t の値に関係しない。つまり，故障率が一定ということだが，次の事象がいつ起こるかは，これまで待った長さによらないということである。事象がランダムに生ずることの本質といえよう。

関連ページ

変動係数　348

ポアソン分布 352

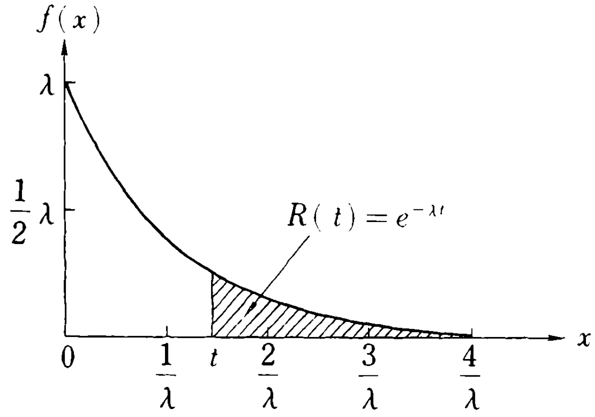


図1　指数分布の密度関数

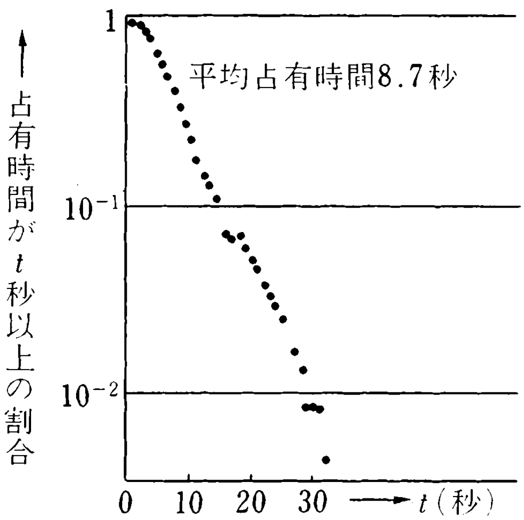


図2　ゲート占有時間の分布(上り方向)

■高速道路料金ゲートでの車の占有時間の分布　指数分布の信頼度関数の対数をとると，

$$\log R(t) = -\lambda t$$

と直線となる。図2は料金ゲートでのサービス時間が t 秒以上となるデータの割合を片対数グラフに記入したものである(〔50〕参照)。ほぼ直線なので，指数分布と考えてよいだろう。

指数分布よりバラツキの大きい分布には**超指数分布**がある。平均の異なった指数分布の混合として表され，計算機の処理時間の分布などによくあてはまる(→413ページ)。

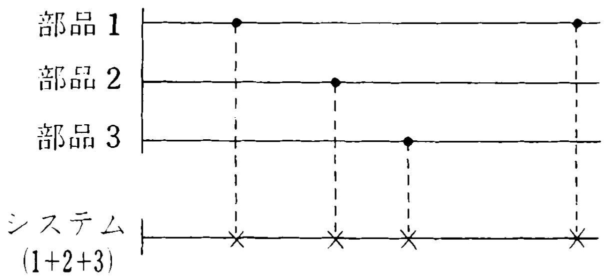


図3　部品とシステムの故障

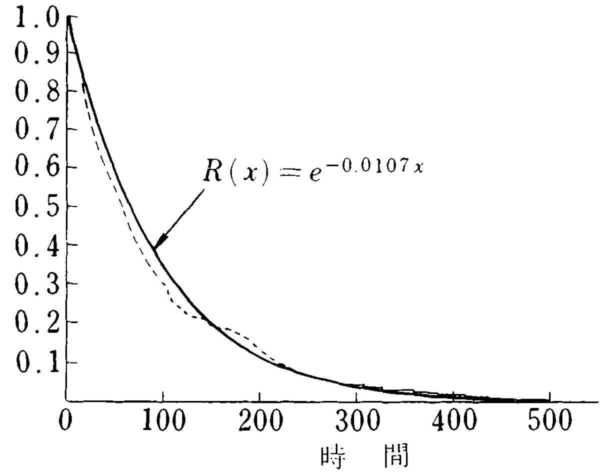


図4　空調の故障分布

■システムの故障　多くの部品からなるシステムの故障時間間隔の分布を考える。部品のどれかが故障すると，システムがダウンする(図3)。つまり，システム・ダウンの発生は，部品の故障発生過程を重ね合わせたものになっている。1つ1つの部品は壊れにくく，その故障発生間隔の分布がまちまちであっても，それが重ね合わされて発生するシステム・ダウンの時間間隔の分布は，部品の数が多くなるにつれ，指数分布に近づくことが知られている。図4はボーイング720型機の空調システムの故障の実測データと指数分布とを比べたものである(〔34〕参照)。適合性を見るには，図2のように対数をとって調べるほうがよいだろう。(森)

指数平滑法

時系列データにおいて，時刻 $t+1$ での観測値 $\hat{y}_{t+1}$ を，

$$\hat{y}_{t+1}=\hat{y}_t+\rho(y_t-\hat{y}_t)$$

によって計算する。ここで，y_t は時刻 t での観測値，$\hat{y}_t$ は当期の予測値，ρ は平滑化指数である。これを指数平滑法という。

★解説

平滑定数 ρ が 1 ならば，上式は $\hat{y}_{t+1}=y_t$ であり，時刻 $t+1$ での予測値 $\hat{y}_{t+1}$ は，時刻 t での観測値 y_t そのもので予測することになる。また，ρ が 0 ならば $\hat{y}_{t-1}=\hat{y}_t$ となり，時刻 t での予測値 $\hat{y}_t$ が，時刻 $t+1$ の予測値そのものになる。一般には平滑化定数は，

$$0<\rho<1$$

であり，この値をどう決めるのかが問題となる。1 つの方法として，観測値 y_1，y_2，…，y_t を用いて予測誤差を求め，その誤差の平方和を小さくするものを ρ として選ぶことが考えられる。

指数平滑法の式

$$\hat{y}_{t+1}=\hat{y}_t+\rho(y_t-\hat{y}_t)$$

に，$\hat{y}_t=\hat{y}_{t-1}+\rho(y_{t-1}-\hat{y}_{t-1})$，$\hat{y}_{t-1}=\hat{y}_{t-2}+\rho(y_{t-2}-\hat{y}_{t-2})$，… を次々に代入していくと，

$$\hat{y}_{t+1}=\rho y_t+\rho(1-\rho)y_{t-1}+\rho(1-\rho)^2 y_{t-2}+\cdots$$
$$+\rho(1-\rho)^{k-1}y_{t-k+1}+(1-\rho)^k\hat{y}_{t-k}$$

を得る。この式からわかるように，指数平滑法は過去のデータに対する加重移動平均を求めている。その点では，自己回帰モデルに共通するところがある。

関連ページ

移動平均法　18

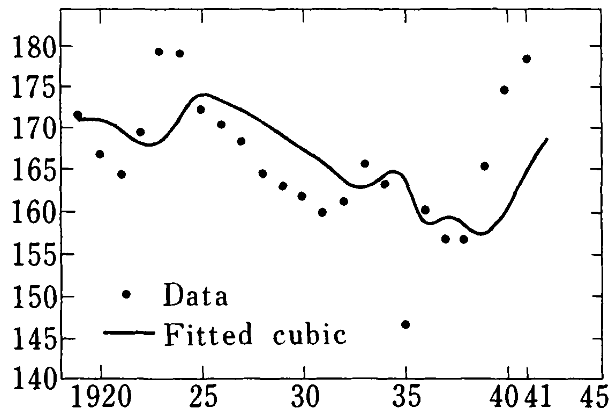


図 1　平滑指数 ρ = 0.3 のときの系列

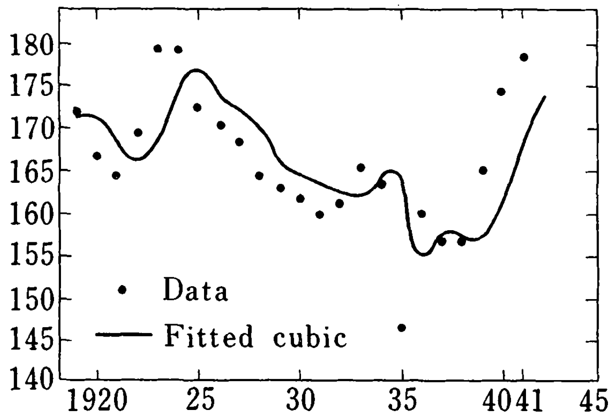


図 2　平滑指数 ρ = 0.5 のときの系列

次の数値は，米国における 1919 年から 1941 年までの 1 人当たり年間の肉消費量(単位：ポンド)である。

年度	t	y_t	年度	t	y_t
1919	1	171.5	1931	13	160.2
1920	2	167.0	1932	14	161.2
1921	3	164.5	1933	15	165.8
1922	4	169.3	1934	16	163.5
1923	5	179.4	1935	17	146.7
1924	6	179.2	1936	18	160.2
1925	7	172.6	1937	19	156.8
1926	8	170.5	1938	20	156.8
1927	9	168.6	1939	21	165.4
1928	10	164.7	1940	22	174.7
1929	11	163.0	1941	23	178.7
1930	12	162.1			

指数平滑法における予測を行ってみよう。平滑化指数を 0.3 とすると，

$$\hat{y}_{t+1} = \hat{y}_t + 0.3(y_t - \hat{y}_t)$$

より，これは

$$\hat{y}_{t+1} = 0.7\hat{y}_t + 0.3y_t$$

とかける。時点 $t+1$ での予測値 $\hat{y}_{t+1}$ は 1 つ前の時点 t での実測値 y_t に 0.3 をかけ，t での予測値 $\hat{y}_t$ に 0.7 をかけて加えたものである。y_t より $\hat{y}_t$ のほうをより重く見ての予測ということになる。y_1 は 171.5 であるが $\hat{y}_1$ の値がない。そこで，$\hat{y}_1$ として最初の数年間の平均に近い 171 を与えることにすると，

$$\hat{y}_2 = 0.7 \times 171 + 0.3 \times 171.5 = 171.15$$
$$\hat{y}_3 = 0.7 \times 171.15 + 0.3 \times 167 = 169.91$$
$$\cdots\cdots$$

を得る。これを図示したのが図 1 である。

指数平滑化指数 ρ を 0.5 とすると，

$$\hat{y}_{t+1} = 0.5\hat{y}_t + 0.5y_t$$

より

$$\hat{y}_2 = 0.5 \times 171 + 0.5 \times 171.5 = 171.25$$
$$\hat{y}_3 = 0.5 \times 171.25 + 0.5 \times 167 = 169.13$$
$$\cdots\cdots$$

を得る。これを図示したのが図 2 である。$\rho=0.3$ の場合より $\rho=0.5$ の場合のほうが原系列の動きに敏感に応答している。平滑化指数 ρ の決め方としては，過去の時系列データから予測誤差を求め，その予測誤差の平方和が最小になるものを選ぶのも 1 つの考え方であろう。（杉山）

実験計画法

数理統計学の手法の１つで，実験結果とそれに影響を及ぼすと考えられる原因系の関係を明確にするため，

(1) 与えられた実験目的に対して，どのような実験を計画するのが，費用，得られる情報量などからみて，最も効果的であるのか，

(2) 得られたデータをどのように解析すればよいのか，

について，それらの考え方，方法を提供するもの。

★解説

実験計画法は，(1)実験の設計と，(2)実験データの解析に関することの２つの側面をもっている。この第１の側面が，他の多くの統計的手法にみられない特長の１つである。

実験を設計するためには，結果を表す特性値(収率や純度など)と，それに影響を及ぼすと考えられる原因系のうちで実験でとりあげるもの(温度，圧力，加工速度などで**因子**という)を明らかにする必要がある。因子については，その水準(とりうる種々の条件)も決める必要がある。

１つの実験でとりあげる因子の数によって，「**１因子実験**」，「２因子実験」などとよぶが，２つ以上の因子を同時にとりあげる実験を多因子実験とよぶ。また，とりあげた因子の水準のあらゆる組み合わせの実験をするのを**要因実験**といい，その一部を行うのを**一部実施法**という。因子の数が多くなると，一部実施法により実験回数を減らす必要が生じるが，これには**直交表**を用いるとよい。

実験データの解析法は，検定と推定に大別することができる。実験データには必ず実験誤差が伴うが，この誤差を考慮してある因子(正確には**要因効果**)が特性値に影響があるか否か検定するのが**分散分析**である。その後に，因子の水準の最適な組み合わせ(最適水準)を求めることになる。これに統計的手法のなかの推定の手法が役に立つ。

関連ページ

因子　20

単一因子実験　232

要因実験　396

乱塊法　402

ラテン方格法　400

直交表　238

無作為化　384

分割法　330

要因効果　394

構造模型　110

分散分析　336

■歴史的背景　実験計画法(design of experiment)は，1920年代に，英国ロザムステット(Rothamsted)農事試験場の技師であったフィッシャー(R. A. Fisher)によって，その基礎が築かれた。当初は施肥効果の比較などの農業試験に主として用いられていたが，第2次世界大戦後はパイロットプラントや工場実験にも広く用いられるようになってきて，工業技術の発展に大きな寄与をしている。

■考え方　実験計画法の考え方が生まれる前には，実験に伴う誤差をできる限り小さくするように精密な実験を行うことが指向されていた。しかし，このような管理された実験の場によって得られた結論を，現実の場に広く適用することは難しい場合が多い。また，農業，工場実験では，実験の場そのものを管理することが難しい。

フィッシャーは，実験データに誤差が伴うことは当然であるという立場に立ち，重要なのは誤差が統計的管理状態にあることを保証することであると考えた。そのため，つぎの3つの重要な原則を提示した。

(1)　反復の原則

(2)　無作為化の原則

(3)　局所管理の原則

反復とは同じ条件で2回以上の実験を行うことで，これによりデータの全変動を実験に伴う誤差による部分とそれ以外の要因による部分に分離することができるので，誤差分散が統計的に推定できる。また，実験順序の無作為化によって，誤差を統計的に偏りのない誤差(偶然誤差)に転化することができる。一方，実験の場をできる限り均一な部分(ブロックとよぶ)に区切って実験の精度をあげるのが局所管理である。

■方法　実験計画法を，上に述べた3つの原則の適用の相違により分類する。

実験全体を無作為化して行う実験を，完全無作為化法という。この方法による1因子実験，2因子実験などは，それぞれ，1元配置法，2元配置法などとよばれている。

局所管理のためのブロック因子を導入した計画を乱塊法という。実験の場をいくつかのブロックに分割し，各ブロック内を局所管理し，処理(因子と水準)のわりつけは各ブロックごとに無作為に行う。ブロック因子を2つとり入れた計画をラテン方格法という。

各ブロックに比較したい処理の一揃いすべてが実験されるような計画を完備型計画，ブロックの大きさに制限があって比較したい処理の一部しか配置できない計画を不完備型計画という。

無作為化が何段階かにわけて行われる実験を分割法といい，実際の場でよく用いられる。また，ブロック因子と，重要でない交互作用を交絡させる実験方法を交絡法という。　　(宮村)

質的データの解析

データの解析を通して，分析対象についての情報を引き出す作業をデータ解析という。データ解析の手法としては，いわゆる多変量解析とよばれる量的データを扱う手法が中心である。しかし，社会科学に代表される文科系の諸科学や医学分野などでは，いわゆる質的データ(名義尺度または分類尺度)が多い。これらの質的データを解析する手法が数量化理論とよばれる手法群である。林の数量化理論，属性相関係数，多重分割表，多次元尺度構成法，双対尺度法などがある。

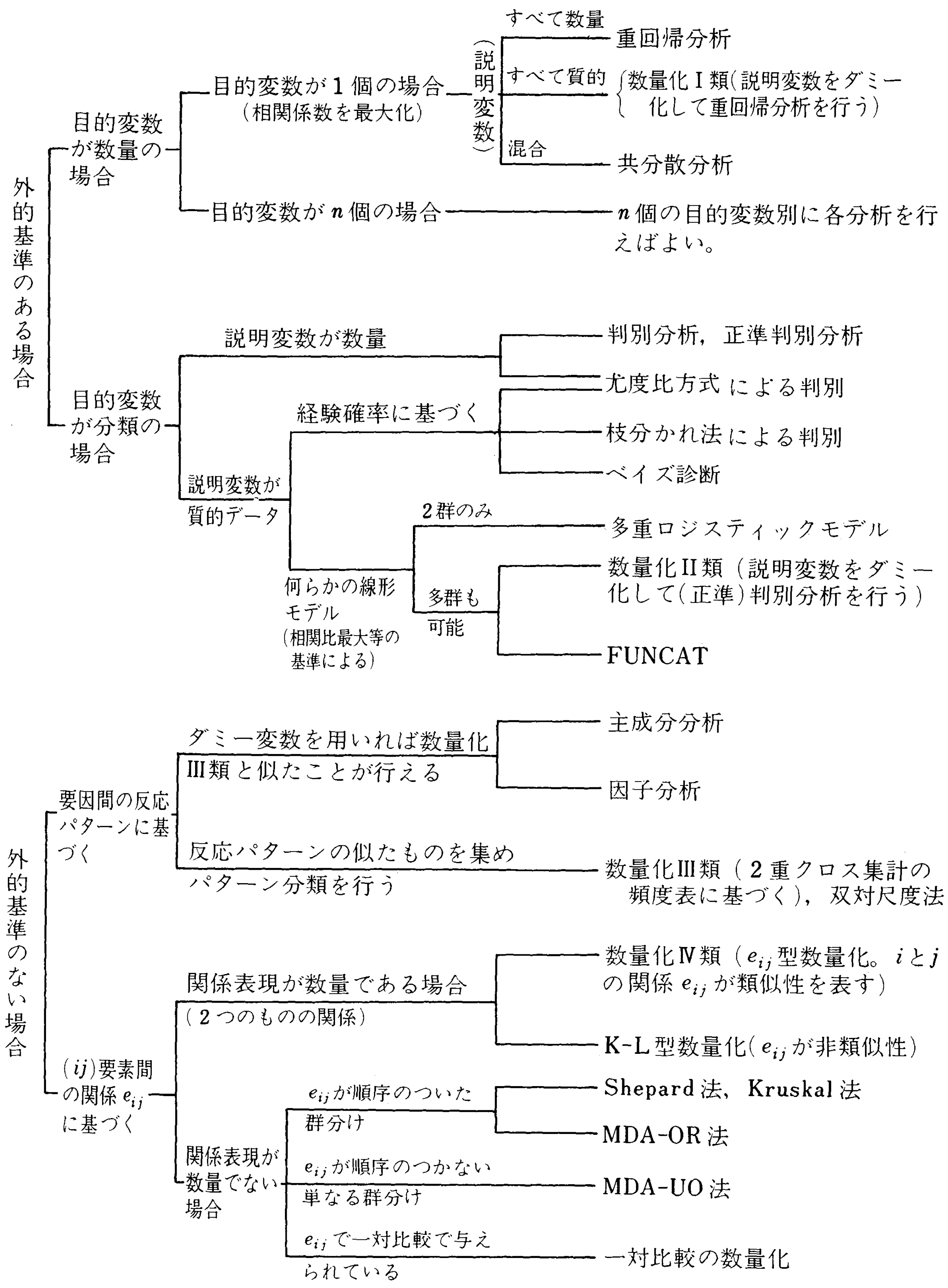
数量化理論は量的データの解析手法の拡張であることが多いから，数量化を学習する前に，少なくとも重回帰分析や主成分分析などの基礎知識を得ておくことが望ましい。

★解説

■質的データ　品種，銘柄，実験計画における因子水準，質問表に対する回答などは，AであるかBであるかといった形のもので，このようなデータを質的データという。これをカテゴリカル・データとよぶこともある。また，品質の等級，疾病の重症度，嗜好の順位などの「順序関係」のあるものを順序尺度という。この尺度は，質的データとして扱う場合もあるし，何らかの数量を与えて間隔尺度として扱うこともある。

■手法　**林の数量化理論**を中心に，データ解析手法を右ページに示す。判別分析手法では他にも多くの手法が提案されている。数量化Ⅰ類とⅡ類は，ダミー変数の導入または**デザイン行列**を学ぶことにより従来の重回帰分析の拡張(一般線形回帰モデル)と考えてもよい。データ解析の心がまえとしては，一手法の適用に終始するのではなく，クロス集計による変数の検討に始まり，数量化Ⅲ類による変数とサンプルのパターン分類を経た後，予測や判別手法を適用すべきである。

関連ページ
林の数量化理論　298
数量化Ⅰ類　188
数量化Ⅱ類　190
数量化Ⅲ類　194
数量化Ⅳ類　196
デザイン行列　250
クロス集計　90
FUNCATによる判別　320
2重クロス表の独立性と属性相関係数　278
ベイズの定理を用いた判別　346



(注) 数量化に際しては，事前準備としてクロス集計によるデータチェック，再カテゴリー化，2項目間の独立性（χ^2 検定）が重要である。

（新村）

図1

資本回収係数と許容投資額

資本回収係数(→164ページ)を投資額に乗じると，その投資から回収すべき年当たりの最低必要額が求まる。また，投資から得られる年当たりの効果額を資本回収係数で割る(年金現価係数を乗じる)と，許容投資額が求まる。

★解説

資本回収係数は，利子を考慮したもとでの現価(たとえば投資額)の年平均値を求めるのに使われる係数である。資本回収係数は次式で表される。

$$資本回収係数=[P \rightarrow M]_n^i=\frac{i(1+i)^n}{(1+i)^n-1}$$

たとえば，現在の投資額の100万円を，$i=10\%$ のもとで5年間にわたり平均すると，

$$100 \times [P \rightarrow M]_5^{10\%}=100 \times 0.2638=263.8 \ (万円)$$

になる。利子率が0のもとでは資本回収係数は $1/n$ になる。上記の数値例では，利子を考慮しなければ年当たりの値は200万円となり，63.8万円は利子ぶんをさしていることになる。また，十分に長い年数，たとえば50年から100年ぐらい，のもとでは，資本回収係数の値は i そのものになることがわかる。

投資の経済性の計算では，投資による必要回収額を求めるとき，あるいは年当たりの効果額から投資の限度額，すなわち許容投資額を求めるときなどに資本回収係数は便利な道具として使われる。

必要回収額＝投資額×資本回収係数

許容投資額＝年当たり効果額÷資本回収係数

関連ページ
資金の時間的価値の換算係数　146
年金価値　144
投資案の経済性指標　268

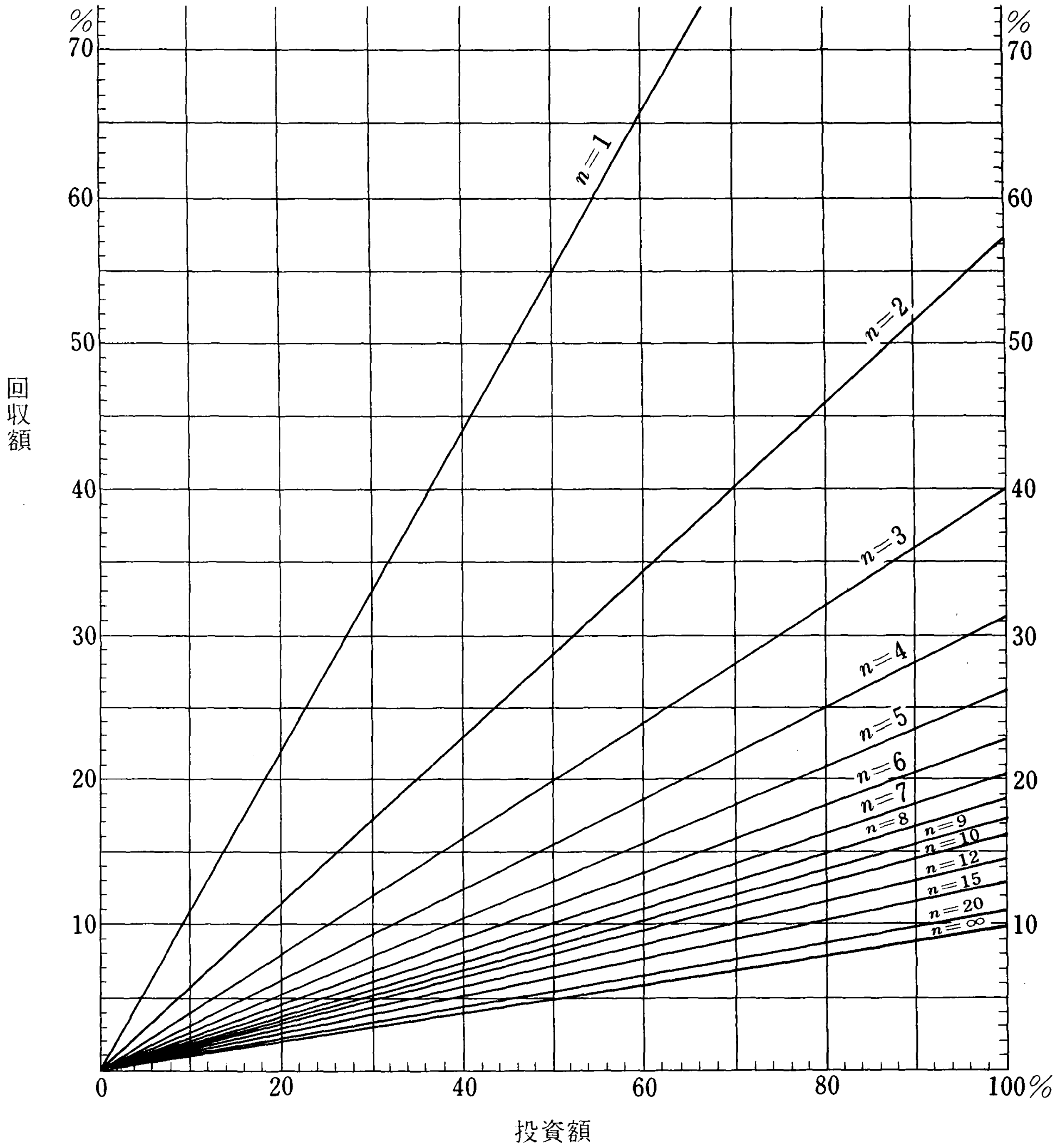


図1　投資・回収図表（$i=10\%$）

投資額がわかっていて回収額を求めたいとき，あるいは年当たりの回収額(効果額)から(許容)投資額を知りたいとき，上に示す図表を用いると便利である(この図表は $i=10\%$ のときのものになっている)。

(例 1)　投資額が 3000 万円のとき横軸の 100 % を 3000 万円にとると，$n=5$ 年では 3000×26％＝720 万円の必要回収額になることが読みとれる。

(例 2)　年当たり 300 万円の効果があがるとき，3 年で回収するならば許容投資額はおよそ 750 万円になることが読みとれる。　(中村)

シミュレーション (Simulation)

模倣するというラテン語の simulo からきた言葉といわれ，広義には「まねをすること」，狭義には「数値実験」を意味する。複雑な現実のシステムを，簡単なシミュレーション・モデル(数式モデル・論理モデル)に置き換えてコンピュータなどで数値解を求める場合や，実物または相似模型等の物理モデルを用いたシミュレーションなどがある。前者で乱数を利用するものはモンテカルロ法とよばれる。

★解説

■シミュレーション・モデルの分類　シミュレーションに用いられるモデルは次のように大別される。

- 論理モデル(ソフトモデル)
  - 数式モデル(計量経済モデルなどの連立方程式系や微分方程式で表されるモデル)
  - 手続きモデル(複雑なシステムの挙動はアルゴリズムや流れ図で記述され，シミュレーション言語で実行できる)
- 物理モデル(ハードモデル)
  - 実物モデル(自動車や航空機の運転練習)
  - スケールモデル(風洞実験などの小型モデル)
  - アナログ・モデル(機械系の振動現象や経済活動などを同じ特性をもつ電気回路で模倣し，アナログ・コンピュータでシミュレートする)

■シミュレーションの長所　原子炉の安全解析や天気予報にみる通り，実物を使うより①安全，②経済的，③再現性，④現実にはあまり起こらないケースの検討，⑤実際にかかる時間を縮める，など。また，数学解析的に解けない問題の数値的な答えを出すためにも多用される。

関連ページ
モンテカルロ法　388
シミュレーション言語　168
統計シミュレーション　256
コンピュータ・ハードウェア　126
コンピュータ・ソフトウェア　124

■コンピュータ・シミュレーション

論理モデルはデジタル・コンピュータで，アナログ･モデルはアナログ･コンピュータでシミュレーションできる。これらを混合したハイブリッド・コンピュータも従来使われていたが，計算精度と汎用性から，今日ではデジタル・コンピュータが主として用いられる。このため,デジタル･コンピュータのシミュレーション言語にも,アナログ･コンピュータの機能を模したものがある。特にアナログ･シミュレーションで特定問題向きの専用機をシミュレータとよぶが,シミュレーション言語によるデジタル･シミュレーションをシミュレータとよぶこともある。

■シミュレーションの分類　シミュレーションは，対象とするシステムの状態変化が**離散的**か**連続的**かで大別できる。

離散型のシステムには,待ち行列,ジョブショップ,在庫,スケジューリングなどの多くのOR問題がある。代表的な言語としては，SIMSCRIPT, GPSS，SIMULAなどがある。

連続型のシステムには，工学上の多くの問題(機械系の振動現象，電気回路の過渡現象，制御系)や，生物の増殖，マクロな社会現象などがある。従来はアナログ・シミュレーションが多かったが，今日では連続型のシミュレーション言語を用いたデジタル・シミュレーションが多くなってきている。代表的な言語としては，CSMP系言語のMIDAS，CSSL系言語のCSSLやSL-1，DYNAMOなどがある。

また，シミュレーションの実行形態からも分類できる。すなわち，単位時間ごとに時間を進める時間進行方式か，次の状態変化が起きた時点まで時刻を進める事象—事象進行方式かである。

変化の記述方法も，事象を中心としたものと，トランザクションまたはプロセス中心のものとで分類され，更にシミュレーションモデルが**確率的**か**確定的**かで分けることもできる。モンテカルロ法は確率モデルのシミュレーションを扱うが，確定的な数値計算も扱える。また，有限要素法を用いた環境シミュレーションも広く行われており，その対象は広い。

■シミュレーションの効率化　シミュレーションは数値実験であるから，多数回試行することによりシステムの特性を知ることができる。シミュレーションモデルは，各種要因を入力としてシステム特性を出力するブラックボックスと考えられる。この要因の組み合わせを少なくする方法として，実験計画の直交表やガロア体が用いられる。また，相関関係の利用も有効である。

（新村）

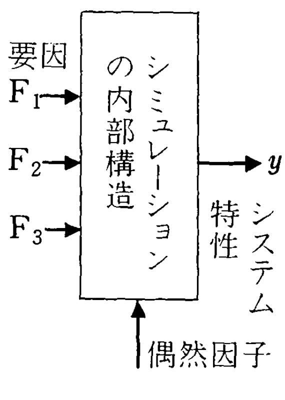


図 1

シミュレーション言語

Fortran, PL/1 などのプログラム言語は，シミュレーション・モデルの記述を簡単に行うことができない。そこで，シミュレーションに特有なモデル記述や演算命令を備えたシミュレーション専用のシミュレーション言語が開発されてきた。GPSS に代表される離散システムを扱うタイプと，DYNAMO に代表される連続システム用とに大別できる。

★解説

▩シミュレーション言語に必要な機能 ①シミュレーションの進行と時間の進行の管理。②モデル構成概念の提供とそれを具体化する論理演算命令等の提供。③任意の乱数の発生。④結果の統計処理やグラフ出力。

(1) 離散システム・シミュレーション言語

GPSS 右ページ参照。初心者向きで，企業などで最も利用されている。

SIMSCRIPT(Simulation Scriptor) M. マルコビッツにより開発された。複雑なシミュレーションに用いることができ，専門家向きの言語といわれる。

SIMULA(Simulation Language) ノルウェーで開発されたシミュレーション言語で，SIMSCRIPT に似ている。

その他 SOL，SIMPL/1 など。

(2) 連続システム・シミュレーション言語

DYNAMO 右ページ参照。

CSMP(Continuous System Modeling Program) アナログ・コンピュータの積分器，加算器，除算器などの演算機能をもち，常微分方程式で表されるシステムを扱う。

(3) 専用シミュレーション言語

上で述べた汎用型に対し，化学工業，回路設計，気象予測等の多くの分野で，Fortran 言語などを用いた専用の言語が開発されている。

関連ページ

シミュレーション 166

モンテカルロ法 388

統計シミュレーション 256

汎用プログラミング言語 308

汎用統計パッケージ 306

コンピュータ・ソフトウェア 124

■ **GPSS** (**General Purpose Simulation System**) R. エフロンと G. ゴードンにより開発され，最初はコンピュータ・システムのバッファやチャネルの能力を決定するためのシミュレーション言語であった。このため，現在でも特に待ち行列に向く言語として評価されている。

待ち行列は，サービス窓口への客の到着→待ち行列の形成→サービスの開始→サービスの終了，という過程をとる。

たとえば，次の例を GPSS でシミュレーションすることを考える。ある銀行の 1 台のキャッシュ・ディスペンサーには，5 秒から 15 秒間隔で客が 1 名到着する。待ち行列に並び，サービスを受ける。サービス時間は 5 秒から 9 秒間隔のランダム性をもつ。GPSS によるブロック・ダイヤグラムを左に，プログラムをその右に示す。

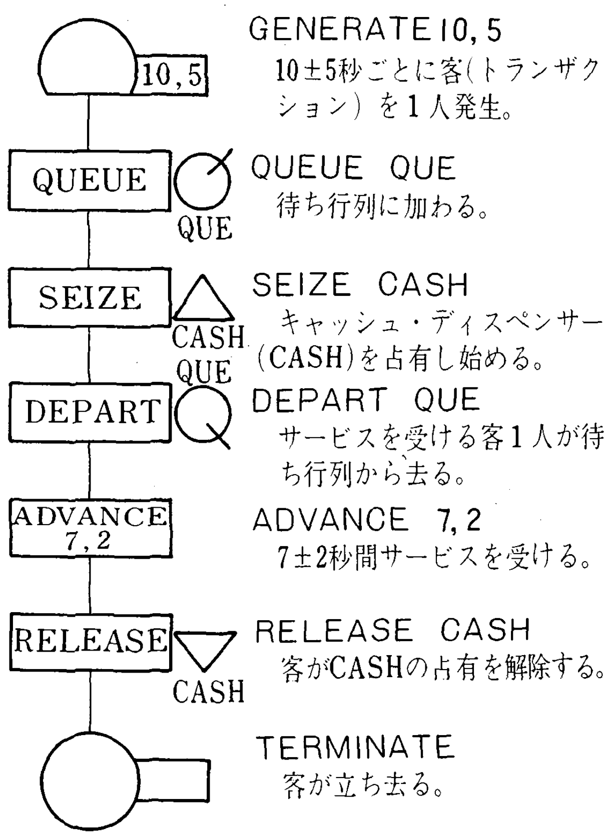


図1 GPSSのブロック図

■ **DYNAMO** (**DYNAmic MOdel**)

MIT の電気工学者フォレスターにより開発された。企業経済活動を自動制御理論のフィードバック系と考えるインダストリアル・ダイナミックス (ID) なる理論に基づく連続システム用の言語である。

ID の考えるシステムは，正と負のフィードバック・ループが多重ループをなし，遅れや増幅のあるシステムを扱う。ここでは，意思決定が次の時点における意思決定に影響を及ぼすことを考えている。この他，指数遅れや非線形性も容易に扱える。

その後，ダイナモの扱う対象は，企業レベル(ID)から都市レベル(アーバン・ダイナミックス，UD)そして世界モデル(ワールド・ダイナミックス，WD)へと発展拡大してきた。このため，これらを総称して今日ではシステム・ダイナミックス(SD)とよばれる。

SD の世界では，①フローとよばれる物・注文・金・人・資本設備・情報などの流れと，②在庫量や受注残などのストックを示すレベルが，③入出荷量や注文量などのレイトにより制御される。 (新村)

重回帰分析

ある特性yを，yに関する情報をもった変数x_1，x_2，…，x_pによって予測あるいは推測したい。このとき，推測式として，

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_px_p+\varepsilon$$

を考える。これは重回帰式とよばれる。目的変数yと説明変数xとの間に，このような関係式が成り立つという事前知識をもっていることが大切である。偏回帰係数βはデータから計算し推定する。

★解説

最小2乗法を用いてデータからβ_0，β_1，…，β_pの推定値を求める。

1変数の場合，xからyを推測するモデルとして，

$$y=\beta_0+\beta_1x+\beta_2x^2+\cdots+\beta_px^p+\varepsilon$$

を考える。このモデルは，

$$x_1=x,\ x_2=x^2,\ \cdots,\ x_p=x^p$$

とおくことにより重回帰モデルになる。積のモデル

$$y'=\beta_0z_1{}^{\beta_1}z_2{}^{\beta_2}\cdots z_p{}^{\beta_p}e^{\varepsilon}$$

の場合，両辺に対数をとり，

$$y=\log y',\ x_1=\log z_1,\ \cdots,\ x_p=\log z_p$$

とおけば重回帰モデルに帰着する。最終的にモデルが線型式で表されれば，重回帰モデルとして扱えることがわかる。

重回帰式のあてはまりのよさ，つまり推測の精度は，**残差分散**の値を参考にする。また，**重相関係数**も1つの目安になる。

関連ページ

残差分散　140
重相関係数　174

全世帯実質消費支出yを，勤労者世帯実質実収入x_1と消費者物価指数x_2で説明する重回帰式

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\varepsilon$$

を考える。データは以下に示す1973年1期から1980年2期までの4半期データである。

年	期	全世帯実質消費支出(y)	勤労者世帯実質実収入(x_1)	消費者物価指数(x_2)
73	1	145000	180863	67.37
73	2	147645	217986	70.83
73	3	151429	213447	72.97
73	4	177244	300704	76.27
74	1	138277	168474	83.83
74	2	146817	222351	87.77
74	3	151903	224118	91.07
74	4	169281	297106	95.03
75	1	147033	185243	96.93
75	2	152051	230523	99.57
75	3	155318	229307	100.47
75	4	176644	296469	103.13
76	1	150457	190179	105.53
76	2	153228	225644	108.87
76	3	156517	223379	110.07
76	4	178330	301346	112.80
77	1	152348	195265	115.33
77	2	156666	232919	118.43
77	3	158261	231472	118.80
77	4	177820	308160	119.77
78	1	155783	201158	120.30
78	2	157651	242106	122.73
78	3	160735	235893	123.57
78	4	183678	313651	123.87
79	1	160185	208742	123.47
79	2	164394	250563	126.60
79	3	165840	244316	127.87
79	4	185120	319923	130.03
80	1	162600	206462	132.77
80	2	161216	246369	137.13

第i番目のデータを$(y_i,\ x_{1i},\ x_{2i})$で表すことにすると，

$$\sum_{i=1}^{n}\{y_i-(\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i})\}^2$$

を最小にするβ_0，β_1，β_2を求めることになる。この最小2乗推定値は，

$$\hat{\beta}_0=\overline{y}-\hat{\beta}_1\overline{x}_1-\hat{\beta}_2\overline{x}_2$$

$$\hat{\beta}_1=S^{11}S_{1y}+S^{12}S_{2y}$$

$$\hat{\beta}_2=S^{12}S_{1y}+S^{22}S_{2y}$$

になる。

S_{iy}は変数x_iとyの積和，S_{ij}は変数x_iとx_jとの積和であり，S^{ij}は行列(S_{ij})の逆行列のij要素である(〔27〕参照)。データから重回帰式を計算すると，

$$\overline{y}=86573.6+0.24667x_1+136.52x_2$$

を得る。

1973年1期の説明変数x_{11}とx_{21}は，

$$x_{11}=180863.27 \qquad x_{21}=67.37$$

であるから，回帰による推測値は，

$$\hat{y}=86573.6+0.24667\times180863.27+136.52\times67.37=140384.31$$

を得る。1973年2期の値x_{12}とx_{22}を回帰式に代入して$\hat{y}_2$を，以下同様にして$\hat{y}_3$，$\hat{y}_4$，…，$\hat{y}_{30}$を得る。

説明変数が3以上の場合，計算式は複雑になるが，考え方はまったく同じである。　(杉山)

重回帰分析の変数選択

重回帰式の中に取り入れられる変数はたくさんあることが多い。その中のどのような変数の組み合わせを用いて回帰分析を行うかを決めることが最初の問題となる。理論的な考察により取り入れられる変数が決まってしまう場合は別であるが，そうでない場合は信頼性の高い回帰式を得るという観点から，変数選択は重要な問題である。変数選択としては，変数増減法と変数減増法がよく用いられる。

★解説

関連ページ
相関係数　222

目的変数 y と説明変数 x_1 との**相関係数**，x_2 との相関係数…を計算する。相関係数の一番大きい x_i が，y に関する情報を最も多くもっていることになる。x_i が選ばれたという条件のもとで，次に y と大きな相関をもった x_l を x_i と組み合わせて，(x_i, x_l) によって回帰式を作るのが推測の精度を最もよくするかというと，必ずしもそうではない。x_i と x_l が強く相関していれば，x_i の y に関してもっている情報と，x_l の y に関してもっている情報とは，かなりの部分で重複していることになる。このようなことから，y と x との相関の大きさだけでなく，実際には説明変数 x_1，x_2，…，x_p の相関関係の強さまで考慮に入れて，回帰式に用いる変数の組を選ぶことになる。コンピュータのパッケージプログラムとしてよく組み込まれているのは，変数増減法と変数減増法である。

変数増減法は，初めに y との相関が最大の説明変数を1つ選ぶ。次にそれと残りのどの変数との組み合わせが推測の精度を最もよくするか，さらに2つが選択された状況のもとで次にどれを追加すべきか，新たに変数が追加されたときこれまでに選択した変数はもはや有用なものでなく除去すべきか否かを判定しながら，説明変数の追加と除去とを繰り返し行う。追加と除去は1つずつ行うのが規則である。

ある国立大学某学部の昭和58年度受験者1753名の入学試験成績がある。二次試験の成績をyで，共通一次の得点をx_1(国語)，x_2(社会)，x_3(数学)，x_4(理科)，x_5(英語)で表す。このとき，二次試験の成績yが，共通一次の成績x_1，x_2，x_3，x_4，x_5のどのような組み合わせによって説明されるかを見るために，重回帰分析を行う。それぞれの平均点と標準偏差，相関行列は次のようになる。

	平均点	標準偏差
x_1(国語)	61	10.9
x_2(社会)	57	10.1
x_3(数学)	74	15.7
x_4(理科)	65	10.4
x_5(英語)	49	12.8
y(二次)	82	4.7

相関行列

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ y \end{array}
\begin{array}{c} \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\
\begin{bmatrix} 1 & & & & & \\ .16 & 1 & & & & \\ .11 & .09 & 1 & & & \\ .20 & .21 & .20 & 1 & & \\ .25 & .12 & .29 & .19 & 1 & \\ .18 & .12 & .28 & .28 & .28 & 1 \end{bmatrix} \end{array}$$

目的変数yと説明変数x_1，x_2，x_3，x_4，x_5とのそれぞれの相関係数は，0.18，0.12，0.28，0.28，0.28と小さく，目的変数と説明変数の相関は弱い。このことから高い説明力をもった重回帰式を得ることは期待できないことが推察できよう。

変数選択において，変数を取り入れる基準値F_{IN}と変数を取り除く基準値F_{OUT}は，データ数が多いときは2.0が妥当であることから，2.0に設定して計算を行った。変数選択の各段階での結果を以下に要約する。

第1段階　x_4(理科)を取り入れる
第2段階　x_5(英語)を取り入れる
第3段階　x_3(数学)を取り入れる
第4段階　x_1(国語)を取り入れる

この段階でのF値は，
取り入れた変数　x_1：6.4
x_2：31.8　x_4：39.5　x_5：26.7
取り入れていない変数　x_2：0.85
である。変数x_1のF値は0.85で，変数を取り入れる基準値$F_{IN}=2.0$より小さいので，この段階で計算は終了する。重回帰式は，

$$\hat{y}=65.1+0.035x_1+0.054x_3 +0.090x_4+0.062x_5$$

であり，重相関係数は0.41である。

残差の標準偏差は4.3であり，つぎのF値(→141ページ)を得る。

$$F=\frac{\text{回帰分散}}{\text{残差分散}}=46.0$$

F分布の自由度(4, 1748)のF分布の1％点は2.37ほどであるから，46.0はこれに比べて著しく大きい。

(杉山)

重相関係数・寄与率

実測値と回帰推定値との相関係数を重相関係数という。重相関係数の2乗を寄与率あるいは決定係数という。データ数が説明変数の個数に近い，つまり標本の数が少ないときは，重相関係数がほとんど1に近い値になることがある。このときは見かけ上のあてはまりのよさを差し引いた修正済重相関係数の値をみることになる。

★解説

実測値 y の平均からの偏差の平方和が全変動

$$(y_1-\overline{y})^2+(y_2-\overline{y})^2+\cdots+(y_N-\overline{y})^2$$

である。回帰による推定値 $\hat{y}$ の平均からの偏差平方和

$$(\hat{y}_1-\overline{y})^2+(\hat{y}_2-\overline{y})^2+\cdots+(\hat{y}_N-\overline{y})^2$$

が回帰による平方和である。y と $\hat{y}$ との差の平方和が残差平方和であるが，これらの間には，

全変動＝回帰平方和＋残差平方和

という関係が成り立つ。上式の両辺を全変動で割ると，

$$\frac{\text{回帰平方和}}{\text{全変動}}+\frac{\text{残差平方和}}{\text{全変動}}=1$$

とかける。上式の第1項が重相関係数の平方，すなわち寄与率である。全変動の中に占める回帰による変動の割合が寄与率である。回帰からの残差の変動が小さければ小さいほど第2項は小さくなり，相対的に第1項の寄与率は大きくなる。逆に寄与率が1に近ければ近いほど，残差平方和は小さい。これは全変動に占める残差平方和が小さいということであり，第2項の値(言いかえれば第1項の寄与率の値)から，残差平方和の値を直接知ることはできないことに注意してほしい。

関連ページ
重回帰分析　170

重回帰分析のところで用いた，全世帯実質消費支出yの勤労世帯実質実収入x_1と消費者物価指数x_2による重回帰式

$$\hat{y}=86573.6+0.24667x_1+136.52x_2$$

で説明する。観測値y_iの平均$\overline{y}$と，回帰による推測値$\hat{y}_i$の平均$\overline{\hat{y}_i}$とは同じ値になる。全変動は，

$$y_1=145000.27 \cdots\cdots y_{30}=161216.46$$

より

$$(145000.27-\overline{y})^2+\cdots+(161216.46-\overline{y})^2=42286\times10^5$$

になる。回帰による平方和は，

$$\hat{y}_1=140384.31 \cdots\cdots \hat{y}_{30}=166067.49$$

より，

$$(140384.31-\overline{y})^2+\cdots+(166067.49-\overline{y})^2=38577\times10^5$$

になる。これから寄与率

$$R^2=\frac{\text{回帰平方和}}{\text{全変動}}=\frac{38577\times10^5}{42286\times10^5}=0.9123$$

を得る。

横軸を観測値y，縦軸を推測値$\hat{y}$として，点$(y_i,\ \hat{y}_i)$，$(i=1,\ \cdots,\ 30)$をプロットしたのが図1である。これらの点が45°の直線のまわりに集まっていればいるほど，回帰による当たりはまりがよいことが推察できよう。R

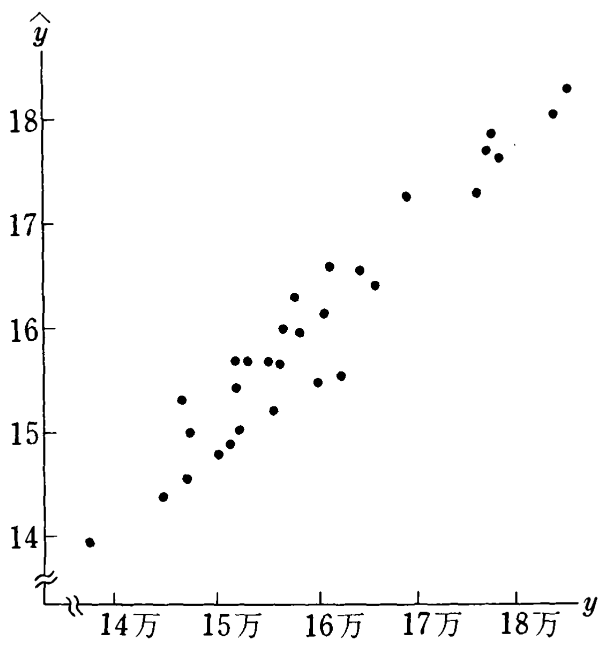


図1　観測値yと推測値$\hat{y}$の相関図

はyと$\hat{y}$との相関係数であり，

$$R=\frac{(y_1-\overline{y})(\hat{y}_1-\overline{y})+\cdots+(y_N-\overline{y})(\hat{y}_N-\overline{y})}{\sqrt{y\text{の平方和}}\sqrt{\hat{y}\text{の平方和}}}$$

によっても求められる。

説明変数の個数Pが標本数Nに近いときは，

$$R'^2=\frac{N-1}{N-P-1}R^2-\frac{P}{N-P-1}$$

により計算される修正済寄与率R'^2の値を見ることになる。この平方根R'を修正済重相関係数という。全世帯実質消費支出の例では$N-R-1=27$であり，

$$R'^2=0.9058$$

になる。$R^2=0.9123$であり，意味のあるほどの差はない。　(杉山)

主成分分析

多変量 x_1, x_2, …, x_p の内容を最もよく説明している1次式

$$y=h_1x_1+h_2x_2+\cdots+h_px_p$$

をデータから求める。これが第1主成分 y_1 であり，x のあらゆる1次式の中で分散が最大である。第2主成分 y_2 は y_1 と関連性がないという条件のもとで，分散が最大となる x の1次式である。以下同様にして，y_m の分散は y_1, y_2, …, y_{m-1} のすべてと関連性がない1次式の中で最大である。p 変数であれば p 個の主成分を得る。

★解説

主成分分析のモデルは，主成分が x_1, x_2, …, x_p の1次式で表されることである。主成分の分散を**固有値**(あるいは固有根)という。主成分を決めている係数(h_1, h_2, …, h_p)を**固有ベクトル**という。

各変数の分散と共分散は，10変数であれば55個，20変数であれば210個と，変数が増せば増すほど多くなり，それらの数値の意味を同時に考察することは容易でない。そこでデータのもっている多次元的特性をなるべく損なわないように要約した，いいかえれば各変量の散らばりの大きさと，変量間の関連性がもっている情報を縮約した主成分に基づいて考察を進めることになる。2つあるいは3つほどの主成分に要約できれば，データを散布図などに可視的に表現したりできるので，考察は容易になる。

主成分分析で得られた新たな変量(主成分)が，どのような意味をもっているかは，固有ベクトルの数値の大きさから推察する。

関連ページ
固有値　178
固有ベクトル　180

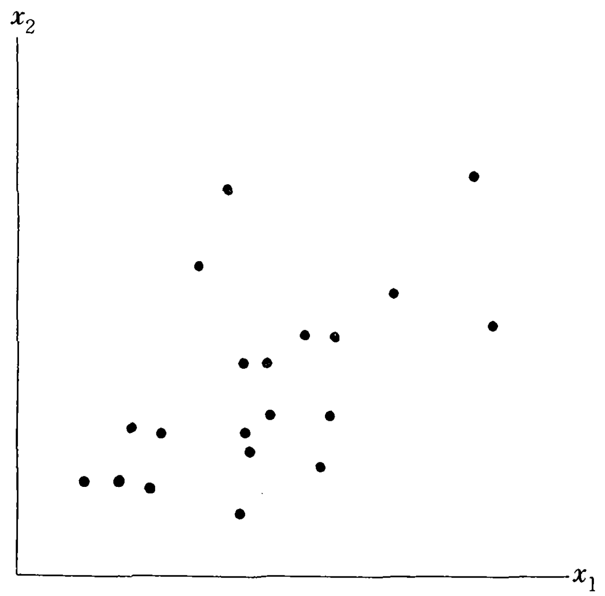


図 1　2 変量のデータの散布図

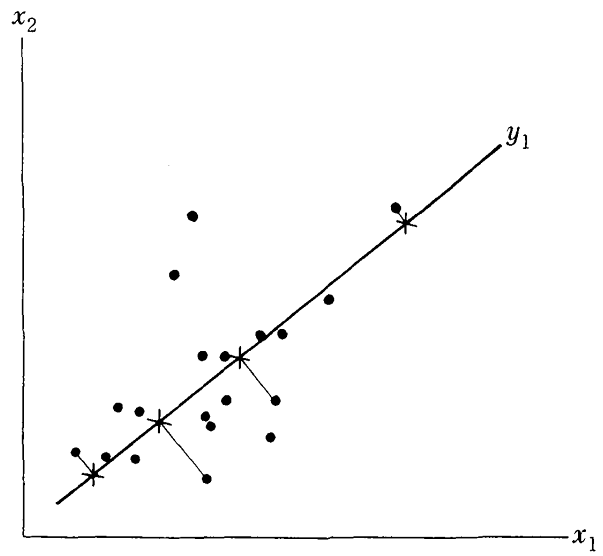


図 2　第 1 主成分 y_1 の図

まず初めに 2 変量で考える。図 1 は標本を 2 次元空間にプロットしたものである。標本の散らばりの有り様を最もよく説明する直線(1 次元空間)を求める。その直線は平均を通り，標本を直線上に写した点(図 2 ではその一部分を×点で示してある)の分散が最大のものである。これが第 1 主成分 y_1 である。

中学 3 年生 137 人の 2 学期の成績を次のように表す。

	1	2	3	・・・	137
国語 x_1	58	23	44	・・・	60
社会 x_2	92	60	80	・・・	64
数学 x_3	85	0	42	・・・	41
理科 x_4	71	38	60	・・・	37
音楽 x_5	52	36	29	・・・	48
美術 x_6	80	53	30	・・・	35
体育 x_7	85	12	40	・・・	68
技家 x_8	47	30	41	・・・	43
英語 x_9	86	17	62	・・・	41

9 変量で考えると，標本は 9 次元空間の 137 個の点として表すことができる。主成分分析の目的は，137 個の標本点の散布の有り様ができる限り保たれるように，9 より低い m 次元空間に標本点を写すことである。9 変量を例にしてのべているが，p 変量と書きかえても同じである。多次元でのデータの散らばりを視覚的にとらえられるように，より低次元の空間へのデータの配置換えを行うことになる。

考える対象とする空間の次元を減らすことによって，もとのデータのもっている情報量は多少失われることになるが，その際に失う情報の損失をできる限り少なくするという条件のもとで，低次元化を行うことになる。

この成績データの場合，主成分分析を行うと，第 1 主成分 y_1 が 76 %の情報をもっている。y_1, y_2 の 2 次元では約83%の情報を抽出している。(杉山)

主成分分析における固有値

第 1 主成分の分散が固有値 l_1 である。第 2 主成分の分散を l_2，第 3 主成分の分散を l_3 で表すと，固有値に関しては $l_1 \geqq l_2 \geqq \cdots \geqq l_p > 0$ という関係がある。第 1 主成分がどの程度多くの情報を把握しているかは，l_1 の大きさからわかる。固有値 l_1 が大きければ大きいほど，情報をたくさん有している重要な主成分といえる。第 2 主成分の固有値 l_2，第 3 主成分の固有値 l_3，…についても同じである。

★解説

固有値はデータから計算した**分散共分散行列**あるいは**相関行列**の固有値を求める方法で得られる。

固有値 l_1，l_2，…，l_p の合計は，もとの変数の分散の合計に等しい。これを総分散という。主成分分析では，情報量という言葉を分散の大きさの意味で用いる。p 変量のもっている情報量の合計を 100 としたときに，第 1 主成分は何パーセントの情報を把握しているかを見るために，

$$\frac{l_1}{\text{総分散}} \times 100\ (\%)$$

を計算する。これは第 1 主成分の寄与率とよばれる。第 2 主成分の寄与率は，上式で l_1 のかわりに l_2 を代入すれば求まる。以下同様である。固有値の大きさの順に第 m 主成分 $(m \leqq p)$ までとったときの寄与率

$$\frac{l_1 + l_2 + \cdots + l_p}{\text{総分散}} \times 100\ (\%)$$

を累積寄与率という。これらは m 個の主成分を選ぶという条件のもとでは，情報量の最も多いものを選んだことになっている。つまり，累積寄与率が最大になっているのである。

関連ページ
分散共分散行列　334
相関行列　334

中学3年生137人について，9教科，国語，社会，数学，理科，音楽，美術，体育，技家，英語の成績得点に基づき，相関行列を計算する(→334ページ)。相関行列を出発点として主成分分析を行ったところ，固有値は次のようであった。

第1主成分の固有値　6.810
第2主成分の固有値　.586
第3主成分の固有値　.470
第4主成分の固有値　.302
第5主成分の固有値　.263

相関行列から主成分分析を行っているので，総分散は相関行列の対角要素の和

$$1+1+1+\cdots+1=9$$

になる。これより第1主成分の寄与率は，

$$\frac{6.810}{9}\times 100=75.8\ (\%)$$

を得る。寄与率を100をかけないで，0.758と表すこともある。以下同様に，第2主成分の寄与率6.5％，第3主成分の寄与率0.052％…を得る。

第2主成分までの累積寄与率は，

$$\frac{6.810+0.586}{9}\times 100=82.3\ (\%)$$

になる。同様に第3主成分までの累積寄与率は，

$$\frac{6.810+0.586\times 0.470}{9}\times 100$$

$$=87.4\ (\%)$$

を得る。第1主成分と第2主成分で，

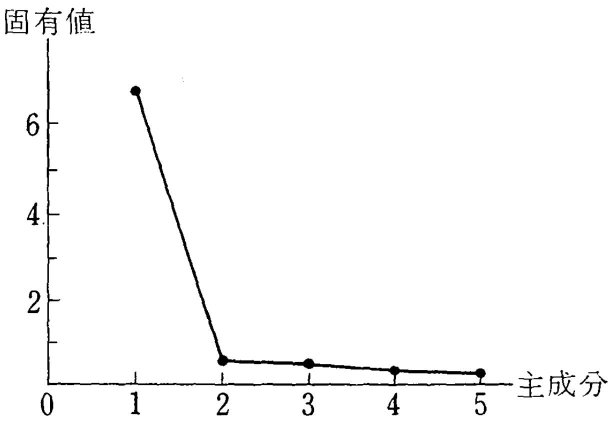


図1

約82％の情報を抽出したことになっている。残りの情報

$$100-82=18\ (\%)$$

は，第3主成分以下の7つの主成分がもっていることになる。もとの9変数で考えるかわりに，情報の縮約されている2つの主成分で考えると扱いやすく好都合だが，18％ほどの情報の損失をその代償として支払うことになる。

何パーセントになるまで主成分をとればよいのか，という質問を受けることがある。60％でよい，あるいは80％でなければいけないなどという基準はない。それぞれの固有値の大きさと，主成分の意味あいを考えながら，結果を活用する人が決めることになる。

寄与率という用語は，重回帰分析でも用いる。主成分分析の寄与率と，重回帰分析の寄与率とは，定義そのものからしてまったく異なるものであるが，この例のように定義が同じでない場合が統計学ではときどきあるから，気をつけてほしい。　(杉山)

主成分分析における固有ベクトル

主成分 $y=h_1x_1+h_2x_2+\cdots+h_px_p$ の意味づけは，x の係数である固有ベクトル$(h_1,\ h_2,\ \cdots,\ h_p)$の数値の大きさと，正負の符号から推し測る。$h_i$ の絶対値が大きければ大きいほど，主成分 y を x_i はよく説明していると解釈する。大きな係数をもった x の変数の組み合わせから，主成分 y がどのような因子であるかをさぐり，意味づけを行う。

★解説

第 j 主成分を

$$y_j=h_{j1}x_1+h_{j2}x_2+\cdots+h_{jp}x_p$$

とかくと，固有ベクトル$(h_{j1},\ h_{j2},\ \cdots,\ h_{jp})$は**固有値** l_j に対応する固有ベクトルとよばれる。係数については，

$$h_{j1}^2+h_{j2}^2+\cdots+h_{jp}^2=1$$

になっている。主成分の係数の大小は，2 乗したもので比較することが多い。また，固有ベクトルのかわりに，**因子負荷量**で表現することがよくある。

変数 x_1，x_2，…，x_p の計測単位が，kg で測られていたり，cm で測られていたりすれば，それらの加法和である主成分は，単位の異なるものを加え合わせたものになり，意味がない。そのような状況で主成分分析を適用したいときは，**無単位になる操作**

$$\frac{x_i-\text{平均}}{\text{標準偏差}}$$

を行う。その意味で，主成分分析は**分散共分散行列**から分析を始める場合と，無単位とした**相関行列**から始める場合とがある。

関連ページ
固有値 178
因子負荷量 24
分散共分散行列 334
相関行列 334

中学3年生137人の9教科の成績に基づく相関行列から主成分分析を行い，固有ベクトルを求めると次のようになる。

主成分	y_1	y_2	y_3
国語 x_1	.350	.097	.065
社会 x_2	.355	.131	.024
数学 x_3	.337	.074	.281
理科 x_4	.359	.056	.161
音楽 x_5	.339	.057	.183
美術 x_6	.292	.322	−.845
体育 x_7	.263	−.923	−.214
技家 x_8	.340	−.053	−.127
英語 x_9	.352	.062	.285

相関行列から主成分分析を行ったということは，各変数に

$$\frac{x_i-\text{平均}}{\text{標準偏差}}$$

という標準化を行って，各教科の分散を等しくしたことを意味している。標準化は解説のところで説明したように，計測単位が異なるときに行う場合と，この例のように各変数の分散を同じにそろえたりする場合に用いられる。いずれの場合も，相関行列を分析の出発点とするのは同じである。

各教科の平均は x_1 から順に，45.9，66.0，46.0，55.2，51.6，50.9，54.3，52.1，56.0であり，標準偏差も同じく18.5，23.1，29.3，26.1，24.7，22.6，21.5，24.9，27.8である。これより，第1主成分は，

$$\begin{aligned}y_1=\,&0.350\left(\frac{x_1-45.9}{18.5}\right)+0.355\left(\frac{x_2-66.0}{23.1}\right)\\&+0.337\left(\frac{x_3-46.0}{29.3}\right)+0.359\left(\frac{x_4-55.2}{26.1}\right)\\&+0.339\left(\frac{x_5-51.6}{24.7}\right)+0.292\left(\frac{x_6-50.9}{22.6}\right)\\&+0.263\left(\frac{x_7-54.3}{21.5}\right)+0.340\left(\frac{x_8-52.1}{24.9}\right)\\&+0.352\left(\frac{x_9-56.0}{27.8}\right)\end{aligned}$$

とかける。それぞれの教科の係数は，すべてが正で，美術，体育が多少小さめであるが，その他は0.34前後の値でほぼ同じである。この係数の符号と大きさから判断して，第1主成分 y_1 は総合点の因子と考えられる。寄与率は0.758（75.8％）であり，y_1 は約4分の3の情報を抽出している重要な因子である。第2主成分 y_2 は体育の係数が −0.923 と他に比べて著しく大きく，y_2 は体育の因子である。第3主成分 y_3 は美術の係数 −.845 と大きく，y_3 は美術の因子である。y_2，y_3 は寄与率は6.5％，5.2％と小さく，1教科のもっている情報量

$$\frac{100}{9}\fallingdotseq 11\ (\%)$$

の約半分にすぎない。これらは，主成分で求めている総合特性値とは考えられない。この場合，総合特性値として意味のあるのは第1主成分 y_1 だけである。　　（杉山）

新聞売り子の問題

新聞スタンドで毎日売る新聞を何部仕入れるのが最適か，という形の問題。売れ残りの新聞が多いと損失を招き，仕入れが少な過ぎると儲け損なう。毎日の売れ行きも確率的に変動するから，期待利益を最大にするという形で最適な仕入れ数を求める。調達期間0で，毎期首の在庫量を決定する問題とも解釈できるが，ひとり在庫管理に限らず，生鮮食品の生産計画などこの形の問題とみなされる例は多い。

★解説

関連ページ
定期発注方式
244

■新聞に関する費用　新聞は1部 C_1 円で仕入れ，C_2 円で売るものとする。売れ残った新聞は1部 C_3 円で売り払われる。もちろん，$C_2>C_1>C_3\geqq 0$ という大小関係が成り立っている。

■新聞の需要分布　毎日の新聞の売れ行きは，ある確率分布に従って変動するものとする。その確率分布は，k 部売れる確率を P_k $(k>0)$ として表される。もちろん，P_k をすべて加えたものは1に等しい(確率の基本性質)。

■仕入れ量の決定　ある日の新聞の仕入れ量を x 部，販売需要を y 部とすると，仕入れに要する費用は C_1x 円である。一方，売り上げは，C_2y 円であり，売れ残りの売却で得る金額は $C_3(x-y)$ 円である。もちろん，$x\leqq y$ のとき，つまり，需要が仕入れ量を上回っていれば，売り切れとなって，せっかく客が買いに来ても断らなければならない。このように儲けの機会を失うことを**機会損失**という。結局 x 部仕入れたときの期待利得 $B(x)$ は，

$$B(x)=C_2\sum_{y=0}^{x} yp_y+C_2x\sum_{y=x+1}^{\infty} p_y+C_3\sum_{y=0}^{x}(x-y)p_y-C_1x$$

という式で与えられ，次式をみたす x で最大値をとる。

$$p_1+p_2+p_{x-1}\leqq\frac{C_2-C_1}{C_2-C_3}<p_1+p_2+\cdots+p_x \qquad (1)$$

■**例題**　右の表1のように，新聞の需要が与えられているとする。また，仕入れ額 C_1 は45円，販売額 C_2 は80円，売れ残り売却額 C_3 は10円としよう。

k	P_k	k	P_k
5	.01	35	.16
10	.02	40	.20
15	.03	45	.22
20	.04	50	.12
25	.06	55	.03
30	.10	60	.01

表1

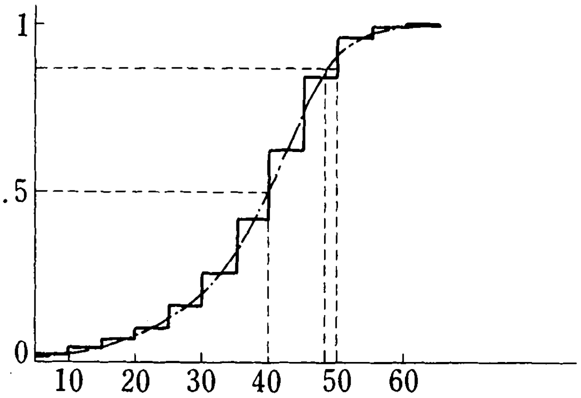


図1

式(1)の不等号の間にはさまれた量は $(80-45)/(80-10)=0.5$ となり，これを $p_1+\cdots+p_x$ の値と比べると，$x=35$ までの和は0.42，$x=40$ までの和は0.62であるので，40部の仕入れが最適ということになる。上の図1を利用すると，C_1，C_2，C_3 の値が変わっても，容易に最適仕入れ部数を求めることができる。

■**連続化**　図1は階段関数として描かれているが，本来5部ずつ売れるものではないから，連続化して滑らかな曲線で表した方がかえってすっきりしているであろう。このような場合，式(1)はもっと簡単になり，図1に示す連続な分布関数と $(C_2-C_1)/(C_2-C_3)$ の値の交点の横座標が最適仕入れ数を与えることになる。たとえば，図1の鎖線のように連続化したとき，最適仕入れ部数はやはり40部である。

■**機会損失の見積もり**　いままでの考え方には，機会損失は暗々裡にはいっていたに過ぎない。つまり，せっかく客が来たのに在庫がなくて売り損なうときの損失は，1部あたり C_2-C_1 円と見ているだけで，それに伴う信用損失までは見積もっていない。しかし，例題のように，40部仕入れたとすると，45部以上の需要に対しては売り切れとなるわけで，その割合は38％にも達する。こんなにも多くの客を断っていると，客の信用を失ってだんだん客は別のスタンドに足を向けるようになるであろう。こういった心配のある場合には，信用損失をもっと多く見積もっておく必要がある。

売り切れで客を断るときの損失を1回 C_4 円と見積もると，式(1)の C_2 の代わりに C_2+C_4 を用いればよい。もし，$C_4=200$ 円と見積もったときには $(C_2+C_4-C_1)/(C_2+C_4-C_3)$ の値は，87になるので，最適仕入部数は50部(連続化したときは48部)となり，売れ残りもかなり出る代わりに売り切れはほとんど生じなくなる。

おいしいパンや菓子など，売り切れを多くしてかえって信用を高める場合は $C_4=0$ とし，新聞などでは C_4 を大きくするのが実際的であろう。(森村)

信頼性

信頼性とはJISによれば,「アイテムが与えられた条件で規定の期間中,要求された機能を果たすことができる性質」である。

技術が進歩し,われわれの使用するシステム(航空機,自動車およびプラントシステムなど)が複雑になると,これらのシステムには故障が少なく,長く使用に耐え,しかも修理しやすいという性質が重視される。このような性質を信頼性という。

★解説

信頼性はその内容を大別すると,耐久性,保全性および設計信頼性の3つによって説明される。

耐久性とは長く寿命が続き,かつ,故障の少ない性質をいう。したがって,耐久性が高いことは,**平均故障寿命**(MTTFと略称する)が大きく,故障率が小さいことである。この故障率の逆数を**MTBF**という。

保全性とは故障したシステムが修理しやすく,または,修理時間の短いことを示した性質をいう。システムやコンポーネントの耐久性を増やすことは,システムなどの製品コストが増加し,かつ,重量などが大きくなるために,不経済となることがある。このために,故障がシステムにとって致命的とはならないときには,コストなどを低くするために耐久性を上げる代わりに保全性を高めることが多い。つまり,故障しやすくても直しやすくする方法をとるのである(図1)。

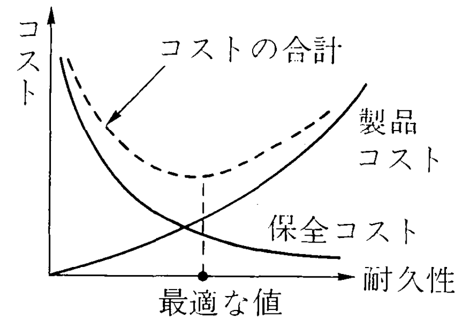


図1　耐久性と保全性

また,設計のときに,冗長系や人間機械系を考慮して,系の信頼度を上げたり,システムを使いやすく,誤動作や誤操作のないように**設計信頼性**を上げることも重要である。

関連ページ

MTTF　40
平均故障寿命　40
故障率　156
MTBF　40

■信頼性のおこり　1940年代の前半に，当時，第二次世界大戦のさなかにおいて新兵器レーダーシステムを開発することに成功した米国は，このシステムによって戦いを優位に進めたが，反面このシステムの故障には頭を悩まし続けたといわれる。このために，故障対策の科学的管理手法として信頼性の考え方が生まれた。また，当時，ナチスのドイツでは，ブラウンらがロケットV2号の開発に信頼性の考え方を応用して成果をあげたということも有名である。

■歴史　1940年代に芽ばえた信頼性の概念が本格的に研究されるようになったのは，米国にAGREE(Advisory Group on Reliability of Electronic Equipment)が発足してからといえる。AGREEは1952年に，当時複雑なシステムの信頼性を確保するために，官，民，学の3者の人々よりなる研究グループで，米国国防省の中に設置された。AGREEは信頼性の概念を明確に，これを向上するための方策を研究し，1957年にはAGREE報告が発行され，ついに，このころに信頼性の基礎が固められた。

一方，この時代には，寿命分布として重要な役割を持つ指数分布やワイブル分布などの統計的な研究も急速に進み，かつ，FMEAやFTAなども1950年代の末には試用され始めていた。

1960年代に入ると，信頼性は単なる研究の対象としてではなく，実用的に役立てられる方向へ向かった。たとえば，1959年の自動車のマイル保証や1961年にスタートしたアポロ計画などは，信頼性工学の発展を促進したといえよう。また，1954年と1955年の間に発生した世界で初めてのジェット旅客機コメットの3回の墜落事故は，その原因が信頼性に帰せられたことにより，航空工業界が1960年ごろには信頼性研究に力を入れたこともあげなければならない。

信頼性の理論も1960年代には大きな飛躍をとげ，OR的手法である待ち行列論なども信頼性研究に役立てられた。

1970年代に入ると，とくにわが国では，品質管理の方面より信頼性が注目されるようになった。そして，次第に品質管理の中の一分科を形成することになった。

■信頼性の位置　航空機や自動車などを利用するに当たっては，これらが故障しないことは重要なことであろう。このような複雑なシステムの安全性を保証し，向上せしめる科学的手法として，種々の固有の工学に横断的にまたがって発展した信頼性工学は，現在の社会において欠くことのできない意義をもつ。参考文献〔25〕〔49〕〔51〕（真壁）

数理計画法

広く解釈すれば，「最も望ましい計画を作成するために用いられる数理的方法」(JIS－Z－1821)であるが，ふつうはもっと狭く，「ある制約式のもとで，ある目的関数を最適化する手法」を指す。

★解説

上の定義に使った用語の説明をしておく。

制約式　変数間の関係を示す等式または不等式，たとえば生産計画を作るための問題では，原料や諸資源，更には納期や倉庫の容量といったさまざまな局面での制限を表現する式である。それ故，この式を満たす変数の組が実行可能な計画を表している。

目的関数　生産計画ならば，その計画によって得られる総利益などで，その計画のよさや悪さを測る尺度である。この関数の値を大きくしたり小さくしたりすることを「目的」とするからこの名がある。

最適化　制約式を満たす変数の組の中で，目的関数の値を最大(もしくは最小)にするものを見つけること。

数理計画法のうち最も大切なものは**線形計画法**である。これは，制約式も目的関数も1次式で表されるものであって，計算方法が確立しているので利用しやすい。これらの式の一方が1次式でないときは，一般には**非線形計画法**に分類される。最近，その計算方法についての理論の発展は目覚ましいので応用例も増加しつつある。

変数が整数に限られるときは，たとえ制約式と目的関数の双方とも1次式であっても線形計画法の枠内にははいらないので，特に**整数計画法**とよばれる。整数計画法の問題を解く手間は一般には大変だが，最近は実用的に許される方法が開発されるようになった。

このほか，**動的計画法**や**ネットワーク・フロー**の問題なども数理計画法に属する手法である。

関連ページ

線形計画法　218
非線形計画法　312
整数計画法　208
動的計画法　270
ネットワーク・モデル　284

■計画の問題　どの工場で何をどのくらい作ったらよいか，または何をいつ作るのが適当か，あるいは，それぞれの販売店に製品を配送するのは，どの倉庫からいくらぐらい送るのが合理的か。企業にあっては，この種の問題は枚挙にいとまがないほど山積している。

多くの場合，従来からの経験があるので，大体において満足のできる計画が立てられているであろう。しかし，昔はどうしても必要だった条件を満たすために定められていた慣習的な決め方が，いまでは全く必要がなくなってしまったのに，なお慣習的に使われていたり，もともと制約条件もよくわからなかったため適当に決めていたりした場合には，従来のやり方で作る計画では満足しかねることもあるであろう。

全く白紙の状況で計画を立てなければならないことも多い。そういったとき，何が制約条件で，何が目的関数であるかを意識するだけでも，有意義であろう。それは，複雑なシステムを，それを構成する互いに関連した諸活動に還元して考えることに当たっているからである。まして，その結果，望ましい計画が数値的にきちんと求められれば大変好都合である。

■数理計画法の活用　はじめに例示した生産計画の問題や輸送配送の問題のみならず，設備計画や投資計画，人事計画など経営全般にわたって，それら，もしくはその一部を数式化して，具体的な計画立案の助けとする試みは非常に多く行われ，かなりの成果を上げている。そして，広義の解釈に従えばもちろんのこと，狭義の解釈に従っても数理計画法の範疇としてとらえられるものが，それらの大勢を占めている。

狭義の数理計画法，すなわち，「ある制約式のもとである目的関数を最適化する」方法が計画立案のために有効であるのは，制約を満たさない限り実行不可能であるし，実行可能な計画の中では「できるだけよい計画」を求めることを万人が望むためである。いいかえると，万人に理解されるように，制約式と目的関数を設定することが大切である。そのためにも定式化に当たっては，多くの人の意見も聞いて納得のいく形で定式化を志すべきであろう。

■線形計画法　左のページでも述べたように，数理計画法の中で中心をなすのは線形計画法である。歴史的に見ると，ダンツィーク(Dantzig, G.)らによって線形計画法が提唱されたのが1947年，電子計算機を用いて成功裡に解かれた最初の例は1952年であって，以来各方面でさまざまな応用が試みられてきている。人類の諸活動には，基本的に比例関係で記述されるものが多い。生産費も輸送費も，変動費は(単価×数量)という形で表される。それで1次式を用いて表現できる現象が多く，この事実が線形計画法を最も重要な手法にしているといえよう。　(森村)

数量化Ｉ類

林の数量化Ｉ類は，数量で与えられた外的基準を説明変数に用いる質的データにカテゴリー数量とよばれる値を決めて予測する手法である。説明変数が質的データでなく数量の場合は，重回帰分析になる。したがって，適用分野や結果の解釈等は重回帰分析と同じに考えてよい。質的データをダミー変数で表して，重回帰分析として扱ってもよい。

★解説

■図的解説　n 個の対象に関して，表1に示す通り，外的基準 y_i と p 個のアイテムのデータがとられているものとする。アイテムのどのカテゴリーをとったかを V 印で示す。1アイテム内では，必ず1個のカテゴリーの値をとるものとする。

アイテム	1				2			・・・	P			サンプルスコア
カテゴリー	1	2	3	4	1	2	3		1	2	3	
カテゴリースコア	x_{11}	x_{12}	x_{13}	x_{14}	x_{21}	x_{22}	x_{23}		x_{p1}	x_{p2}	x_{p3}	
y_1	V					V					V	$\alpha_1=x_{11}+x_{22}+\cdots+x_{p3}$
y_2		V				V			V			$\alpha_2=x_{12}+x_{22}+\cdots+x_{p1}$
⋮												⋮
y_n				V	V					V		$\alpha_n=x_{14}+x_{21}+\cdots+x_{p2}$

表1　数量化Ｉ類のデータ構成

i アイテムの j カテゴリーに，x_{ij} というカテゴリースコアが与えられたものとすれば，1番目のサンプルについて対応するカテゴリースコアの和 α_1（サンプルスコアとよぶ）が求まる。これは y_1 の予測値になる。

■アルゴリズム　カテゴリースコアは，最小2乗法を用いて誤差平方和 $\Sigma(y_i-\alpha_i)^2$ を最小にするように決められる。この基準は y_i と α_i の相関係数（重相関係数）を最大にすることに等しい。第1カテゴリーを除く残りのすべてのカテゴリーに，ダミー変数を割り当て重回帰分析を行えばよい。

関連ページ

重回帰分析　170

重相関係数　174

質的データの解析　162

林の数量化理論　298

数量化Ⅱ類　190

数量化Ⅲ類　194

数量化Ⅳ類　196

デザイン行列　250

■簡単な例 アイテム X_1（3カテゴリー）と X_2（2カテゴリー）の反応パターンが，表2で示されているデータを用いて数量化Ⅰ類を行う。

外的基準	X_1			X_2		サンプル
Y	X_{11}	X_{12}	X_{13}	X_{21}	X_{22}	スコア
0	∨			∨		$X_{11}+X_{21}$
0	∨				∨	$X_{11}+X_{22}$
1		∨		∨		$X_{12}+X_{21}$
1			∨	∨		$X_{13}+X_{21}$
2		∨			∨	$X_{12}+X_{22}$
3			∨		∨	$X_{13}+X_{22}$

表2　データ

カテゴリーに反応（∨）したものを1，無反応のものに0を与え，ダミー変数に置き換えて考える。

■解析1 X_{11} 列と X_{21} 列を落として，定数項を入れたモデル(1)を考える。

$$y=\beta_0+\beta_{12}x_{12}+\beta_{13}x_{13}+\beta_{22}x_{22} \quad (1)$$

デザイン行列 X_1, 逆行列 $(X_1'X_1)^{-1}$, 係数 $\boldsymbol{\beta}_1$ は次の通りである。

$$\boldsymbol{X}_1=\begin{array}{c}\begin{array}{cccc}\text{定} & X_{12} & X_{13} & X_{22}\end{array}\\ \begin{bmatrix}1&0&0&0\\1&0&0&1\\1&1&0&0\\1&0&1&0\\1&1&0&1\\1&0&1&1\end{bmatrix}\end{array},\quad \boldsymbol{X}_1'\boldsymbol{X}_1=\begin{bmatrix}6&2&2&3\\2&2&0&1\\2&0&2&1\\3&1&1&3\end{bmatrix}$$

$$(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}=\begin{bmatrix}2/3&-1/2&-1/2&-1/3\\-1/2&1&1/2&0\\-1/2&1/2&1&0\\-1/3&0&0&2/3\end{bmatrix}$$

$\boldsymbol{\beta}_1=[-1/2,\ 3/2,\ 2,\ 1]$

■解析2 定数項と X_{21} 列を省いた次のモデル(2)を考える。

$$y=\beta_{11}x_{11}+\beta_{12}x_{12}+\beta_{13}x_{13}+\beta_{22}x_{22}+\varepsilon \quad (2)$$

デザイン行列 X_2, 逆行列 $(X_2X_2)^{-1}$, 係数 $\boldsymbol{\beta}_2$ は次の通りである。

$$\boldsymbol{X}_2=\begin{array}{c}\begin{array}{cccc}X_{11} & X_{12} & X_{13} & X_{22}\end{array}\\ \begin{bmatrix}1&0&0&0\\1&0&0&1\\0&1&0&0\\0&0&1&0\\0&1&0&1\\0&0&1&1\end{bmatrix}\end{array},\quad \boldsymbol{X}_2'\boldsymbol{X}_2=\begin{bmatrix}2&0&0&1\\0&2&0&1\\0&0&2&1\\1&1&1&3\end{bmatrix}$$

$$(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}=\begin{bmatrix}2/3&1/6&1/6&-1/3\\1/6&2/3&1/6&-1/3\\1/6&1/6&2/3&-1/3\\-1/3&-1/3&-1/3&2/3\end{bmatrix}$$

$\boldsymbol{\beta}_1'=[-1/2,\ 1,\ 3/2,\ 1]$

■両係数の比較 解析1で省いた X_{11} と X_{21} の係数と，解析2の定数項と X_{21} の係数を0として，表3にまとめて比較する。解析1のアイテム X_1 の

	β_0	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}
解析1	$-\frac{1}{2}$	0	$\frac{3}{2}$	2	0	1
解析2	0	$-\frac{1}{2}$	1	$\frac{3}{2}$	0	1

表3　カテゴリースコア

カテゴリースコアから1/2を減じたものが解析2のカテゴリースコアになる。ただし，定数項には1/2を足して0にする。一方，$\beta_{21}x_{21}+\beta_{22}x_{22}$ で，アイテム X_2 の値を求めた合計は3である。よってデータ平均値は1/2になるので，β_{21} と β_{22} から1/2を引けば，カテゴリー X_2 でのデータ平均がゼロになるよう変換できる。（新村）

数量化Ⅱ類

林の数量化Ⅱ類は，説明変数が質的データの場合の判別分析に相当する。ただし，3群以上の判別はいわゆる次元の減少を伴う判別分析になる。2群判別の場合には，外的基準として一方の群に0，他方の群に1を与えて，重回帰分析(数量化Ⅰ類)を行っても同じ結果になる。

★解説

■図的解説　n個の対象が，G_1群(n_1)とG_2群(n_2)の2群に分かれ，p個のアイテムのデータがとられているとする。∨印でアイテムのどのカテゴリーをとったかを示す。

アイテム		1			2		…	P		サンプルスコア
カテゴリー		1	2	3	1	2	…	1	2	
カテゴリースコア		x_{11}	x_{12}	x_{13}	x_{21}	x_{22}	…	x_{p1}	x_{p2}	
G_1	$y_1^1=0$	∨			∨			∨		$\alpha_1^1=x_{11}+x_{21}+\cdots+x_{p1}$
	⋮									
	$y_{n1}^1=0$		∨			∨			∨	$\alpha_{n1}^1=x_{12}+x_{21}+\cdots+x_{p2}$
G_2	$y_1^2=1$			∨		∨			∨	$\alpha_1^2=x_{13}+x_{22}+\cdots+x_{p2}$
	⋮									
	$y_{n2}^2=1$			∨	∨				∨	$\alpha_{n2}^2=x_{13}+x_{21}+\cdots+x_{p2}$

表1　数量化Ⅱ類のデータ構成

数量化Ⅰ類と同じく，i アイテムの j カテゴリーにx_{ij}というカテゴリースコアが与えられたものとすれば，対応するカテゴリースコアの和でもってサンプルスコアが求まる。

■アルゴリズム　G_i群のサンプルスコアの平均を $\alpha_i=\sum_k \alpha_{ik}/n_i$ とし，全体の平均を $\alpha_{..}=\sum_i\sum_k \alpha_{ik}/n$ と表す。全体の分散は $\sigma_T^2=\sum_i\sum_k(\alpha_{ik}-\alpha_{..})^2/n$，群間の分散は $\sigma_b^2=\sum(\alpha_{i.}-\alpha_{..})^2 n_i/n$ になる。この比を，相関比 $\eta^2=\sigma_b^2/\sigma_T^2$ という。このη^2を最大にするようなカテゴリースコアを求めれば，2群が最もよく判別されることになる。G_1群の外的規準として0，G_2群に対しては1を与えて重回帰分析を行ってもよい。η^2は3群以上でも同様に定義できる。

関連ページ

判別分析　302

質的データの解析　162

林の数量化理論　298

数量化Ⅰ類　188

数量化Ⅲ類　194

ベイズの定理を用いた判別　346

FUNCATによる判別　320

デザイン行列　250

決定行列　104

ROC曲線　14

■**実例**　2アイテム5カテゴリーのデータを用いて，数量化Ⅱ類を説明する。6件のデータを2群 (G_1, G_2) と3群 (G_1', G_2', G_3') に分ける。

カテゴリースコア		x_1	x_2	x_3	x_4	x_5	サンプルスコア
G_1	G_1'	1	0	0	1	0	X_1+X_4
		0	1	0	1	0	X_2+X_4
	G_2'	1	0	0	0	1	X_1+X_5
G_2		0	0	1	0	1	X_3+X_5
	G_3'	0	1	0	0	1	X_2+X_5
		0	0	1	1	0	X_3+X_4
2群	f_i	2	2	2	3	3	
	f_i^1	2	1	0	2	1	
	f_i^2	0	1	2	1	2	

f_i はカテゴリー x_i の合計を，f_i^1 は G_1 群での合計を示す。$f_i=f_i^1+f_i^2$ である。この値から，次の行列 $\boldsymbol{S}_B$ と $\boldsymbol{S}_T$ を計算する。

$$\boldsymbol{S}_B=1/3(f_{ij}^1)+1/3(f_{ij}^2)-1/6(f_{ij})$$

$$=\begin{bmatrix} 2/3 & 0 & -2/3 & 1/3 & -1/3 \\ 0 & 0 & 0 & 0 & 0 \\ -2/3 & 0 & 2/3 & -1/3 & 1/3 \\ 1/3 & 0 & -1/3 & 1/6 & -1/6 \\ -1/3 & 0 & 1/3 & -1/6 & 1/6 \end{bmatrix},$$

$\mathrm{rank}(\boldsymbol{S}_B)=1$

ここで，行列 (f_{ij}) は i 行 j 列の要素が f_i*f_j の積であることを示す。これらの行列をそれぞれのデータ件数で割ったものの和と差で $\boldsymbol{S}_B$ が与えられる。元の0/1データの級間平方和積和行列になる。たとえば，$\boldsymbol{S}_B$ の (1, 1) 要素は x_1 の級間平方和を与える。

$$\boldsymbol{S}_B(1,\ 1)=3(2/3-2/6)^2+3(0/3-2/6)^2$$
$$=2/3$$

また，(1, 3) 要素は x_1 と x_3 の級間積和を与える。

$$\boldsymbol{S}_B(1,\ 3)=3(1/3)(-1/3)$$
$$+3(-1/3)(1/3)=-2/3$$

一方，$\boldsymbol{S}_T$ は元のデータを (6×5) の行列 D とすれば次で与えられる。

$$\boldsymbol{S}_T=D'D-1/6(f_{ij})$$

$$=\begin{bmatrix} 4/3 & -2/3 & -2/3 & 0 & 0 \\ -2/3 & 4/3 & -2/3 & 0 & 0 \\ -2/3 & -2/3 & 4/3 & 0 & 0 \\ 0 & 0 & 0 & 3/2 & -3/2 \\ 0 & 0 & 0 & -3/2 & 3/2 \end{bmatrix},$$

$\mathrm{rank}(\boldsymbol{S}_T)=3$

$\boldsymbol{S}_T$ (1, 1)=4/3 は x_1 の全平方和を，$\boldsymbol{S}_T$ (1, 2)=−2/3 は x_1 と x_2 の積和を表す。

■**アルゴリズム**　判別分析は，次の相関比 η^2 を最大にするようにカテゴリースコア $\boldsymbol{x}$ を決める。

$$\eta^2=\frac{\boldsymbol{x}'\boldsymbol{S}_B\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{S}_T\boldsymbol{x}}\quad\rightarrow\text{最大}$$

これは，$\boldsymbol{x}'\boldsymbol{S}_T\boldsymbol{x}=1$ の条件の下で次の固有値問題になる。

$$(\boldsymbol{S}_B-\eta^2\boldsymbol{S}_T)\boldsymbol{x}=0$$

すなわち，行列式 $|\boldsymbol{S}_B-\eta^2\boldsymbol{S}_T|=0$ を満たす固有値 η^2 と固有ベクトル $\boldsymbol{x}$ を求めればよい。これは $\boldsymbol{S}_T$ が存在すれば，$|\boldsymbol{S}_T^{-1}\boldsymbol{S}_B-\eta^2\mathbf{E}|=0$ になる。$\boldsymbol{S}_B$ のランクは群間の自由度 $(g-1)$ を越

えない。すなわち，2群判別では $g=2$ より1個の固有ベクトル(カテゴリースコア)が求まる。

■計算過程 数量化Ⅱ類では，アイテム内のカテゴリーの和が $1(X_1+X_2+X_3=1,\ X_4+X_5=1)$ という制約があるので，普通は第1カテゴリー X_1 と X_4 を省くが，本データは件数が少なく計算上不都合が起こるので，X_2 と X_4 を省くことにする。

$$S_T^{-1}=\begin{bmatrix} 4/3 & -2/3 & 0 \\ -2/3 & 4/3 & 0 \\ 0 & 0 & 3/2 \end{bmatrix}^{-1}=\begin{bmatrix} 1 & 1/2 & 0 \\ 1/2 & 1 & 0 \\ 0 & 0 & 2/3 \end{bmatrix}$$

$$S_T^{-1}S_B=\begin{bmatrix} 1/3 & -1/3 & -1/6 \\ -1/3 & 1/3 & 1/6 \\ -2/9 & 2/9 & 1/9 \end{bmatrix}$$
rank=1

固有値を $\lambda\ (=\eta^2)$ として，

$$\begin{vmatrix} 1/3-\lambda & -1/3 & -1/6 \\ -1/3 & 1/3-\lambda & 1/6 \\ -2/9 & 2/9 & 1/9 \end{vmatrix}=\lambda^2(\lambda-7/9)=0$$

を解いて $\eta^2=7/9$ が求まる。固有ベクトルは次の連立方程式を解けばよい。

$$\begin{bmatrix} 1/3 & -1/3 & -1/6 \\ -1/3 & 1/3 & 1/6 \\ -2/9 & 2/9 & 1/9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix}=7/9\begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix}$$

$x_1+x_3=0$ と $2x_3=3x_5$ が求まる。l を媒介変数として $(x_1,\ x_3,\ x_5)=(l,\ -l,\ -2/3l)$ になる。この l は，$\boldsymbol{x}'\boldsymbol{S}_T\boldsymbol{x}=1$ の制約条件より $l=\sqrt{3/14}=0.4629$ になる。省いた X_2 と X_4 のカテゴリスコアを0とみなせば，次の解が得られる。

$(x_1\ x_2\ x_3\ x_4\ x_5)=(0.4629,\ 0,\ -0.4629,\ 0,\ -0.3086)$。これからサンプルスコアを計算すれば，次の分布が得られる。G_1 群では，$(y_1\ y_2\ y_3)=(0.463,\ 0,\ 0.154)$。$G_2$ 群では，$(y_4\ y_5\ y_6)=(-0.772,\ -0.309,\ -0.463)$。$G_1$ の平均は0.206で，G_2 の平均は -0.514，そして全体平均は -0.154 になる。全体平均値を判別境界とすれば，誤分類数ゼロになる。

■2群判別の場合の場合には，固有値問題のかわりに次の連立方程式を解く。

$$\boldsymbol{S}_T\boldsymbol{x}=\mathbf{C}\ \Rightarrow\ \boldsymbol{x}=\boldsymbol{S}_T^{-1}\mathbf{C}$$

$\mathbf{C}$ は2群のカテゴリー平均の差である。(1, 1) 要素は，X_1 の平均の差 $2/3(=2/3-0/3)$ になる。$\boldsymbol{x}'=(x_1\ x_3\ x_5)=(1/3,\ -1/3,\ 2/9)$ になる。固有方程式の解とは定数倍だけ異なっているが，判別成績に影響を与えない。

■3群判別 6件のデータを3群 G_1'，G_2'，G_3' の3群に分ける。級間平方和積和行列は次の通りになる。

$$S_B=\begin{bmatrix} 1/3 & -1/6 & -1/6 & 0 & 0 \\ -1/6 & 1/3 & -1/6 & 1/2 & -1/2 \\ -1/6 & -1/6 & 1/3 & -1/2 & 1/2 \\ 0 & 1/2 & -1/2 & 1 & -1 \\ 0 & -1/2 & 1/2 & -1 & 1 \end{bmatrix}$$
rank=2

2群判別と同じく x_2 と x_4 を省略する。

$$S_T^{-1}S_B=\begin{bmatrix} 1/4 & 0 & 1/4 \\ 0 & 1/4 & 1/2 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$
rank=2

固有値問題を解いて $\lambda=11/12,\ 1/4,\ 0$ が求まる。3群判別 $(g=3)$ では2個の正の固有値と固有ベクトルが求まる。$\lambda=11/12$ の場合の固有ベク

トルは$(l_1, 2l_1, 8/3l_1)'$になる。制約条件式より，$l_1^2=3/44$なので，(0.261，0.522，0.696)になる。第1アイテムのカテゴリースコア$X_1=0.261$，　$X_2=0$，$X_3=0.522$を用いて，6サンプルのサンプルスコアを計算すると，0.261, 0, 0.261, 0.522, 0，0.522になり，データ平均は0.261になる。

							第1サンプルスコア	第2サンプルスコア
第1カテゴリースコア(f_1)		0	−0.261	0.261	−0.348	0.348		
第2カテゴリースコア(f_2)		0.426	−0.426	0	0.213	−0.213		
G_1'	y_1	1	0	0	1	0	−0.348	0.639
	y_2	0	1	0	1	0	−0.609	−0.213
G_2'	y_3	1	0	0	0	1	0.348	0.213
	y_4	0	0	1	0	1	0.609	−0.213
G_3'	y_5	0	1	0	0	1	0.097	−0.639
	y_6	0	0	1	1	0	−0.097	0.213

データ平均がゼロになるようにカテゴリースコアから0.261を引いたものが，最終的なカテゴリースコアとしてもよい。すなわち $X_1=0$，$X_2=-0.261$，$X_3=0.261$，$X_4=-0.348$，$X_5=0.348$になる。

$\lambda=1/4$ の固有ベクトルは，データ件数が少ないためゼロになる(普通には起こらない)。そこで仮にこれを，$X_1=0.426$，　$X_2=-0.426$，　$X_3=0$，$X_4=0.213$，$X_5=-0.213$ とする。

表2はこの2組のカテゴリースコアと，これから計算されたサンプルスコアを示す。

■結果の解釈　第1カテゴリーのスコアの組を第1軸(f_1)とし，第2カテゴリーの組を第2軸(f_2)として，カテゴリースコアの値をプロットして図1に示す。この平面上で近いものほど似ているカテゴリーと考えられる。たとえ

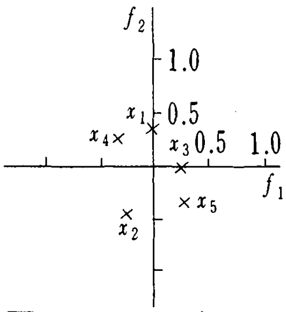


図1　カテゴリースコアのプロット

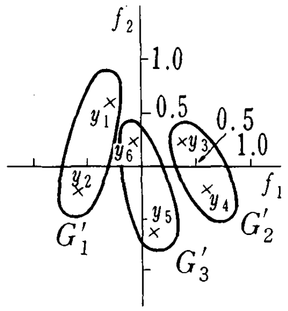


図2　サンプルスコアのプロット

ば，X_1とX_3は似ていて，X_2とは異なる。図2は同じ平面上にサンプルスコアをプロットした。第1軸上では完全に分かれているが，第2軸上では3群の判別が難しいことが分かる。また，G_3'群は他の2群の中間に位置する群であることが分かる。g個の群がある場合，$(g-1)$個の軸が得られるが，3軸から5軸程度が利用の限界である。また，多群判別よりも2群判別の組み合わせをうまく考えた方が実用的なことも多い。　(新村)

数量化Ⅲ類

林の数量化Ⅲ類は，外的基準をもたない質的データの分類手法である。アイテムの各カテゴリーと対象に同時に数値を与え，似たカテゴリーと対象を近くに集める方法である。これと同じ方法が，双対尺度法，最適尺度法，**analyse factoriel des correspondances** などと，種々の名称で研究されている。

★解説

■図的解説　n 個の対象が，アイテムの区切りをつけずに表した q 個のカテゴリーで，∨印で示される特性をもっていたとする。

対象＼特性		1	2	3	4	・	・	・	q
		x_1	x_2	x_3	x_4				x_q
1	y_1	∨	∨						∨
2	y_2		∨	∨					
3	y_3	∨			∨				
⋮	⋮								
n	y_n	∨	∨						∨

表1　数量化Ⅲ類のデータ構成

i 番目の対象に y_i，j カテゴリーにカテゴリースコア x_j を同時に与える。対象1とカテゴリー1のように(∨印で示す)，∨印のあるもののみをあらためて2変数 y と x のデータ (y_1, x_1) とみなし，この相関係数 ρ_{xy} を最大にするように y_i と x_j を決める。

表1の∨印のものを，(y_i, x_j) の値に応じて2次元平面にプロットすれば散布図が得られる。この散布図は，y_i $(i=1, \cdots, n)$ と x_j $(j=1, \cdots, q)$ に任意の値を与えて得られるものの中で，直線相関(分布が直線上に並ぶ)の最も高いものになる。また，x_i と y_j の値をプロットすれば，対象とカテゴリーで似たものが近くにプロットされ分類できる。

関連ページ
質的データの解析　162
林の数量化　298
数量化Ⅰ類　188
数量化Ⅱ類　190
数量化Ⅳ類　196
MDS　42

$$
\begin{array}{c} \text{(i)} \\ \text{イ} \\ \text{ロ} \\ \text{ハ} \\ \text{ニ} \end{array}
\begin{array}{c} a\ b\ c \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \end{array}
\quad
\begin{array}{c} \text{(ii)} \\ \text{イ} \\ \text{ハ} \\ \text{ニ} \\ \text{ロ} \end{array}
\begin{array}{c} a\ b\ c \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \end{array}
\quad
\begin{array}{c} \text{(iii)} \\ \text{イ} \\ \text{ハ} \\ \text{ニ} \\ \text{ロ} \end{array}
\begin{array}{c} a\ c\ b \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ y_1\, y_2\, y_3 \end{array}
\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ \\ \end{array}
$$

図1　自動車嗜好アンケート

■例題　3種の自動車 a, b, c の好き(1)，嫌い(0)をイ，ロ，ハ，ニの4人にアンケートを行い，図1の(i)の結果を得た。一見して何の関連もないデータを，ロをニの後ろに並べ換え(ii)，その後で b と c を入れ換えて，(iii)のように対角線上に1が集まる規則的な表が得られる。数量化Ⅲ類は，この並べ換えを行う。

■アルゴリズム　イハニロに，x_1, x_2, x_3, x_4, そして acb に y_1, y_2, y_3 の数値を与える。次に，イと a のように1のあるもののみを$(x_1 y_1)$の値をもつデータと考える。全部で6件ある。この x と y の相関係数 ρ_{xy} を最大にするように，各 x_i と y_i の値 $(\boldsymbol{x}, \boldsymbol{y})$ を決めるのが数量化Ⅲ類である。

P を図1の(iii)の行列。P_x を個人ごとに1と回答した数(イは1，ハは2など)を対角要素とする対角行列。P_y は車ごとに好きと回答された数(acb とも各2)を対角要素とする対角行列。この時，$\boldsymbol{y}$ は次の固有値問題で求まる。

$$(P_y^{-\frac{1}{2}}P'P_x^{-1}PP_y^{-\frac{1}{2}}-\rho_{xy}{}^2\boldsymbol{E})P_y^{\frac{1}{2}}\boldsymbol{y}=0$$

$P_y^{-\frac{1}{2}}P'P_x^{-1}PP_y^{-\frac{1}{2}}$ の固有値 $(\rho_{xy}{}^2)$ は，$\lambda=1$, $3/4$, $1/4$ になる。第1固有値は必ず1になり，数量化Ⅲ類では用いない。第2，第3固有値に対する固有ベクトル $P_y^{\frac{1}{2}}\boldsymbol{y}_2$ と $P_y^{\frac{1}{2}}\boldsymbol{y}_3$ は次の通りになる。

$P_y^{\frac{1}{2}}\boldsymbol{y}_2=(1/\sqrt{2},\ 0,\ -1/\sqrt{2})$

$P_y^{\frac{1}{2}}\boldsymbol{y}_3=(1/\sqrt{6},\ -2/\sqrt{6},\ 1/\sqrt{6})$

これを解いて $\boldsymbol{y}_2$ と $\boldsymbol{y}_3$ が求まる。

$\boldsymbol{y}_2=(1/2,\ 0,\ -1/2)$

$\boldsymbol{y}_3=(1/2\sqrt{3},\ -1/\sqrt{3},\ 1/2\sqrt{3})$

第2，第3固有値に対する $\boldsymbol{x}$ の値は，

$$\boldsymbol{x}_i=\frac{1}{\sqrt{\lambda_i}}P_x{}^{-1}P\boldsymbol{y}_i \quad (i=2,\ 3)$$

で求まり，次の値をとる。

$\boldsymbol{x}_2=(1/\sqrt{3},\ 1/2\sqrt{3},\ -1/2\sqrt{3},\ -1/\sqrt{3})$

$\boldsymbol{x}_3=(1/\sqrt{3},\ -1/2\sqrt{3},\ -1/2\sqrt{3},\ 1/\sqrt{3})$

以上を表2にまとめて示す。この値を平面にプロットしたのが図2である。a がイとハ，b がロとニ，c がニとハと関係していることが分かる。図1の(i)のデータでも当然同じ結果になる。

	$\lambda_2=3/4$	$\lambda_3=1/4$
イ	0.577	0.577
ハ	0.289	−0.289
ニ	−0.289	−0.289
ロ	−0.577	0.577
a	0.5	0.289
c	0.0	−0.577
b	−0.5	0.289

表2　x と y の値

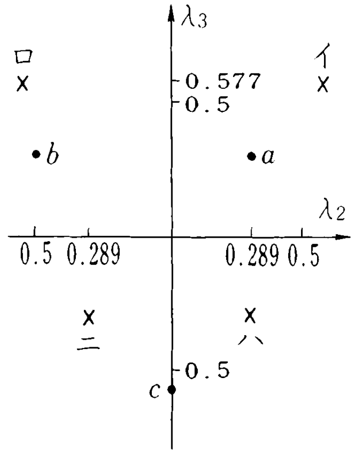


図2　車と人の配置

数量化Ⅳ類(e_{ij}型数量化)

n個の対象の2者間 (i, j) に類似度(または非類似度) e_{ij} が与えられたとき, $(n \times n)$ の類似度行列が作られる。数量化Ⅳ類はこの行列の情報に基づいて, n個の対象に座標点を与える方法である。座標にn個の点をプロットすれば, n個の対象の関係を示す地図を描くことができる。この地図上で近いものどうしは, 互いに似た対象と考えられる。

★解説

■類似度と非類似度　2つのモールス信号が, 同じ信号か否かを100人の人間に回答させる。同じ信号を聞かせて95人が同じ信号と回答した場合や, 異なった信号を聞かせて60人が異なった信号と答えた場合, これらの値を類似度という。2地点間の情報量なども類似度になる。これに対して, 2地点間の距離は値が大きいほど離れているので非類似度とよばれる。

■類似度行列　n個の対象の2者間 (i, j) の類似度 e_{ij} を, (i, j) 要素とする行列 E を類似度行列という。e_{ij} はどのような距離でもよいので, $e_{ij} \geqq 0$ として値が大きいほど i と j の類似度が高いと考える。ただし, 対角要素 e_{ij} は存在しなくてもよく, 普通はゼロと考える。また, $e_{ij} \neq e_{ji}$ でもよい。相関行列も一種の類似度行列であり, 因子分析は相関行列に基づく一種の数量化Ⅳ類と考えられる。

■アルゴリズム　対象 i に与える数値を x_i とし, 類似度 e_{ij} が大きいほど x_i と x_j の距離が小さくなるように数量化することを考える。そこで, $\sum_i x_i^2 =$ 一定(たとえば1)のもとで次の基準 Q を最大にする x_i $(i=1, \cdots, n)$ を求める。

$$Q = -\sum_{i \neq j}\sum e_{ij}(x_i - x_j)^2$$

右ページに示すように, 行列の固有値問題に帰着される。

関連ページ

質的データの解析　162

林の数量化理論　298

数量化Ⅲ類　194

MDS　42

■例題　5地点abcde間の距離が，図1のように与えられている。ab間の距離は$\sqrt{2}$であり，値が大きいほど2地点間の交流が少なくなるので，非類似度と考えられる。これを行列にまとめたものが図2の非類似度行列である。この行列の$(i,\ j)$要素をe_{ij}とする。たとえば，$e_{12}=\sqrt{2}$である。この行列から，$i \neq j$の場合には$a_{ij}=-(e_{ij}+e_{ji})$，$i=j$の場合には$a_{ii}=\sum_{k=1}^{5} a_{ik}$として求めた行列が次の$A$である。1行と2行のみを示す。

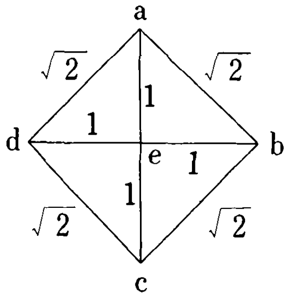


図1　5点間の距離

$$
\begin{array}{c|c}
 & \begin{array}{ccccc} \mathrm{a} & \mathrm{b} & \mathrm{c} & \mathrm{d} & \mathrm{e} \end{array} \\
\begin{array}{c} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \\ \mathrm{d} \\ \mathrm{e} \end{array} &
\begin{bmatrix}
0 & \sqrt{2} & 2 & \sqrt{2} & 1 \\
\sqrt{2} & 0 & \sqrt{2} & 2 & 1 \\
2 & \sqrt{2} & 0 & \sqrt{2} & 1 \\
\sqrt{2} & 2 & \sqrt{2} & 0 & 1 \\
1 & 1 & 1 & 1 & 0
\end{bmatrix}
\end{array}
$$

図2　非類似度行列

$$
A=\begin{bmatrix}
6+4\sqrt{2} & -2\sqrt{2} & -4 & -2\sqrt{2} & -2 \\
-2\sqrt{2} & 6+4\sqrt{2} & -2\sqrt{2} & -4 & -2
\end{bmatrix}
$$

この行列の固有値問題，すなわち，

$$A\boldsymbol{x}=\lambda \boldsymbol{x}$$

を求めることが数量化Ⅳ類になる。

固有値は，-15.657，-15.657，-13.314，-10.0になる。ゼロ以外の固有値に対応する4個の固有ベクトルをまとめて行列Bとして表す。

$$
\begin{array}{cc}
B=\begin{bmatrix}
-0.576 & 0.410 & -0.5 & -0.224 \\
0.410 & 0.576 & 0.5 & -0.224 \\
0.576 & -0.410 & -0.5 & -0.224 \\
-0.410 & -0.576 & 0.5 & -0.224 \\
0 & 0 & 0 & 0.894
\end{bmatrix} &
\begin{array}{c} \mathrm{a} \\ \mathrm{b} \\ \mathrm{c} \\ \mathrm{d} \\ \mathrm{e} \end{array}
\end{array}
$$

すなわち，1列目は固有値-15.657に対応する固有ベクトルである。1行目は地点aの座標$(-0.576,\ 0.410,\ -0.5,\ -0.224)$を表す。1列目と2列目を$x$軸と$y$軸の座標値として5地点をプロットすると，図3が求まる。図1にほぼ似た位置関係を示すプロット図になる。このようなプロット図を**データの布置**とよぶ。

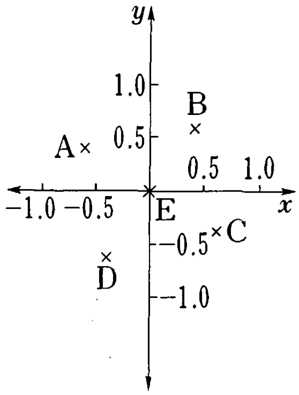


図3　5地点のプロット

すなわち，数量化Ⅳ類は，2地点間の距離のような非類似度や，2つの刺激間の類似度等が与えられた時，それらの関係を表す地図を与えてくれる。この問題では，5地点が4次元空間の点として求まる。

数量化Ⅳ類の欠点は，2次元平面上のユークリッド距離として厳密に与えられた関係がそのまま復元されないことである。B行列から各地点間の距離の2乗を計算すると，すべて2になる。たとえば，$(\mathrm{ab})^2=0.986^2+0.166^2+1^2+0^2 \fallingdotseq 2$。

多次元尺度法(MDS)の一手法であるALSCALなどを用いれば，図1の地図は復元される。　(新村)

正規確率紙

統計データの解釈を容易にするために工夫された特別の方眼紙がいくつかある。正規確率紙はその1つである。右ページに例示してあるように，横軸は普通目盛りであるが，縦軸は正規分布の特別な目盛りになっている。データから累積度数百分率の表をつくり，変量と累積度数百分率をそれぞれ横座標，縦座標として，つぎつぎと打点したとき，それらの点が同一直線上にならぶようであれば，データは正規母集団から独立にとられたものであるとみなしてよい。そのとき，平均や標準偏差がいくらぐらいかを，図から読みとることもできる。

★解説

正規確率紙の横軸は普通目盛りであるが，縦軸は特別の確率目盛りになっている。中ほどに50という目盛りがあり，それから上に60，同じ長さだけ下がったところに40，というように，50を中心にして，上下に70，30；80，20；……のように目盛られている。また，確率紙の右側に目盛ってあるμは，左側の目盛り50の位置にあるが，このμを通って横軸に平行に引いた直線と，データから打点してつくられる直線との交点の横座標を読むと，それが**母集団**の**平均**値の**推定値**になる。さらに，右側の$\mu+\sigma$は，左側の目盛り84，17に対応し，これらを結ぶ線分とデータからの直線の交点の横座標が(平均＋標準偏差)を表しているので，これから先に求めた平均を引くことにより，**標準偏差**の推定値が求まる。ふつう用いられている確率紙として，正規確率紙のほかに，**対数正規確率紙**，**ワイブル確率紙**などがあるが，使い方はどれも同じで，データから累積度数百分率を計算し，確率紙の上に打点したとき，直線上にならぶかどうかをみればよい。

関連ページ
正規分布　200
母集団　260
平均　272, 276, 344
推定値　254
標準偏差　272,332

(例) 次表は昭和55年度の国公立大学入学試験のための共通一次試験の総合得点(1000点満点)の得点分布である。

得　点	人　数	累積人数	累積人数百分率
300点以下	2,400人	2,400人	0.7 %
301～350	5,700	8,100	2.4
351～400	11,400	19,500	5.9
401～450	17,400	36,900	11.1
451～500	25,500	62,400	18.7
501～550	34,800	97,200	29.2
551～600	44,100	141,300	42.4
601～650	49,800	191,100	57.4
651～700	48,600	239,700	72.0
701～750	41,700	281,400	84.5
751～800	29,700	311,100	93.4
801～850	16,200	327,300	98.3
851～900	5,100	332,400	99.8
901 以上	600	333,000	100.0

表１　共通一次得点の度数分布(昭和55年度)

■共通一次試験　毎年，国公立大学などへの入学志願者に対する共通一次試験が実施され，総合得点(1000点満点)の度数分布などが，大学入試センターから発表される。昭和55年度の度数分布として，正式に発表された表は，241点から920点まで，10点きざみで，68の階級に区分されているが，上の度数表は，それに基づいて，大きく14段階にまとめたものである。大学入試センターで発表した，この年の受験者数は333,000人であり，また平均点＝617.4，標準偏差＝128.1となっている。

左の表に基づき，累積人数百分率を正規確率紙の上に打点すると，次図のようになる。

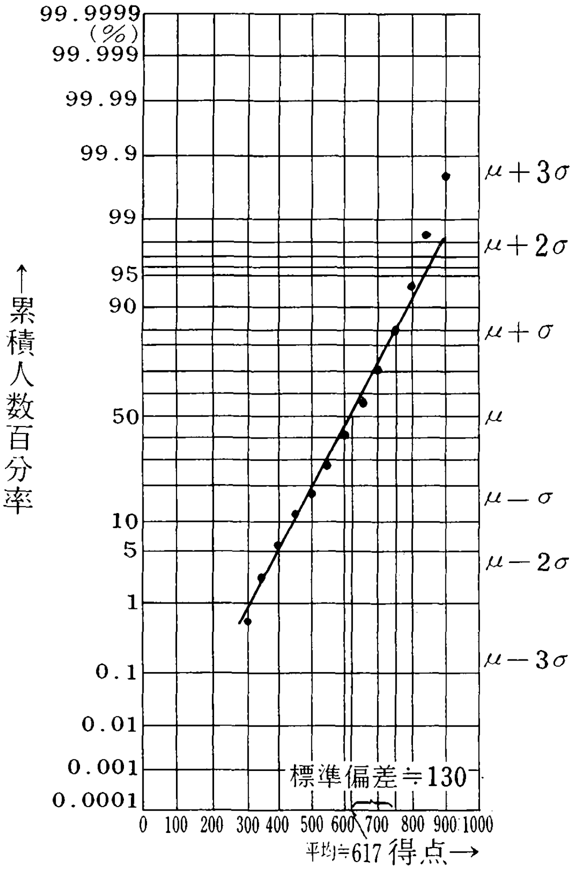


図１　正規確率紙への打点

■正規確率紙の使い方　上の例では，図からわかるように，上位得点階級については，少し直線から外れている。しかし，データから累積度数百分率を計算し，正規確率紙に打点したとき，それがぴたりと直線上にならぶということはまれで，上のように多少わん曲してくるのがふつうである。これを直線的であるとみなして，上のように直線で結び，平均，標準偏差を推定してみると，大まかに 平均点＝617，標準偏差＝130 と読みとることができる。

(牧野)

正規分布

最も基本的な確率分布の1つであり，密度関数は釣鐘型をしている。たとえば，ビール瓶の中をのぞくと，瓶ごとにビールの量が多少異なる。工場では内容量が一定となるように機械をセットしているが，何らかの影響によりこのような差が生ずる。このような誤差の分布は，**正規分布**(normal distribution)となる。人間の身長の分布，ある製品の週当たりの需要量の分布など，生物や社会の現象に適合することが多く，最も大切な確率分布である。

★解説

正規分布の**密度関数**と**特性値**は次のとおりである。

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (-\infty<x<\infty)$$

平均　$E(X)=\mu$,　　　分散　$\mathrm{Var}(X)=\sigma^2$

この正規分布を簡単に $\mathrm{N}(\mu,\ \sigma^2)$ と略記する。図1に見るように，密度関数は平均 μ に関して対称な釣鐘型をしており，分散が小さいほど平均 μ の近くの値が出やすく山が高くなっている。μ は分布全体の位置を示すので，**位置パラメータ**といい，これに対し σ^2 は測定単位の大きさによって変化するので，**尺度パラメータ**とよばれる。

とくに平均0の正規分布 $\mathrm{N}(0,\sigma^2)$ は，偶然誤差の従う分布法則としてガウス(Gauss,1777～1855)によって導かれたので，ガウス分布とか誤差分布とかよばれることもある。その後，リヤプノフ(Lyapunov,1857～1918)らによって**中心極限定理**が証明されたことにより，いろいろな現象になぜ正規分布がよくあてはまるかが理解されるようになった。

平均から 3σ 以上離れた値のでる確率はおよそ千分の三であり(図2)，きわめて小さい。これによって，並みから大部外れた人を，“彼は 3σ だ”とか“千三ツ”だともいう。

関連ページ
密度関数　56
特性値　272
中心極限定理　226

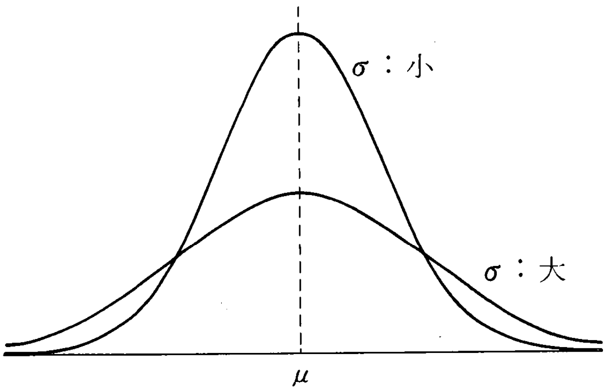


図 1　正規分布の密度関数

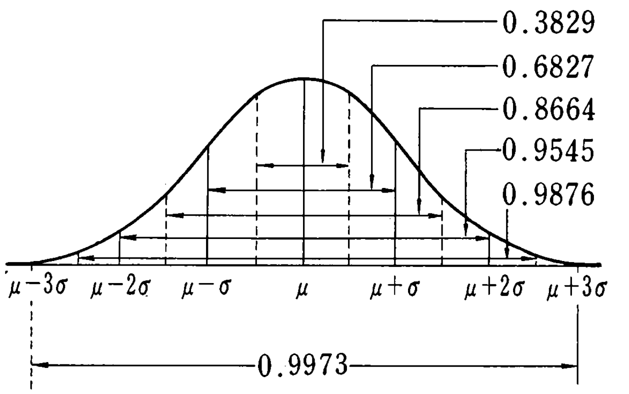


図 2　正規分布の確率

■**標準正規分布**　平均 0，分散 1 の正規分布 $N(0,\ 1)$ を，とくに**標準正規分布**とよぶ。密度関数は次のとおり。

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\ (-\infty<x<\infty)$$

標準正規分布 $N(0,\ 1)$ に従う確率変数を Z とする。表 1 は $\{Z>z\}$ となる確率，つまり図 3 の斜線部の面積 $I(z)$ の値を表にしたもので，**正規分布表**とよばれる。つまり，$I(z)$ は，

$$I(z)=\mathrm{P}\{Z>z\}=\int_z^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

を計算して求めてある。また，

$$\mathrm{P}\{Z>K_P\}=P$$

なる K_P を**上側 100P %点**といい，

$$\mathrm{P}\{|Z|>U_P\}=P$$

なる U_P を**両側 100P %点**という。平

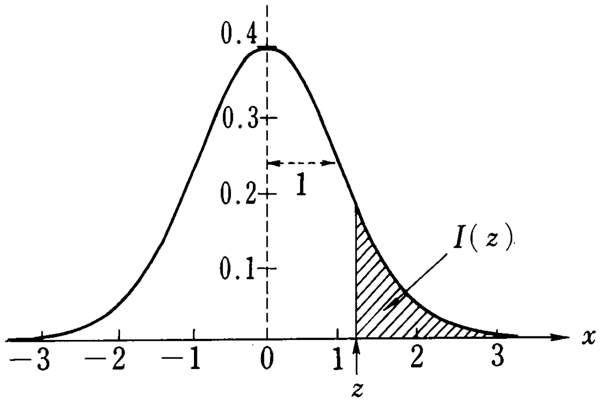


図 3　標準正規分布の密度関数

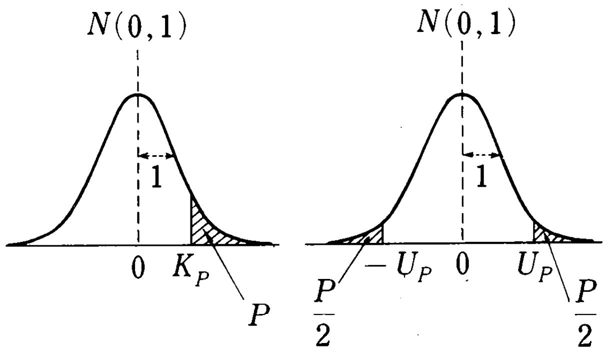


図 4　上側および両側100P%点

均値の検定や区間推点に用いる(図 4)。

任意の正規分布 $N(\mu,\ \sigma^2)$ に関する確率も，この正規分布表から以下のように求めることができる。

■**標準化**　確率変数 X を

$$X^*=\frac{X-\text{平均値}}{\text{標準偏差}}$$

と置きかえると，X^* の平均は 0，分散が 1 となる。X の**標準化**という。

X が正規分布 $N(\mu,\ \sigma^2)$ に従うとき，これを標準化した $X^*=(X-\mu)/\sigma$ は標準正規分布 $N(0,\ 1)$ に従う。

この事実を用いて，$\{X>x\}$ なる確率を正規分布表より求めてみよう。

事象 $\{X>x\}$ は事象 $\{(X-\mu)/\sigma>(x-\mu)/\sigma\}$ と全く同じである。

$$z=(x-\mu)/\sigma$$

z	.00	.02	.04	.06	.08
.0	.5000	.4920	.4840	.4761	.4681
.1	.4602	.4522	.4443	.4364	.4286
.2	.4207	.4129	.4052	.3974	.3897
.3	.3821	.3745	.3669	.3594	.3520
.4	.3446	.3372	.3300	.3228	.3156
.5	.3085	.3015	.2946	.2877	.2810
.6	.2743	.2676	.2611	.2546	.2483
.7	.2420	.2358	.2296	.2236	.2177
.8	.2119	.2061	.2005	.1949	.1894
.9	.1841	.1788	.1736	.1685	.1635
1.0	.1587	.1539	.1492	.1446	.1401
1.1	.1357	.1314	.1271	.1230	.1190
1.2	.1151	.1112	.1075	.1038	.1003
1.3	.0968	.0934	.0901	.0869	.0838
1.4	.0808	.0778	.0749	.0721	.0694
1.5	.0668	.0643	.0618	.0594	.0571
1.6	.0548	.0526	.0505	.0485	.0465
1.7	.0446	.0427	.0409	.0392	.0375
1.8	.0359	.0344	.0329	.0314	.0301
1.9	.0287	.0274	.0262	.0250	.0239
2.0	.0228	.0217	.0207	.0197	.0188
2.1	.0179	.0170	.0162	.0154	.0146
2.2	.0139	.0132	.0125	.0119	.0113
2.3	.0107	.0102	.0096	.0091	.0087
2.4	.0082	.0078	.0073	.0069	.0066
2.5	.0062	.0059	.0055	.0052	.0049
2.6	.0047	.0044	.0041	.0039	.0037
2.7	.0035	.0033	.0031	.0029	.0027
2.8	.0026	.0024	.0023	.0021	.0020
2.9	.0019	.0018	.0016	.0015	.0014
3.0	.0013	.0013	.0012	.0011	.0010

表 1　正規分布表(zから$I(z)$を求める)

とおくと，２つ目の事象は $\{X^*>z\}$ と書き直される。X^* は $N(0,\ 1)$ に従っているから，

$$P\{X>x\}=P\{X^*>z\}=I(z)$$

となり，正規分布表より該当の確率が求まる。この計算例を次に示す。

■脳の重さの分布　スウェーデンの成人男子の脳の重さを測定したところ，平均 1400.5g，分散 106.3^2 であった。図 6 は実測データの相対度数のグラフ(全面積が 1 となるように描いてある)に，正規分布 $N(1400.5,\ 106.3^2)$ をあてはめている。あてはまりのよさを見るには，本当は正規確率紙(→ 198 ページ)の上にプロットする方がよい。

このあてはめが適切として，脳の重さが 1300～1550g の間にある人の割合

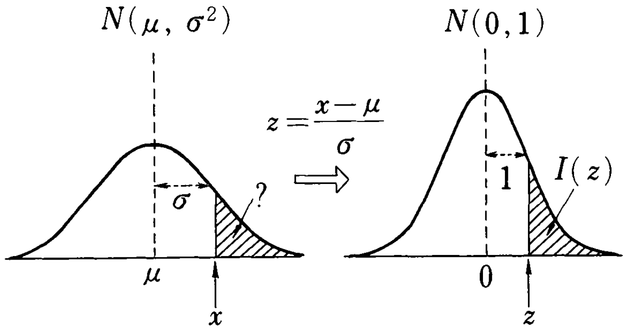


図 5　標準化

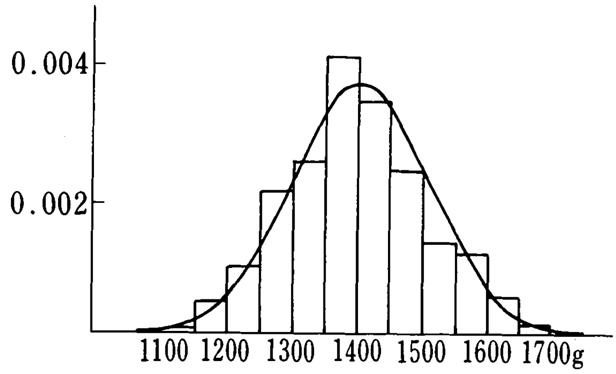


図 6　脳の重さの分布

を求める。これは先に述べたことより，

$$=P\left\{\frac{1300-1400.5}{106.3}\leqq X^*\leqq\frac{1550-1400.5}{106.3}\right\}$$

$$=P\{-0.95\leqq X^*\leqq 1.40\}$$

となる。密度関数が原点に関して対称だから，これは，

$$=1-I(0.95)-I(1.40)$$

$$=1-0.171-0.080=0.749$$

となり，求める割合は約 75 ％であることがわかる(図 6 参照)。

■正規分布の再成性　２つの確率変数 X_1，X_2 が独立(→ 229 ページ)で，それぞれ正規分布 $N(\mu_1,\ \sigma_1)$，$N(\mu_2,\ \sigma_2^2)$ に従うとき，２つの変数の和及び差はそれぞれ正規分布$N(\mu_1+\mu_2,\ \sigma_1^2+\sigma_2^2)$

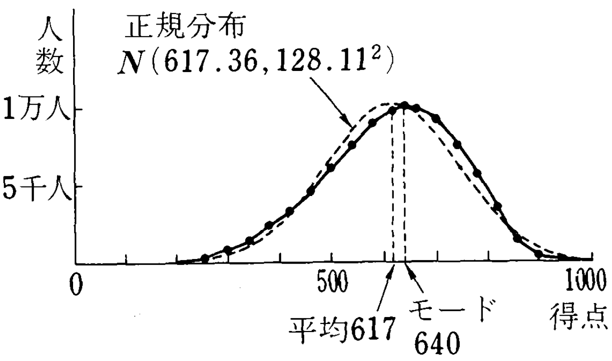


図 7　共通一次試験の得点分布
（1980.2.6付の毎日新聞を参考にして）

および $N(\mu_1-\mu_2,\ \sigma_1^2+\sigma_2^2)$ に従う。これを称して，正規分布は**再成性**をもつという。この応用を1つ示す。

あるサイズのボルトの外径 X_1 とナットの内径 X_2 がそれぞれ $N(15.0,\ 0.1^2)$ および $N(15.3,\ 0.1^2)$ に従っているとする（単位 mm）。それぞれ任意に選んで組み合わせたとき，ボルトが太すぎてはまらない確率を求めたい。

要は $X_1>X_2$，つまり $X_1-X_2>0$ の確率を求めればよい。上より，この分布は $N(-0.3,\ 0.1^2+0.1^2)$ だから，

$$\mathrm{P}\{X_1-X_2>0\}=I([0-(-0.3)]/\sqrt{0.02})$$
$$=I(1.23)=0.111$$

と求まる。

偏差値　校外模擬テストを受けると，偏差値がいくつとついてくる。偏差値は次のように計算されている。

偏差値

$$=\frac{\text{自分の得点}-\text{平均点}}{\text{標準偏差}}\times 10+50$$

もし得点の分布が正規分布であれば，偏差値の分布は $N(50,\ 10^2)$ となるはずである。偏差値68の人であれば，上位3.6％以内にいると見積もれる。

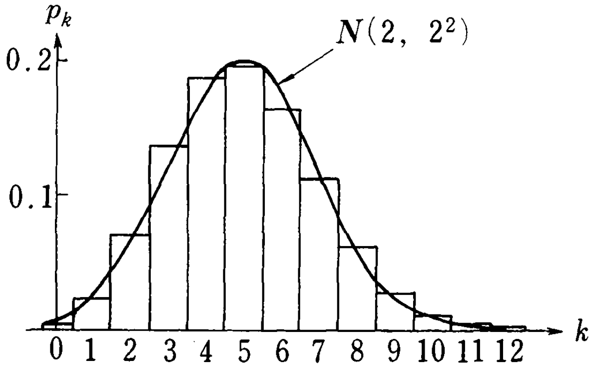


図 8　二項分布の正規分布近似
（$n=25,\ p=0.2,\ npq=4.0$ のとき）

人間の能力や学力も正規分布に従うと信じられて，通知表の5段階評価などに用いられているが，あやしいことも多いようだ。図7は国立大学共通一次試験の総得点の分布である（受験者数33.5万人）。これに同じ平均，分散をもつ正規分布をあてはめてみたが，あてはまりがよくない。得点分布はかなり右に片寄っている。たとえば，700点以上の人は，実際は95000人以上いるのに，正規分布だとすると86000人と，かなりくい違う。

二項分布の正規分布近似　図8は，$n=25,\ p=.2$ の**二項分布**（→276ページ）に，同じ平均，分散をもつ正規分布 $N(np,\ npq)$ をあてはめたものである。二項分布は正規分布で，

$${}_n\mathrm{C}_k p^k q^{n-k}$$
$$\fallingdotseq I\left(\frac{k-\frac{1}{2}-np}{\sqrt{npq}}\right)-I\left(\frac{k+\frac{1}{2}-np}{\sqrt{npq}}\right)$$

と近似され，n が大きくて，$npq\geqq 5$ のときよく合う。左辺を計算するのは大変だが，右辺なら正規分布表があればよい（→226ページ）。（森）

生産統計

経済活動の中心は生産であり，その実態を明らかにしている統計が生産統計である。農業，林業，水産業の生産については農林水産省が，鉱業，製造業，商業については通商産業省が，輸送活動については運輸省あるいは日本国有鉄道が統計を作成している。生産の推移は，自動車の生産，テレビの生産のみで判断することはできないから，数万あるいは数十万といわれる生産品目を代表する主な生産品目についてのみ生産量を把握し，これを総合して生産活動の指標としている。これが生産指数である。農業，林業，水産業，鉱業，製造業，電気・ガス業について生産指数があり，運輸については総合輸送活動指数が作成されている。

★解説

生産額の状況は，年1回(または数年に1回)行われる悉皆(しっかい)調査(一部は抽出またはカット)によって把握される。製造業については工業統計調査がある。これらの調査では品目別に生産額がわかるから，そのうちの代表的な品目についてのみ，定期的に生産量の報告を求め，その調査結果を総合して生産指数を作っている。鉱工業生産指数の作り方を現行の55年基準指数について述べると，採用品目数は，石炭，鉄鋼，鋼管，アルミニウム，掘さく機，電気洗たく機，カラーテレビ，電子計算機，乗用車，鋼船，時計，板ガラス，セメント，ポリエチレン，重油，医薬品，板紙，綿糸，綿織物，製材，飲用牛乳，小麦粉，たばこ，フィルム，電力，ガスなど534品目について毎月，生産量の調査がなされ，その品目別生産指数を使って，これにウエートをつけて総合している(ラスパイレス方式)。

鉱工業生産指数の関連指数として，生産者出荷指数，生産者製品在庫指数，同在庫率指数，生産能力指数，稼働率指数，原材料消費指数，販売業者在庫指数などがある。

関連ページ
主な指定統計一覧　424
指数　154

消費，投資，輸出などの需要が増大するとモノの生産活動が活発になり，出荷や販売も増加し，在庫は減少する。

企業は増大した需要に応じられるように生産量を増やし，生産者在庫を増やそうとする。しかし不況となって最終的な需要が停滞すると，在庫がたまるので生産量を減らすことになる。こうした景気の変動を鉱工業生産・出荷・在庫の各指数は敏感に反応する。このような性格をもつところから，これら一連の指数は景気指標の代表的な指標に数えられている。

生産指数の上昇は景気が上向いているといえ，下降は景気が後退しているといえるが，生産と景気の間には若干の時間的ズレがあるので，出荷指数や在庫指数の動きともあわせて総合的な判断を下す必要がある。

ここ10数年間の動向では，49，50年に生産水準の後退があり，51年以降ゆるやかな上昇となったが，56，57年には再び後退している。このような生産水準の変動は，業種によっても，また投資関連の生産品目であるか，消費関連の品目であるかによって変化の時期や変動の程度が異なるので，各指数の変動要因や景気変動の局面との関係を考えた上で利用するとよい。

一方，需要・供給のバランスを設備の稼働状況からみる指標として生産能力指数と稼働率指数が作成されている。前者は，「固定設備が正常な状態でフル生産したときの生産量」を指数化したものである。後者は，100 tの生産能力がある工場で80 tの生産を行っている場合，80 %の稼働率ということになるが，この稼働率を総合し，指数化したものである。

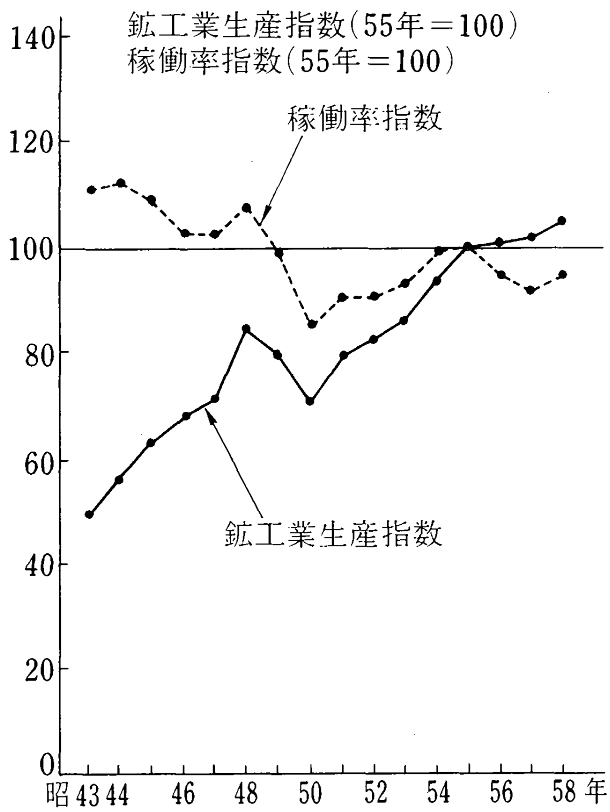


図1　鉱工業生産指数と稼働率指数

稼働率指数は，需給バランスを反映して動き，需給がひっぱくしてくれば上昇し，供給が過剰になれば低下する。

石油危機以降の低下と50年以降のゆるやかな上昇や56，57年における再度の低下が図1に示されている。なお稼働率指数が100であるというのは，稼働率が100 %であるという意味ではなく，基準時(55年)の稼働率水準との相対水準であることに注意する必要がある。　(市野)

正準相関分析

変数 x_1，x_2，…，x_q，x_{q+1}，…，x_p において，q 変数の組(x_1，x_2，…，x_q)と $p-q$ 変数の組(x_{q+1}，…，x_p)の間にどのような関係があるかを知る目的で，相関係数を最大とする q 変数の線形結合と，$p-q$ 変数の線形結合を推定する方法論を正準相関分析という。

★解説

q 個の変数と $p-q$ 個の変数の間の相関関係は，

$$q \times (p-q) \text{ 個}$$

の相関係数を調べることによって推察できる。しかしながら，この個数が多いときは容易なことでなく，各組の1次式である総合特性値

$$y = a_1x_1 + a_2x_2 + \cdots + a_qx_q$$

$$z = b_1x_{q+1} + b_2x_{q+2} + \cdots + b_{p-q}x_p$$

を考えて，y と z の相関の大きさを見ることになる。正準相関分析は，この y と z の相関係数を最大にする線形結合の係数(a_1，a_2，…，a_q)と(b_1，b_2，…，b_{p-q})を推定することである。これを y_1，z_1 で表す。

一般性を失うことなく，$q<p-q$ であるとする。y_1 と無相関である(x_1，x_2，…，x_q)の線形結合を y_2，z_1 と無相関である(x_{q+1}，…，x_p)の線形結合を z_2 とし，無相関であるという条件のもとで，y_2 と z_2 との相関が最大のものを推定する。一般性を失うことなく $q<p-q$ とすると，この手順の繰り返しにより，それぞれの q 個の正準変量

$$(y_1 \quad y_2 \quad \cdots \quad y_q)$$

$$(z_1 \quad z_2 \quad \cdots \quad z_q)$$

を得る。(y_j，z_j)に相関係数の値が対応している。

関連ページ

相関関係　222

高等学校での5教科の成績と，共通一次および二次試験の成績を調べる。高校の5教科には，国語 x_1，社会 x_2，数学 x_3，理科 x_4，英語 x_5 である。共通一次の成績は国語 x_6，社会 x_7，数学 x_8，理科 x_9，英語 x_{10}，二次試験 x_{11} である。

北国のある大学の某学部を受けた受験生の中から，400人を無作為抽出して調べたところ，相関行列は以下のようであった。

高校の	国語	社会	数学	理科	英語
	x_1	x_2	x_3	x_4	x_5
x_2	.60	1			
x_3	.47	.48	1		
x_4	.57	.62	.60	1	
x_5	.64	.54	.58	.58	1

共通一次の	国語	社会	数学	理科	英語	二次
	x_6	x_7	x_8	x_9	x_{10}	x_{11}
x_7	.15	1				
x_8	.12	.10	1			
x_9	.21	.23	.21	1		
x_{10}	.27	.14	.30	.19	1	
x_{11}	.19	.13	.29	.29	.30	1

正準相関分析で重要な役割を果たすのは，次の相関行列の部分である。

	x_1	x_2	x_3	x_4	x_5
x_6	.39	.18	.00	.14	.13
x_7	.10	.37	.08	.22	.09
x_8	.24	.17	.45	.35	.25
x_9	.28	.29	.14	.40	.23
x_{10}	.33	.23	.19	.21	.46
x_{11}	.65	.64	.60	.67	.68

これより第1次正準変量として，

$$y_1 = .22x_1 + .25x_2 + .15x_3 + .30x_4 + .31x_5$$

$$z_1 = .01x_6 + .10x_7 + .10x_8 + .10x_9 + .11x_{10} + .86x_{11}$$

を得る。y_1 と z_1 の相関係数(正準相関係数)は0.826である。y_1 は高校の成績の総合点によく似た変量であり，z_1 は二次試験の変数である。

第2正準変量は，

$$y_2 = -.63x_1 - .20x_2 + 1.10x_3 + .27x_4 - .40x_5$$

$$z_2 = -.57x_6 - .03x_7 + .80x_8 - .13x_9 - .46x_{10} + .06x_{11}$$

である。y_2 は数字が正で1.10，国語と英語は負で -0.63，-0.40，z_2 は数字が正で0.80，国語と英語は負で -0.57，-0.46 と比較的大きい。正準変量 y_2 と z_2 は数学，国語，英語に関する高校の成績と共通一次の成績との関係を示しており，正準相関係数は0.486である。

第3正準変量は，

$$y_3 = .41x_1 + .97x_2 - .38x_3 + .51x_4 - .71x_5$$

$$z_3 = -.02x_6 + .78x_7 - .22x_8 + .32x_9 - .56x_{10} + .00x_{11}$$

である。第3正準変量は社会，英語，それに理科に関連したものであり，正準相関係数は0.448である。　(杉山)

整数計画法

広い意味では，変数のとりうる値が整数に限るという条件(整数条件という)のついた数理計画法を指すが，通常は線形計画法で整数条件のついたものを指す。変数が0または1の値しかとらないとき，0-1整数計画問題という。また，一部の変数だけが整数であればよいとき，混合整数計画問題とよぶ。整数条件をつけることで定式化できる対象は飛躍的に増加するが，解くことは逆に難しくなる。

★解説

■整数計画法による定式化の例 (1)研究開発計画を立てる際，提案された研究テーマがn件で，それらが成功した場合の予定収益，成功確率は知られているとする。そして資金，実験室スペース，研究員数などいくつかの資源利用について制約がある。このとき，収益額最大を狙って研究テーマの選定を行いたい。これは**ナップサック問題**とよばれる型の問題になる。(2)ある工場から製品を配送する地域とその中の配送先が定まっている。輸送はトラックによるが，1台のトラックは，何個所かの配送先を回りながら配送をする。うまくルートを作って，総走行距離最小にしたい。このとき配送先があるルートに含まれるか否かで1か0かをとる変数と，あるルートにトラックを走らせるか否かで1か0をとる変数を導入すると整数計画問題になる。この型の問題は**集合被覆問題**とよばれる。(3)n個の都市があり，都市間の移動に要する費用は与えられている。n個の都市すべてをちょうど一度ずつ巡って出発した都市に戻って来る道筋の中から，移動費用の総和を最少にするものを見つけたい。これは**巡回セールスマン問題**とよばれる。このほか，**割当問題**，**固定費つき輸送問題**，**ジョブショップ・スケジューリング**の問題なども整数計画の問題として知られている。

関連ページ
数理計画法　186
線形計画法　218

割当問題　420
輸送問題　392
日程計画　280

■整数条件はなぜ必要か　長さ5000㎜のアルミ棒からサッシの材料を切り取る。部品A，B，Cの長さはそれぞれ1800，900，650㎜であるとする。A，B，Cをそれぞれx，y，z本ずつ切り出すとすれば，

$$1800x+900y+650z\leqq 5000$$

でなければならないが，この変数はいずれも，0, 1, 2, …のいずれかである。

これは確かに整数条件である。しかし，問題の性格上，A，B，Cの総数は適当にバランスしていなければならない。そうすると，上の不等式を満たすx，y，zの組はかなり限られてきて，可能な組み合わせごとに目的関数の値を調べ，簡単に最適値が見つかる可能性がある。

一方，輸送問題(→392ページ)での輸送単価は恐らくトラック1台分で計算されているであろう。393ページの解法例示に用いた例のように，供給量も需要もともに整数であれば，最適解も整数値の組なので問題はないが，整数値でない場合には，輸送費が輸送量に比例するとは限らないのが実際であろう。しかしながら，供給量や需要がある程度大きければ，ことさら整数条件を入れることもあるまい。得られた値を基に適当にトラックの手配をすることは多分可能であろう。

どうしても整数条件が必要な場合の典型的な例は，左ページの諸例のように，ある研究テーマを採択するか否かとか，ある配送ルートを設定するか否かとか，ある場所に倉庫を設置するか否かというように，0-1整数条件が本質的に必要な場合である。

これらも本質的には，サッシ材料の例のように，可能な組み合わせについて調べればよいという種類のものになる。ただ，その組み合わせの数が多い場合には，整数計画法の計算手段に訴えることが大変有効になる。

■整数計画の解法　代数的方法，特に**切除平面法**とよばれる方法が理論的には大切であるが，実用性に乏しい。

分枝限定法は，基本的には，あらゆる可能な組み合わせを列挙して，それぞれの場合における目的関数の値を調べるという原始的な方法であるが，これを基に線形計画法を利用したりする方法が最も実用的である。現在存在している商用の数理計画ソフトウェアでは，ほとんどがこの方式に拠っているといわれている。分枝限定法とは，もとの問題に適当な条件をつけた，いくつかの子問題に分けるという**分枝操作**と，子問題の最適値の近似値や下限値を求め，分枝操作の手間を省く**限定操作**とを組み合わせたものである。限定操作のためには，整数条件を外した**連続緩和問題**や，不等式制約条件を外した**ラグランジュ緩和問題**が役に立つ。

整数計画法及びそれを含む組み合わせ最適化の実用的解法については〔24〕を参照。　　（森村）

積率

統計学において重要な概念の1つは，資料を要約し，特定化することである。そのために度数分布をつくって資料を要約し，特性値を求めたりするが，それらのほとんどを積率から得ることができる。積率とは次のような量である。

いま，x_i が f_i 個 ($i=1, 2, \cdots, k$) である $n(=f_1+f_2+\cdots+f_k)$ 個のデータが得られているとき，任意の定数 a を用いて算出された量

$$\frac{1}{n}\{(x_1-a)^r\cdot f_1+(x_2-a)^r\cdot f_2+\cdots+(x_k-a)^r\cdot f_k\}$$

のことを，a の回りの r 次の積率という。

★解説

関連ページ

■原点の回りの積率　上式で $a=0$ とおくと，**原点の回り**(単に，原点回りともいう)**の r 次の積率**

$$m_1{}'=\frac{1}{n}(x_1^r\cdot f_1+x_2^r\cdot f_2+\cdots+x_k^r\cdot f_k) \tag{1}$$

になる。とくに $r=0, 1, 2$ に対しては，

$m_0{}'=1$，$m_1{}'=$**平均値** $\overline{x}$，

平均値　344

$$m_2{}'=\frac{1}{n}(x_1^2\cdot f_1+x_2^2\cdot f_2+\cdots+x_k^2\cdot f_k)=2\text{ 乗の平均値}$$

となる。

■平均値の回りの積率　上の定義式で，$a=\overline{x}$ にとると，統計学でもっとも重要な**平均値の回り**(単に，平均回りともいう)**の r 次の積率**

$$m_r=\frac{1}{n}\{(x_1-\overline{x})^r\cdot f_1+(x_2-\overline{x})^r\cdot f_2+\cdots+(x_k-\overline{x})^r\cdot f_k\} \tag{2}$$

が得られる。とくに，$r=0, 1, 2$ に対しては，

$m_0=1$，$m_1=0$，$m_2=$**分散**

分散　332

になる。

■**歪度と尖度**　分布のひずみの度合い，いいかえれば左右非対称の度合いを表す量を歪度といい，とがりの度合いを表す量を尖度という。歪度や尖度は，式(2)の m_r を用いて，次のように定義されている。

$$\text{歪度}\ \alpha=\frac{m_3}{m_2}\qquad \text{尖度}\ \beta=\frac{m_4}{m_2^2}-3$$

分布型を，分布関数(→57ページ)として表現するのでなく，積率で表現できるかどうかを調べることは，たいせつである。一方，確率分布(→56ページ)の中で，中心的な役割をはたすのは正規分布(→200ページ)である。歪度や尖度は，そのような意味から，正規分布を念頭においた尺度である。すなわち，正規分布においては，上の α，β はいずれも0になる。歪度や尖度は，これらを規準にした量である。

ところで，歪度や尖度を求めるには，平均回りの3次の積率 m_3 や4次の積率 m_4 を計算しなければならない。これを直接計算するよりも，原点回りの積率 m_3'，m_4' を計算したほうがよい場合が多い。平均回りの r 次の積率 m_r は，原点回りの積率 m_r' を用いて，次のように表すことができる。

$$m_r=m_r'-\tbinom{r}{1}\overline{x}\cdot m'_{r-1}+\tbinom{r}{2}\overline{x}^2\cdot m'_{r-2}-\tbinom{r}{3}\overline{x}^3\cdot m'_{r-3}+\cdots\cdots$$

ただし，

$$\tbinom{r}{k}={}_r\mathrm{C}_k=\frac{r(r-1)(r-2)\cdots(r-k+1)}{k!}$$

(例) 5個の貨幣を100回投げて，各回で出た“表の数”

表の数	0	1	2	3	4	5	計
度数	5	11	29	35	15	5	100

表1　度数表

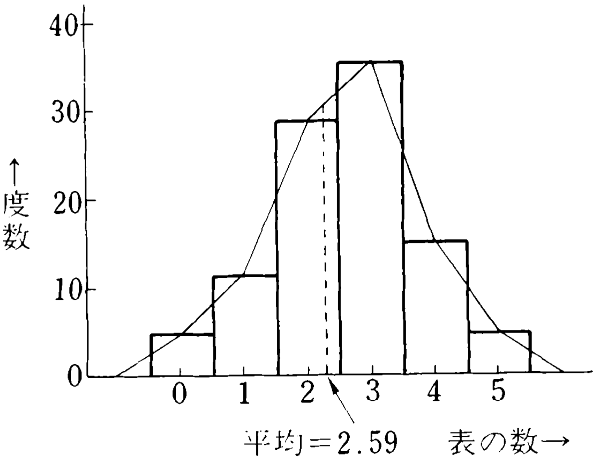


図1　各回で出た表の数の分布

である。

とくに，m_3，m_4 は次のようになる。

$$m_3=m_3'-3\overline{x}\cdot m_2'+2\cdot x^3$$

$$m_4=m_4'-4\overline{x}\cdot m_3'+6x^2\cdot m_2'-3\cdot x^4$$

上の例について，α，β を計算すると，次のようになる。

まず，

$\overline{x}=2.59$，$m_2=1.362$，

$m_3=-0.226$，$m_4=5.273$

となるから，

$$\text{歪度}\ \alpha=-0.226/1.362^{\frac{3}{2}}=-0.142$$

$$\text{尖度}\ \beta=(5.273/1.362^2)-3=-0.157$$

になる。上の表から折れ線グラフをつくると，図1のようになる。多少ひずんでいる様子が，図からもわかる。

（牧野）

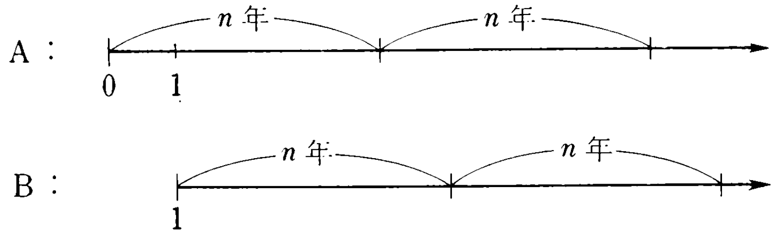
設備更新のタイミングの計算

現在まで使用している設備を，いつ新設備に更新するのが有利かを求めることが，設備更新の問題である。この問題では，現時点で新設備へ更新する案と，1年後に新設備へ更新する案とを比較することになる。

★解説

新設備へ，現時点で更新する案と1年後に更新する案を比較することは，新設備への更新を「1年間遅らせる方策」の有利さを調べることになる。更新を1年遅らせることが不利なら現時点で更新するのがよく，有利ならば更新を1年待つ方がよいことになる。

使用年数(寿命)がn年で，初期時点で表した費用がG円の新設備の使用を1年遅らせることを評価してみよう。



図に示すように，Aは現在から使う案，Bは1年後から使う案である。ここでは，設備を反復的に使用することを仮定している(→286ページ)。AとBの年平均費用(**年価**)はともに同じ値で，$M=G\times[P\to M]^i_n$ になる。2年目以後をみると，A，Bともに年当たりMの費用が共通にかかるとみなせるので，AとBの違いは1年目だけにあらわれ，Aのほうが1年後のMだけ余分に費用がかかることになる。したがって，1年遅らせることにより，1年後の価値でみてM円だけ節約される。

一方，現在使用している設備は更新を1年遅らせると，1年目に生じる費用だけ負担が増す。その額が1年後の価値でF円とすると，GとFの大きさの比較で更新を遅らせる方策の有利さがわかることになる。

関連ページ

設備の経済寿命
214

年価法の役立ち
286

▩新設備への更新の検討例　設備Aは現在使っている汎用機である。この先，長期的にみて専用機に設備を置きかえていく方針がある。そこで，この設備Aを専用機Bに更新することが検討の対象にのぼってきた。

汎用機の設備Aは，現在処分すると800万円で，1年後に処分すると500万円になると見込まれる。この1年間の操業費はおよそ2500万円かかる。他方の専用機Bは，年平均費用が3037万円と計算されている（215ページの経済寿命の計算例の設備をさしている）。

更新を遅らせると，現有の設備Aでは，現在処分すれば得られるはずの処分収入800万円の機会損失（→70ページ），1年後の500万円の処分収入，1年後の2500万円の操業費用が生じることになる。したがって，更新を1年遅らせることによる損失Fは，1年後の価値で表すと次式になる（$i=10\%$）。

$$F=800\times(1+0.1)-500+2500$$
$$=2880\ (\text{万円})$$

更新を遅らせることにより，新設備については$M=3037$万円の節減になる。したがって，節減額Mの方が負担増Fより大きいので，更新を1年遅らせた方が有利となる。

▩新設か改造かの選択問題　ゴミ処理設備の能力が不足してきたので，その対策として，現有設備を改造する案と新設備に置きかえる案を比較検討することになった。

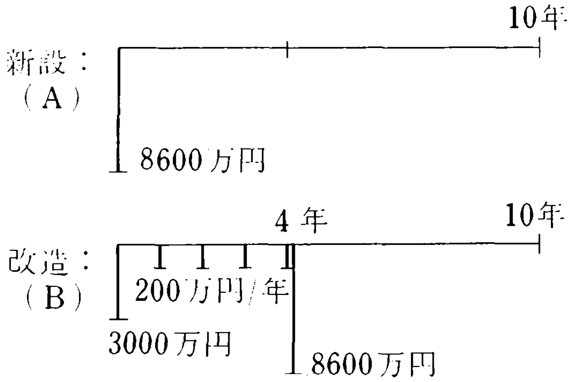


図1　新設と改造案のキャッシュ・フロー

新設備にすると，購入費，据付工事費など合わせて8600万円ほどかかる見込みである。設備の物理的寿命は15年以上は十分ある。一方，改造すると部品資材の購入費，工事費など全部で3000万円が見積もられている。

改造案は能力に限界があり，4年後にはいずれにしても新設備に置きかえることになりそうである。また，10年後には新設備も能力不足になることが予想できる。

改造設備，新設備のいずれも処分価額は，スクラップ同然となり，ゼロと見込まれる。また，改造設備は動力費が新設備よりも年当たり200万円余分にかかる。

新設備案と改造案のキャッシュ・フロー（→358ページ）を明らかにし，比較する（図1参照）。比較の年数は，問題状況から10年とすればよいことがわかる（$i=10\%$）。

$$P_B=8600\times[S\rightarrow P]_4^{10\%}+200\times[M\rightarrow P]_4^{10\%}+3000=9509\ (\text{万円})$$

この計算により，新設案が有利になる。（中村）

設備の経済寿命

設備の経済寿命とは，設備の取得価額と処分価額で決まる設備費と毎期の操業費を加えた総費用の1期当たりの平均費用が，最も小さくなるような使用期間をさすのが普通である。

★解説

通常の設備では，設備を稼動することによって発生する操業費は，使用期間(n)が延びるにつれて設備の性能が劣化したり，故障などのメンテナンス費がかさみ，増大するのが普通である。そこで，使用期間が長くなると，操業費の1期間当たりの平均負担額は増加する。

一方，設備の取得価額の1期間当たりの平均負担額は減少する。これらの2種類の平均負担額の変化は，図1のようになる。これらの和であらわされる総費用の平均負担額の動きは，図に示すように下に凸のグラフになる。この値を最小にする使用期間(n^*)が経済寿命になるわけである。

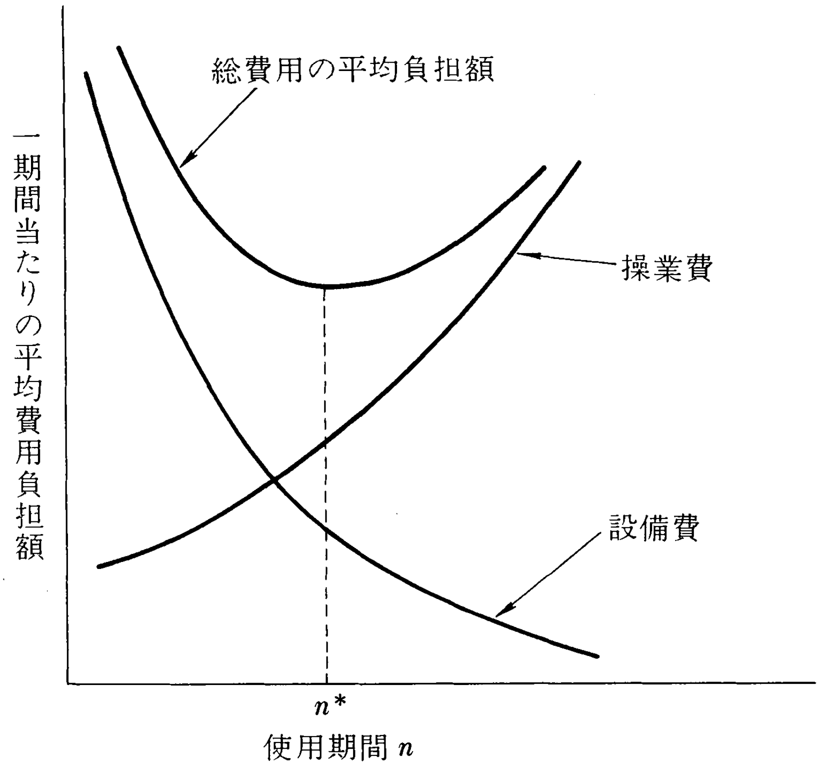


図1　1期間当たりの平均費用負担額の動き

関連ページ

年価法の役立ち　286

（万円）

①	②	③	④	⑤	⑥	⑦	⑧	⑨
使用年数	操業費用	現価係数（$i=10\%$）	操業費用の現価	操業費用の現価の和	資本回収係数	操業費用の年金換算値⑤×⑥	初期投資の年金換算値5 000万円×⑥	総費用の年金換算値⑦＋⑧
1	1 000	0.909 1	909	909	1.100	1 000	5 500	6 500
2	1 400	0.826 5	1 157	2 066	0.576 2	1 190	2 881	4 071
3	1 800	0.751 3	1 352	3 418	0.402 1	1 374	2 011	3 385
4	2 200	0.683 0	1 503	4 921	0.315 5	1 553	1 578	3 131
5	2 600	0.620 9	1 614	6 535	0.263 8	1 724	1 319	3 043
6	3 000	0.564 5	1 694	8 229	0.229 6	1 889	1 148	3 037
7	3 400	0.513 2	1 745	9 974	0.205 4	2 049	1 027	3 076

表１　経済寿命の計算

■経済寿命の計算方法　経済寿命をきめる問題は，設備を１年間使う案，２年間使う案，３年間使う案，……といったなかで最も有利な案を求めることを意味する。これは寿命の異なる案の比較になるので，年価法が役立つことになる(→286ページ)。

設備の取得価額を C_0, その設備を n 期間使用したときの処分価額を D_n, 毎期末の操業費を E_1, E_2, E_3, ……で表す。そこで，設備をおのおのの期間使った場合の総費用の年価(**期当たりの平均費用負担額**) M_n ($n=1, 2, 3,$ ……)は次のようにして求められる。

$$M_1=\left[C_0-\frac{L_1}{(1+i)}+\frac{E_1}{(1+i)}\right]\times[P\to M]_1^i$$

$$M_2=\left[C_0-\frac{L_2}{(1+i)^2}+\frac{E_1}{(1+i)}+\frac{E_2}{(1+i)^2}\right]\times[P\to M]_2^i$$

$$M_3=\left[C_0-\frac{L_3}{(1+i)^3}+\frac{E_1}{(1+i)}+\frac{E_2}{(1+i)^2}+\frac{E_3}{(1+i)^3}\right]\times[P\to M]_3^i$$

このように，順次 M_n の値を計算して，最も小さくなる使用期間 n を求めれば，それが経済寿命になる。

このことを数値例でみてみよう。取得価額が5000万円の設備があり，操業費は第１年度末が1000万円，そのあとは毎年400万円ずつ増加することが見込まれている。この設備は専用機なので，処分しても価値は低く，取りはずしの費用を考えれば処分価額はほとんどゼロとみなせる。$i=10\%$ のもとで，経済寿命を計算してみよう。

経済寿命の計算は，上に示した表１で行うとよい。使用年数を①列に，操業費用を②列に，各年数のもとでの現価係数を③列，資本回収係数を⑥列に記入する。以下は表にあるように計算し，年価の値が⑨列に求められる。この結果，経済寿命は６年で，そのときの年平均費用は約3037万円になる。　（中村）

ゼロ和2人ゲーム

ゲーム理論で扱われるゲームの中で，最も簡単かつ基礎的なものである。2人のプレーヤーで争われ，一方の利得が他方の損失となる形のゲームである。両者がともに合理的な行動，すなわち相手も十分に利口であると考えて，自己の損失を最小限にする手をとるならば，混合戦略(下記参照)の範囲内で両者のとるべき戦略は定まる。

★解説

■**ミニマックス原理**　上で**合理的**と称した行動は，一方のプレーヤーが，相手がどんな手をとったとしても，最悪でも確保しうる利得を最大にするような手をとるということを指している。これは，利得行列の各行(自分のとる手によって定まる)の最小値(相手の手によって定まる最悪の値)のうちの最大のものを見つけ，それがある行を選ぶことに当たる。このような行動原理は，ミニマム(最小値)をマックスに(最大化)するということから，**ミニマックス原理**とよばれている。

■**鞍点**　右ページの図1に示す行列で，○印をつけた位置にある数1は，その行においては最小値であるが，その列においては最大値である。このような行列の要素を**鞍点**という。馬の鞍は左右に対しては最高，前後に関しては最低の位置に乗るように作られているから，それにちなんでつけられた名まえである。

■**純戦略と混合戦略**　プレーヤーがとる手の1つ1つを**純戦略**という。利得行列に鞍点のあるようなゼロ和2人ゲームでは，両者は鞍点に対応する純戦略をとろうとするが，ふつうはいくつかの純戦略をある割合で選ぶことで，負けないように努める(次ページ参照)。そのとき，各純戦略をとる確率分布を**混合戦略**，もしくは単に**戦略**とよぶ。

関連ページ

ゲーム理論　102

■最適戦略　右の図1を利得行列とするゼロ和2人ゲームでは，プレーヤー1が最大の利得4を得ようとして手$\boxed{2}$をとると，プレーヤー2は①をとるであろう。その結果プレーヤー1の利得は −5 になって大損をしてしまう。それ故，彼は最大利得4はあきらめて$\boxed{3}$をとる。相手も損失は小さくしたいから②をとるであろうから，$\boxed{3}$の中では最小の利得1しか得られないがそれでもプラスである。もし，$\boxed{1}$をとると相手が②か③をとったとき，利得はマイナスになる。結局，相手が合理的に行動する限り，プレーヤー1は$\boxed{3}$をとることで，最悪の場合でも利得1を確保するのが合理的である。

$$\begin{array}{c|c}
 & \begin{array}{ccc} ① & ② & ③ \end{array} \\
\begin{array}{c} \boxed{1} \\ \boxed{2} \\ \boxed{3} \end{array} & \left(\begin{array}{ccc} 2 & -2 & -3 \\ -5 & 0 & 4 \\ 3 & ① & 2 \end{array}\right)
\end{array}$$

図1

他方，プレーヤー2にとっても，①をとれば3，③をとれば4を失うかもしれないのに，②をとれば最悪の場合でも1ですむから，②をとるのが合理的である。つまり，両者が合理的な行動をする限り，それぞれ$\boxed{3}$，②の手をとるしかない。この例では，各プレーヤーのとるべき最適の手が1つずつに定まってしまうので最適戦略は純戦略の範囲で得られた。このときの利得1を**ゲームの値**とよび，最適戦略とゲームの値を求めることを**ゲームを解く**という。

■混合戦略の必要性　図1の利得行列には鞍点があったが，図2のそれには鞍点がない。このとき，もしプレーヤー1が$\boxed{2}$をとると，相手は①をとろうとし，それなら$\boxed{3}$をとる方が得で，そうすると相手は②になる。したがってプレーヤー1は$\boxed{2}$をとるべきである，というように，2人のプレーヤーのとるべき手がぐるぐる回ってしまって落ち着くところがない。いいかえると，相手の裏をかこうとすると，そのまた裏をかき，…となってとめどがないのである。現実にその例は多い。

$$\begin{array}{c|c}
 & \begin{array}{ccc} ① & ② & ③ \end{array} \\
\begin{array}{c} \boxed{1} \\ \boxed{2} \\ \boxed{3} \end{array} & \left(\begin{array}{ccc} 2 & -2 & -3 \\ -5 & 1 & 4 \\ 3 & -1 & 2 \end{array}\right)
\end{array}$$

図2

こういうとき，各プレーヤーはもし1つの手をとり続けたとすると，相手にそれと察知されて裏をかかれ，確実に悪い結果になるので，いろいろな手をある確率でとるよう混ぜ合わせ，平均すれば最善の手というものを考えるべきである。たとえば，プレーヤー1が$\boxed{1}$，$\boxed{2}$，$\boxed{3}$をそれぞれ (0，1/2，1/2) でとり，相手が①，②，③をそれぞれ (1/2，1/2，0) でとったとすると，プレーヤー1の期待利得は $(-5+3-1+1)\times(1/4)=-1/2$ である。

このように，相手の戦略を考慮して，最悪の場合の期待利得を最大にするような戦略が最適戦略であり，そのときの期待利得がゲームの値となる。混合戦略の枠内なら最適戦略は必ずある。

（森村）

線形計画法

いくつかの1次式で表される制約式を満たし，かつ，1次式で表される目的関数の値が最小(または最大)となる変数の組を求める手法で，数理計画法ひいてはオペレーションズ・リサーチの手法の中心をなすものである。生産計画などの問題には，この形で定式化できるものが多く，解法も確立していることから，広い範囲で応用されている。

英語のLinear Programming を略して LP ともいう。

★解説

ある工場で2種類の製品A，Bを作っていて，月間生産計画を立てたいとしよう。A，Bの生産量をx_1，x_2(キロ)とおき，Aを1キロ生産するのに電力が3キロワット時，Bを1キロ生産するのに5キロワット時必要とすれば，全部で電力は $3x_1+5x_2$ キロワット時必要になる。もし，この工場では毎時15キロワットまでしか電力を使用できないとしたら，

$$3x_1+5x_2 \leqq 15 \qquad (1)$$

という制約式を満たさなければならない。

このほか，原料その他いくつかの要因についても(1)同様の**制約式**が作られよう。これらがすべて1次不等式であることも実際に多いから，上述の「いくつかの1次式で表される制約式を満たし」という条件が満足される場合はよく起こる。また，この例で，製品Aはキロ当たり5万円，製品Bは6万円の利益があるものと仮定すると，総利益は，

$$5x_1+6x_2 \text{（万円）} \qquad (2)$$

と表される。利益をできるだけ生み出すことが生産の目標であるとするならば，(2)で表される1次式を目的関数とすればよいであろう。こうして，上述のような線形計画の枠組みの中でこの生産計画は定式化されたことになる。

関連ページ
制約式　186
目的関数　186
線形計画の標準問題　234

■線形計画問題の一例　左ページであげた例において，原料について更に次のような制約条件があるものとしよう。

原料 I については，1キロの製品を作るため，Aでは4キロ，Bでは3キロが必要で，全体で毎時12キロしか用意できない。また，原料 II については製品1キロあたり，A，Bとも0.2キロが必要で，毎時0.64キロの使用が可能である。

■制約条件　このような状況が設定されれば，左ページの(1)も含め

$$\left.\begin{array}{l}3x_1+5x_2\leqq 15\\4x_1+3x_2\leqq 12\\0.2x_1+0.2x_2\leqq 0.64\end{array}\right\}\qquad(3)$$

$$x_1\geqq 0,\ \ x_2\geqq 0\qquad(4)$$

という制約条件を満たさなければならない。条件(4)は生産量はマイナスではありえないということから当然出てくるが，このような条件は**非負条件**とよばれている。

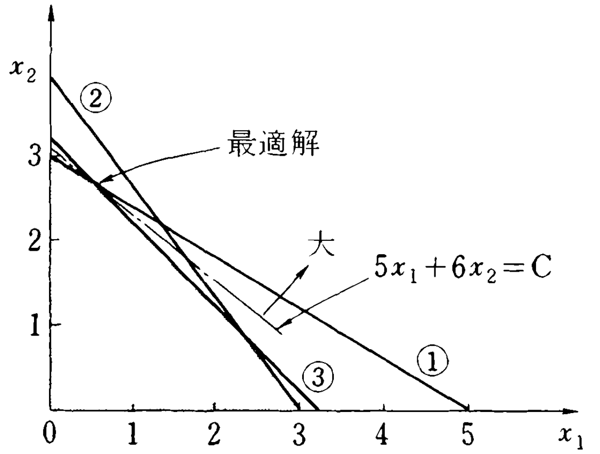
■定式化　左のページに示したような利益を示す目的関数をとるならば，ここで例示した生産計画の問題は，

「条件(3)，(4)のもとで，目的関数(2)を最大にせよ」

という線形計画の問題として定式化されたことになる。

実際問題では製品の種類も制約となる要因もはるかに多いので，変数の数が何千，制約式の数が何百といった大規模な線形計画問題も決して珍しくない。

いま例にあげた問題では変数が2個



であったので，制約条件(3)(4)は上図のように図示することができる。

(3)には3つの1次不等式があるが，それぞれの境界つまり等号で示される1次式を図示すると，上図の①②③で示す直線が得られる。

不等号の向きを考えると，(3)の各不等式を満たす点 $(x_1,\ x_2)$ の存在する範囲は各直線の下側(左側)になる。したがって，3つの不等式および非負条件(4)を同時に満たす点 $(x_1,\ x_2)$ は図の影の部分にあることがわかる。それで，この領域を**許容領域**とよぶ。この線形計画問題の解として許されるという意味である。

目的関数(2)も1次式であるため，これをCに等しいとおくと，図で鎖線で示すような傾きの直線が得られる。Cをいろいろに変えると，これに平行な直線が得られ，Cを増やすと矢印の方向へ移動する。したがって，許容領域の中にあってCを最大にするのは，図中「最適解」と記した点(0.5, 2.7)で，これを**線形計画問題の解**という。参考文献[50]　　(森村)

線形(多重)ロジスティックモデル

線形ロジスティックモデルは，外的基準が2値のデータをp個の説明変数x_1, …, x_pで回帰する。表面上は回帰分析であるが，判別分析として用いられることも多い。線形判別関数は2群が多次元正規分布に従う確率密度関数という強い仮定があるが，線形ロジスティックモデルは，2群の確率密度の比の対数をとったもの(対数尤度)が説明変数の線形和で表されるだけでよい。このため，2群が多次元正規分布から乖離(かいり)しているようなデータでは，線形判別関数を用いるよりは線形ロジスティックモデルを用いる方がよい。たとえば，説明変数も2値変数であったり，2群のデータ数が極端に異なる疫学データに多く用いられる。

★解説

▩モデル　電子部品の故障の生起や，疾病の発症の有無等の2値をとる特性Y_iは，1と0の値をとる確率変数と考えることができる。

$$\mathrm{prob}\ (Y_i=1)=\theta_i,\ \mathrm{prob}\ (Y_i=0)=1-\theta_i \quad (i=1,\ \cdots,\ n) \qquad (1)$$

θ_iはi番目の個体が$Y_i=1$になる確率を表す。確率θ_iをロジスティック変換した$\lambda_i=\log\{\theta_i/(1-\theta_i)\}$が，個体$i$の$p$個の説明変数値$(X_{1i},\ \cdots,\ X_{pi})$の線形モデルで表されるとき，これを線形ロジスティックモデルという。

$$\lambda_i=\log\{\theta_i/(1-\theta_i)\}=\beta_0+\beta_1x_{1i}+\cdots+\beta_p x_{pi}=\boldsymbol{\beta}'\boldsymbol{x} \quad (i=1,\ \cdots,\ n) \qquad (2)$$

これを変形して次の確率θ_iと$(1-\theta_i)$が求まる。

$$\theta_i=\frac{\exp(\boldsymbol{\beta}'\boldsymbol{x})}{1+\exp(\boldsymbol{\beta}'\boldsymbol{x})},\quad 1-\theta_i=\frac{1}{1+\exp(\boldsymbol{\beta}'\boldsymbol{x})} \qquad (3)$$

式(2)は確率θ_iの対数尤度を目的変数とした回帰分析であるが，θ_iが0.5より大きければ$Y_i=1$と判別できる。

関連ページ
質的データの解析　162
数量化Ⅱ類　190
FUNCATによる判別　320

■アルゴリズム　n 個の観測対象で，2値特性 y_i と p 個の説明変数の値が測定されている。左ページの線形ロジスティックモデルの β_i は，i 番目の説明変数 x_i のロジスティック回帰係数である。この値は次の尤度を最大にするように，ニュートン・ラプソン法を用いて反復計算される。

$$\prod_{i=1}^{n}\left\{\frac{\exp(\boldsymbol{\beta}'\boldsymbol{x})}{1+\exp(\boldsymbol{\beta}'\boldsymbol{x})}\right\}^{y_i}\left\{\frac{1}{1+\exp(\boldsymbol{\beta}'\boldsymbol{x})}\right\}^{1-y_i}$$

$$=\prod_{i=1}^{n}\frac{\exp(\boldsymbol{\beta}'\boldsymbol{x}y_i)}{\{1+\exp(\boldsymbol{\beta}'\boldsymbol{x})\}}$$

$$=\frac{\exp(\sum_{i=1}^{n}\boldsymbol{\beta}'\boldsymbol{x}y_i)}{\prod_{i=1}^{n}\{1+\exp(\boldsymbol{\beta}'\boldsymbol{x})\}}\longrightarrow \max \qquad (4)$$

■判別分析　2群 G_i $(i=0,\ 1)$ で，説明変数ベクトル $\boldsymbol{x}$ の確率密度関数を $f_i(\boldsymbol{x})$ とし，2群の事前確率を π_i とする。すなわち，

$$P(G_i)=\pi_i,\quad P(\boldsymbol{x}\,|\,G_i)=f_i(\boldsymbol{x}) \quad (5)$$
$$(i=0,\ 1)$$

ここで観測値 $\boldsymbol{x}$ が与えられたとき，それが G_1 に属する事後確率 $P(G_1|\boldsymbol{x})$ はベイズの定理を用いて式(6)になる。

$$P(G_1|\boldsymbol{x})=\frac{P(G_1)P(\boldsymbol{x}|G_1)}{P(G_0)P(\boldsymbol{x}|G_0)+P(G_1)P(\boldsymbol{x}|G_1)}$$

$$=\frac{\pi_1 f_1(\boldsymbol{x})}{\pi_0 f_0(\boldsymbol{x})+\pi_1 f_1(\boldsymbol{x})}$$

$$=\frac{(\pi_1/\pi_0)(f_1(\boldsymbol{x})/f_0(\boldsymbol{x}))}{1+(\pi_1/\pi_0)(f_1(\boldsymbol{x})/f_0(\boldsymbol{x}))} \qquad (6)$$

ここで，$\log(f_1(\boldsymbol{x})/f_0(\boldsymbol{x}))$ が $\boldsymbol{x}$ の線形結合で表されるものとする（式(7)）。

$$f_1(\boldsymbol{x})/f_0(\boldsymbol{x})=\exp(\boldsymbol{\beta}'\boldsymbol{x}),$$
$$\exp(\beta_0)=\pi_1/\pi_0 \qquad (7)$$

これを式(6)に代入して式(8)が求まる。

$$P(G_1|\boldsymbol{x})=\frac{\exp(\beta_0+\boldsymbol{\beta}'\boldsymbol{x})}{1+\exp(\beta_0+\boldsymbol{\beta}'\boldsymbol{x})} \quad (8)$$

同様にして，

$$P(G_0|\boldsymbol{x})=\frac{1}{1+\exp(\beta_0+\boldsymbol{\beta}'\boldsymbol{x})} \quad (9)$$

になり，従って

$$\log\left\{\frac{P(G_1|\boldsymbol{x})}{P(G_0|\boldsymbol{x})}\right\}=\beta_0+\boldsymbol{\beta}'\boldsymbol{x} \qquad (10)$$

になる。すなわち，i 番目の観測値 $\boldsymbol{x}_i$ に対して，

$$\theta_i=\frac{P(G_1|\boldsymbol{x}_i)}{P(G_1|\boldsymbol{x}_i)+P(G_0|\boldsymbol{x}_i)} \qquad (11)$$

とおけば，式(10)は式(2)を表す。

以上，ロジスティック判別では式(7)に示すように2群の確率密度関数の比の対数が，$\boldsymbol{x}$ について線形であればよい。このため，2群が分散の等しい多次元正規分布に従うという強い仮定から導かれた通常の線形判別関数よりは，広い範囲のデータに適用できる。疫学データのように，正常例が極めて多く，疾病例の少ない場合や，説明変数に質的データが混在する場合には，数量化II類や線形判別分析を用いるよりはよい結果が得られる。初期値を線形判別関数とし，ニュートン・ラプソン法でロジスティック回帰係数を求めるプログラムが多重ロジスティックとよばれ，医学で多用されている。（新村）

相関関係

1回の実験や観測で，2個の属性の値X，Yが

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots\cdots,\ (x_n,\ y_n)$$

というように，2つの数の組として得られているとき，XとYとの関係の強さを記述するのに，散布図や相関係数が用いられる。

★解説

変量X，Yについて，上のようなn組のデータが得られたとき，関係の強さを手軽に判断するのに**散布図**が用いられる。これは，横軸をx軸，縦軸をy軸として，方眼紙の上に，それぞれの点

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots,\ (x_n,\ y_n)$$

を打点したものであって，散布図が図1のようになれば，**正の相関**があるという。この場合，Xの値の大きなものは，Yの値も大きいという傾向が強い。これに対し，図2のような場合には，**負の相関**があるという。この場合には，Xの高いものほど，Yが低くなるという関係がみられる。ところが，図3のようなときには，Xが増加すればYも増加するとか，または逆にYは減少する，といった特別の関係はない。このようなとき，XとYとは**無相関**であるという。

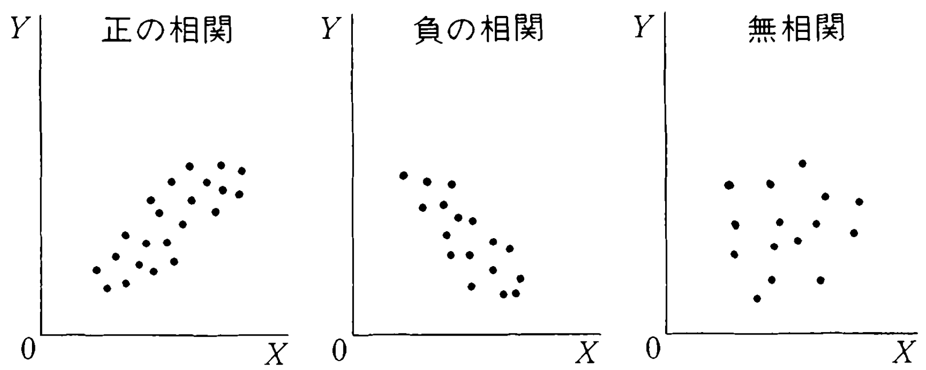


図1　散布図(1)　　図2　散布図(2)　　図3　散布図(3)

上のような関係の強さを数量的に記述するのに，**相関係数**が用いられる。

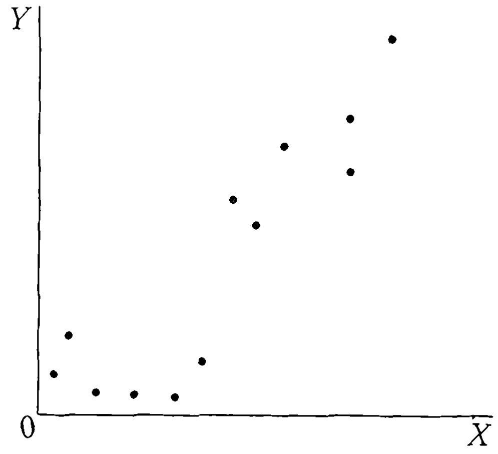
関連ページ

方眼紙　356

（例）　2つの銘柄A，Bの株価について，最近12日間の終値を調べたところ，次表のようになった。

	Aの終値	Bの終値
第1日	105円	68円
2日	102	71
3日	100	69
4日	108	66
5日	112	65
6日	115	70
7日	118	75
8日	116	76
9日	120	78
10日	125	77
11日	125	79
12日	128	82

Aの終値をX，Bの終値をYとするとき，散布図は次のようになる。



上図から，XとYとの間に正の相関があることがわかるが，これを数量的に把握するために，相関係数rを計算してみると，$r=0.86$になる。

■相関係数の計算

一般にn組のデータ

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots,\ (x_n,\ y_n)$$

があったとき，XとYとの間の関係の強さを表す尺度である相関係数rは，次式を用いて算出される。

$$r=\frac{(X\text{と}Y\text{の共分散})}{\sqrt{(X\text{の分散})}\sqrt{(Y\text{の分散})}}$$

ただし，

$$X\text{の分散}=\frac{1}{n}\{(x_1-\overline{x})^2+\cdots+(x_n-\overline{x})^2\}$$

（$\overline{x}$はXの平均）

$$Y\text{の分散}=\frac{1}{n}\{(y_1-\overline{y})^2+\cdots+(y_n-\overline{y})^2\}$$

（$\overline{y}$はYの平均）

XとYの共分散

$$=\frac{1}{n}\{(x_1-\overline{x})(y_1-\overline{y})+\cdots+(x_n-\overline{x})(y_6-\overline{y})\}$$

である。

相関係数rの値は，いつも$-1\leqq r<1$となるものであって，XとYとの増加・減少が完全に一致していれば$r=1$，一致の傾向が弱まれば次第に小さな値になり，無相関のときは0の近くの値をとり，完全に逆の関係のとき，$r=-1$になる。

なお，上の例で，第1日はAが105円であるのに対し，Bは71円というように，すべて1日ずらしてみると，$r=0.90$となり，2日ずらすと$r=0.99$，3日で，$r=0.91$になる。このことからBの株価を予想するのに，2日前のAの株価が役立つことがわかる。　（牧野）

ダービン・ワトソン比

重回帰モデルの残差系列 ε_1, ε_2, …, ε_N に,

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

という関係が存在するか否かを調べるのに，ダービン・ワトソン比は用いられる。これは ε_t とその 1 つ前の ε_{t-1} との間に関連性が存在するか否かをみていることになる。ここで，u_t は互いに独立に平均 0，分散一定の正規分布に従うとしている。

★解説

実測値 y と回帰による推定値 $\hat{y}$ との差

$$e_t = y_t - \hat{y}_t \qquad (t=1,\ 2,\ \cdots,\ N)$$

により，ダービン・ワトソン比 d は次のように定義される。

$$d = \frac{(e_2-e_1)^2+\cdots+(e_N-e_{N-1})^2}{e_2^2+\cdots+e_N^2}$$

パラメータ ρ は,

$$\hat{r} = \frac{e_2e_1+e_3e_2+\cdots\cdots+e_Ne_{N-1}}{e_2^2+e_3^2+\cdots\cdots+e_N^2}$$

で推定する。ρ の推定値 r が 0 から大きく離れていれば，ρ は 0 ではないとし，上記モデルの 1 階の従属構造の存在を考えることになる。d と r との間には近似的に,

$$d \fallingdotseq 2(1-r)$$

が成り立つから, ダービン・ワトソン比 d では, 2 より大きく離れていれば，残差系列の中に上記のような構造の存在を認めることになる。d が 2 に近ければ, そのような関係はないと考える。判定基準はダービン・ワトソンの数表による。

関連ページ

回帰からの偏差 e_t は，残差分散のところで説明したドルベース輸出価格指数では，

$e_1 = 2.12$	$e_{11} = 4.63$	$e_{21} = -4.55$
$e_2 = 2.41$	$e_{12} = 4.71$	$e_{22} = -4.16$
$e_3 = 4.90$	$e_{13} = 4.16$	$e_{23} = -4.01$
$e_4 = 2.11$	$e_{14} = 3.35$	$e_{24} = -5.21$
$e_5 = -2.44$	$e_{15} = 2.21$	$e_{25} = -4.59$
$e_6 = -4.52$	$e_{16} = 1.36$	$e_{26} = 2.59$
$e_7 = -2.51$	$e_{17} = -2.64$	$e_{27} = -3.57$
$e_8 = -2.27$	$e_{18} = -4.29$	$e_{28} = -2.31$
$e_9 = -3.51$	$e_{19} = -5.54$	$e_{29} = 1.95$
$e_{10} = 1.26$	$e_{20} = -2.18$	$e_{30} = 5.39$
		$e_{31} = 5.78$
		$e_{32} = 5.42$
		$e_{33} = 4.95$
		$e_{34} = -1.00$

である。e_1, e_2, …, e_m が観測されたとき，次に観測される数値 e_{m+1} がどのような値であるか，それ以前の数値 e_j の系列から予測できないことが必要である。もし予測できたとすれば，それは残差系列の中に何らかの規則性があることを意味している。残差系列をプロットしたのが図1であり，この図をよく見ると，1つ前の値が正であれば次に続く値も正，1つ前の値が負であれば次に続く値も負の傾向があるのに気づく。このような場合は残差系列の中に，

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \qquad (1)$$

という構造が存在すると考えられる。ダービン・ワトソン比は，

$$r = \frac{2.41\times2.21+\cdots+(-1.00)\times4.95}{2.41^2+\cdots+(-1.00)^2}$$

$$=0.58$$

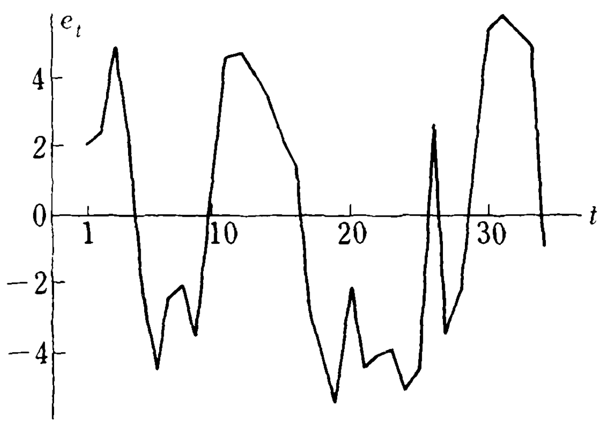


図1　残差 e_t の系列

である。これが2.0から大きく離れていて，残差系列の中に式(1)のような規則性が存在しないという仮説を棄却するか否かは，ダービン・ワトソン比の表による。この場合，説明変数の個数 $p=2$ であり，標本数 $N=34$ であるから，ダービン・ワトソン比の数表より

$$d_L = 1.13 \qquad d_U = 1.36$$

を得る。判定は，

(1) $d < d_L$ のときは，規則性なしという仮説は棄却される。

(2) $d > d_U$ のときは，規則性なしという仮説は棄却されない。

(3) $d_L < d < d_U$ のときは不定。

とする。いまの場合 $0.58 < d_L$ であり，残差系列の中に式(1)のような構造が存在することになる。これは重要な説明変数が欠落していて，このような結果になる場合もあるから，その点の再検討も意味のあることである。d が2.0より大きいときは，$4-d$ について上記と同じ手順で調べる。　(杉山)

大数の法則と中心極限定理

独立で同一分布に従う確率変数をたくさんとって平均(標本平均という)すると，それはその分布の期待値に近い値をとる。これを**大数の法則**という。また，この標本平均と期待値の差は(適当に拡大すると)だいたい正規分布に従っている。これを**中心極限定理**という。

★解説

関連ページ
独立　229
期待値　58
分散　332
母集団　260

いま，確率変数の列 X_1, X_2, …, X_n が互いに独立で，期待値 μ，分散 σ^2 をもつ同一の分布に従うものと仮定する(統計では，“母平均 μ，母分散 σ^2 のある母集団から，大きさ n の標本 X_1, X_2, …, X_n を無作為に抽出する”という)。このとき，次の2つの定理が成り立つ。

■大数の法則　上の仮定の下で，n 個の標本の平均

$$\overline{X_n}=(X_1+X_2+\cdots+X_n)/n$$

は，n を大きくすると，必ず(“確率1で”ともいう)期待値 μ に収束する。この事実を大数の法則(正確には強法則)とよぶ。

■中心極限定理　上の仮定の下で，

$$Z_n=(\overline{X}_n-\mu)/\sqrt{\sigma^2/n}=\sqrt{n}(\overline{X}_n-\mu)/\sigma$$

とおくと，n が大きいとき，Z_n の分布は正規分布 N(0, 1) に近づく。すなわち，

正規分布　200

$$\mathrm{P}\{Z_n\leqq x\}\to\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du \qquad (n\to\infty)$$

となる。これを中心極限定理とよぶ。

Z_n は標本平均 $\overline{X_n}$ と期待値 μ との差を $\sqrt{n}$ 倍に拡大して見ていることに注意する。また，Z_n は期待値が0，分散が1となるよう標準化してある。中心極限定理は確率論の“中心”であり由来は古い。ド・モアブル(De Moivre, 1667－1754)やラプラス(Laplace, 1749－1829)の二項分布の正規分布による近似に端を発し，リヤプノフ(Lyapunov,1857－1917)，リンデベルク(Lindeberg)らにより近代的な形に整理された。

標準化　201

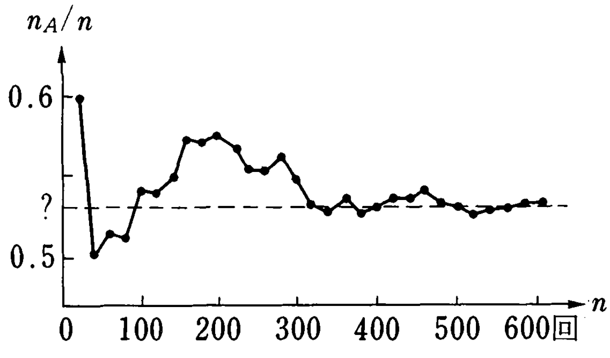
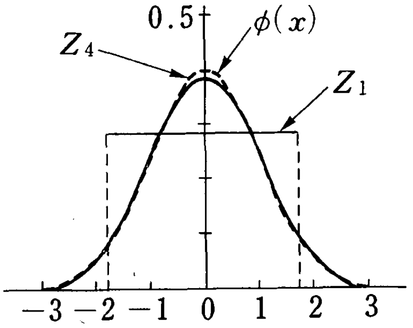
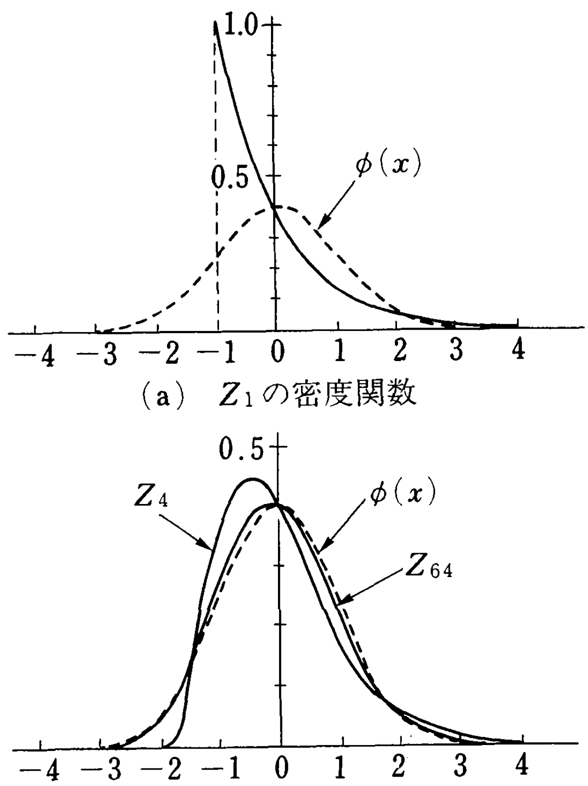


図1　メダル投げで表の出た相対度数のグラフ



Z_1，Z_4 の密度関数

図2　一様分布する確率変数の和の分布の正規分布への近づき方



（a）　Z_1 の密度関数

（b）　Z_4，Z_{64} の密度関数

図3　指数分布する確率変数の和の分布の正規分布への近づき方

■大数の法則と確率の概念　“同じ条件の下で試行が繰り返し行われるとき”，ある事象 A の生起する回数の相対頻度の極限をもって，事象 A の確率 P_A と考える。つまり，n 回の試行の中，A の生じた回数を n_A とすると，

$$n_A/n \to P_A \qquad (n\to\infty)$$

と考える。このような確率の解釈を確率の頻度概念といい，フォン・ミーゼス（Von Mises, 1882～1925）により唱えられた。

大数の法則は，上の条件“…”が成り立つ状況では，この解釈の妥当性を保証している。つまり，各 X_i を

$$X_i=\begin{cases}1 & (A\text{が起きたとき})\\ 0 & (\text{それ以外のとき})\end{cases}$$

とおけばよく，$\overline{X}_n=n_A/n$ であり，$\mu=E(X_i)=P_A$ である。

■中心極限定理の利用上の注意　図2，図3はそれぞれ，X_i の分布が一様分布（→411ページ），指数分布（→156ページ）に従うときの Z_n の分布と $N(0,\ 1)$ とを比べたものである。その正規分布への近づき方は一様分布のほうが指数分布の場合より速い。どのくらいの n で正規分布に近いと考えてよいかは，分布の形に依存する。

203ページには共通一次の得点分布と正規分布とを比べてある。共通一次の得点は数多くの問題の合計点だから中心極限定理があてはまりそうだが，そうはなっていない。それは，各問の成績の間に強い相関があり，定理の条件を満たさないためである。　（森）

多変数の分布

雪の多い年は豊作になるといわれている。この説の真偽を確かめるには，冬の降雪量とその秋の米の収穫量との関係を統計的に調べる必要がある。このように2つ以上の変数を同時に考えて，その確率分布を知りたいことが多い。このような多変数の分布の代表的なものとして離散的な場合の**多項分布**と連続的な場合の**多次元正規分布**がある。

★解説

まず2変数の場合の確率分布の表し方を述べる。

■離散的な場合　2つの確立変数X，Yがそれぞれ離散値$\{x_1, x_2, \cdots\}$, $\{y_1, y_2, \cdots\}$をとるものとする。このとき，

$$p(x_i, y_j)=\mathrm{P}\{X=x_i, Y=y_j\} \qquad (i, j=1, 2, \cdots)$$

を2変数(X, Y)の**同時分布**(あるいは結合分布)とよぶ。表1はある学校での算数の成績Xと国語の成績Y(5段階評価)の同時分布を表したものである。確率であるからその値は非負で，総和は1である。また，

$$p_1(x_i)=p(x_i, y_1)+p(x_i, y_2)+p(x_i, y_3)+\cdots$$

をXの**周辺分布**とよぶ。表1に見るとおり，この値は表の周辺に書き表され，Xそのものの分布，つまり$p_1(x_i)=P\{X=x_i\}$である。Yの周辺分布$p_2(y_i)$も同様に定まる。

■連続的な場合　X，Yが連続的な場合，その分布は

$$\mathrm{P}\{X\leqq x, Y\leqq y\}=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u, v)dudv$$

と表現される。$f(x, y)$は2変数(X, Y)の**同時密度関数**とよばれ，たとえば図2のような形をしている。(X, Y)が図2の微小領域に入る確率が$f(x, y)\Delta\cdot\Delta'$と考えてよい。この総和，つまり$f(x, y)$の曲面と$(x, y)$平面で囲まれる空間の体積は1である。上の式の左辺，$F(x, y)=\mathrm{P}\{X\leqq x, Y\leqq y\}$を$(X, Y)$の**同時分布関数**とよぶ。また離散的な場合と同様，

$$f_1(x)=\int_{-\infty}^{\infty}f(x, v)dv,\ F_1(x)=\int_{-\infty}^{x}f_1(u)du=F(x, \infty)$$

をそれぞれXの**周辺密度関数**，**周辺分布関数**とよぶ。

関連ページ
確率変数　56
確率分布　56
密度関数　56
分布関数　57

■ n 変数の場合　2変数のときと同じように，n 変数 $(X_1, X_2, \cdots, X_n)$ の分布はその同時分布関数

$$F(x_1, x_2, \cdots, x_n)=\mathrm{P}\{X_1 \leqq x_1, \cdots, X_n \leqq x_n\}$$

を与えることによって定まる。連続的な場合には，その同時密度関数

$$f(x_1, x_2, \cdots, x_n)$$

が定まる。

■確率変数の独立性　X，Y が離散的な場合は，すべての x_i，y_j に対して，

$$p(x_i, y_j)=p_1(x_i)\cdot p_2(y_j)$$

が成り立つとき，X と Y は互いに**独立**であるという。連続的な場合は，任意の x，y に対して，

$$f(x, y)=f_1(x)\cdot f_2(y)$$

が成り立つとき，X と Y は互いに独立であるという。さらに，n 変数 X_1，X_2，…，X_n については，任意の x_1，x_2，…，x_n に対して，

$$f(x_1, x_2, \cdots, x_n)=f_1(x_1)\cdot f_2(x_2)\cdots f_n(x_n)$$

が成り立つとき，互いに**独立**であるという。ただし，$f_i(x_i)$ は X_i の周辺密度関数である。

■条件付分布　X，Y が離散的な場合，$p_2(y_j) \neq 0$ のとき，

$$p_{1|2}(x_i | y_j)=p(x_i, y_j)/p_2(y_j)$$

を，$Y=y_j$ という条件の下での，X の**条件付分布**という。$p_{2|1}(y_j | x_i)$ も同様に定義される。連続的な場合は，$f_2(y_i) \neq 0$ なる y に対して，

$$f_{1|2}(x | y)=f(x, y)/f_2(y)$$

を，$Y=y$ という条件の下での，X の**条件付密度関数**という。これに関する期待値を条件付期待値という。

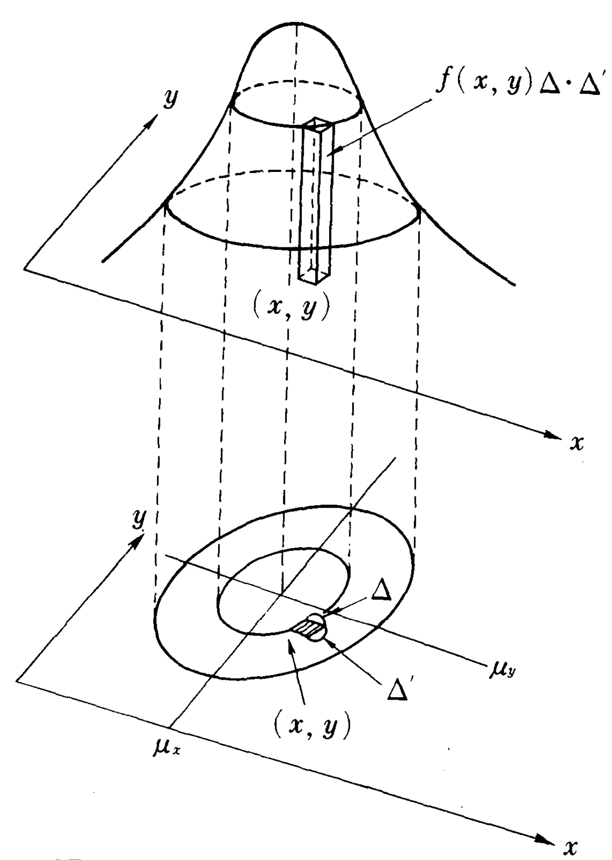


図1　同時密度関数と等高線

■共分散と相関係数　X，Y の期待値をそれぞれ μ_1，μ_2 とするとき，

$$\mathrm{Cov}(X, Y)=E((X-\mu_1)(Y-\mu_2))$$

を (X, Y) の**共分散**という。X，Y が連続的な場合，これは，

$$\mathrm{Cov}(X, Y) =\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dxdy$$

と計算される。和 $X+Y$ の分散は，

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y) +2\mathrm{Cov}(X, Y)$$

となる。もし，X，Y が独立ならば，

$\mathrm{Cov}(X, Y)=0$ となり，

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$$

となる。つまり，共分散は (X, Y) のからみをうまく表している。

この (X, Y) の関係の強さを量的

算数 x_i \ 国語 y_j	1	2	3	4	5	$p_1(x_i)$
1	.04	.03	.01	0	0	.08
2	.04	.10	.05	.01	0	.20
3	0	.04	.30	.08	.02	.44
4	0	.03	.07	.08	.02	.20
5	0	0	.01	.03	.04	.08
$p_2(y_j)$	.08	.20	.44	.20	.08	1.00

表1　算数と国語の成績の同時分布

に把握するには，共分散を次のように正規化した，**相関係数**

$$\rho(X,\ Y)=\frac{(X,\ Y)\text{ の共分散}}{\sqrt{X\text{の分散}}\cdot\sqrt{Y\text{の分散}}}$$

が役立つ。これは次の性質をもつ。

① $-1\leqq\rho(X,\ Y)\leqq 1$

② 任意の定数 $a, b, c, d(\neq 0)$ に対して，$\rho(aX+b,\ cY+d)=\rho(X,\ Y)$

③ X と Y が独立なら，$\rho(X,\ Y)=0$

（無相関という）である。ρ の絶対値が1に近いほど相関が強いという。

（例）　算数と国語の成績　2つの成績 $(X,\ Y)$ の同時分布が表1のように与えられるとき，その関係を調べてみる。まず独立かどうかみる。表より，$p(2,\ 3)=.05$ だが $p_1(2)\cdot p_2(3)=.20\times .44=.088$ と等しくないので，X と Y は独立ではない。国語の成績 $Y=4$ の人の算数の成績 X の条件付分布は，

$p_{1|2}(1|4)=0/.20=0$

$p_{1|2}(2|4)=.01/.20=.05$

$p_{1|2}(3|4)=.08/.20=.40$

$p_{1|2}(4|4)=.08/.20=.40$

$p_{1|2}(5|4)=.03/.20=.15$

であり，X の周辺分布と異なる。その条件付期待値 $E(X|Y=4)$ は，

$$(1\times 0)+(2\times .05)+(3\times .40)+(4\times .40)+(5\times .15)=3.65$$

となる。また，X，Y の平均，分散は，

$\mu_1=\mu_2=3.00$, $\mathrm{Var}(X)=\mathrm{Var}(Y)=1.04$

である。共分散は，

$$\mathrm{Cov}(X,\ Y)=\sum_{i=1}^{5}\sum_{j=1}^{5}(i-3)(j-3)p(i,\ j)=0.70$$

と計算される。相関係数は，

$$\rho(X,\ Y)=\frac{\mathrm{Cov}(X,\ Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}=\frac{.70}{\sqrt{1.04}\cdot\sqrt{1.04}}=.67$$

となり，表1での算数と国語の成績はかなり高い正の相関がある。

■多項分布　日本人で血液型がA型，B型，AB型，O型の人の割合はそれぞれ p_1, p_2, p_3, p_4 であるとする。ランダムに n 人選んだとき，各血液型の人数をそれぞれ X_1, X_2, X_3, X_4 とする。この同時分布は，

$$p(x_1,\ x_2,\ x_3,\ x_4)=\frac{n!}{x_1!\cdot x_2!\cdot x_3!\cdot x_4!}p_1^{x_1}p_2^{x_2}p_3^{x_3}p_4^{x_4}$$

で与えられ，**多項分布**とよぶ。ただし，$x_1+x_2+x_3+x_4=n$ を満たす。これは二項分布を拡張したものである。

A型の人数 X_1 だけの分布（周辺分布）を求めるには，A型とそれ以外の型と2つに分けて考えればよい。この分布は二項分布（→276ページ）

$$\mathrm{P}\{X_1=x_1\}=\binom{n}{x_1}p_1^{x_1}q^{n-x_1}$$

となる。ただし，$q=1-p_1$ である。他の X_1 の周辺分布についても同様に導かれる。これより，

$$E(X_i)=np_i,\ \mathrm{Var}(X_i)=np_i(1-p_i)$$

となる。X_i と X_j の共分散，相関係数は，

$$\mathrm{Cov}(X_i,\ X_j)=-np_ip_j,$$

$$\rho(X_i,\ X_j)=-\sqrt{p_ip_j/(1-p_i)(1-p_j)}$$

となり，負の相関をもつ。

■2次元正規分布　同時密度関係が

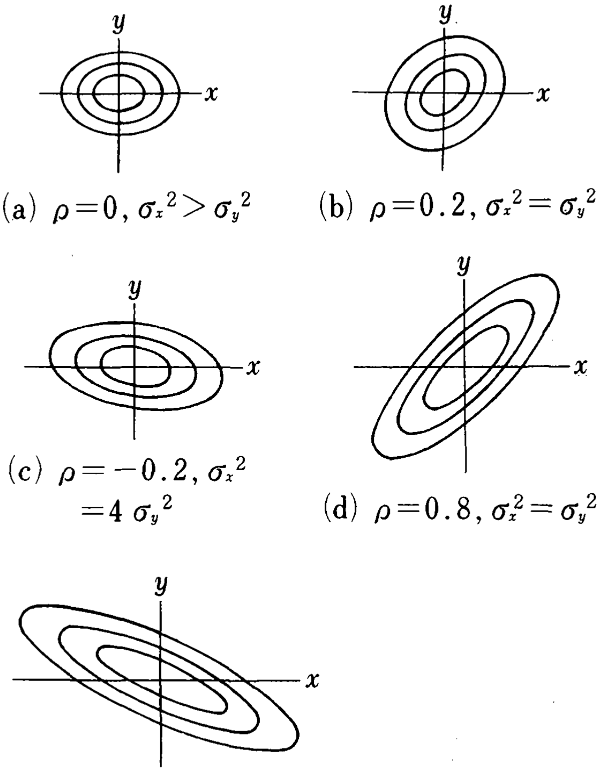
$$f(x,\ y)=\frac{1}{\sqrt{(2\pi)^2\sigma_1^2\sigma_2^2(1-\rho^2)}}e^{-\frac{1}{2(1-\rho^2)}Q(x,\ y)}$$

で与えられる分布をいう。ただし，

$$Q(x,\ y)=\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\times\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2$$

である。密度関数は釣鐘型をしており，その等高線は楕円形となる。X，Yの平均はそれぞれ μ_1, μ_2, 分散はσ_1^2, σ_2^2, 相関係数は ρ となる。この2次元正規分布を $\mathrm{N}(\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2,\ \rho)$ と略記する。ρ, σ_1^2, σ_2^2 の値により等高線の形が変わる(図2)。多次元正規分布も同じように扱うことができる。

また，X，Yの周辺分布はそれぞれ正規分布 $\mathrm{N}(\mu_1,\ \sigma_1^2)$, $\mathrm{N}(\mu_2,\ \sigma_2^2)$ となる。$X=x$ という条件の下での，Yの条件付密度関数は $N(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\ \sigma_2^2(1-\rho^2))$ となる。この条件付期待値



(a) $\rho=0,\ \sigma_x{}^2>\sigma_y{}^2$　(b) $\rho=0.2,\ \sigma_x{}^2=\sigma_y{}^2$

(c) $\rho=-0.2,\ \sigma_x{}^2=4\,\sigma_y{}^2$　(d) $\rho=0.8,\ \sigma_x^2=\sigma_y^2$

(e) $\rho=-0.8,\ \sigma_x{}^2=4\,\sigma_y{}^2$

図2　2次元正規分布の等高線と相関係数
($\mu_x=\mu_y=0$ とする。$\rho=\rho(X,\ Y)$ とおく)

$$E(Y\,|\,X=x)=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$

は，YのXに対する**回帰直線**とよばれる。XのYに対する回帰直線も同様に定まる。

$(X_1,\ Y_1)$, …, $(X_n,\ Y_n)$ を2次元正規母集団 $\mathrm{N}(0,\ 0,\ \sigma_1^2,\ \sigma_2^2,\ \rho)$ からの n 組の標本とする。その積和を

$$W_{xx}=X_1^2+X_2^2+\cdots+X_n^2$$
$$W_{xy}=X_1Y_1+X_2Y_2+\cdots+X_nY_n$$
$$W_{yy}=Y_1^2+Y_2^2+\cdots+Y_n^2$$

とおく。このとき，3変数 $(W_{xx},\ W_{yy},\ W_{xy})$ の従う分布を2次元のウィシャート(**Wishart**)**分布**とよぶ。

(森)

単一因子実験

1つの因子の水準を何段階かにとって，それぞれの水準の下で何回か実験を繰り返すが，その他の原因系の条件は一定に保って行う実験。因子の組み合わせにより効果が異なるときは，最適条件が求められない。

★解説

温度，時間，原料の種類，加工法など多くの**因子**をとりあげて実験するとき，1つの因子のみの水準を変えて，他の因子の水準を固定し，いわゆる1元配置の実験を逐次行って，因子の最適な組み合わせを求めるのが単一因子実験である。これは，2つ以上の因子をとりあげて，それらのすべての組み合わせについて行う**要因実験**に対するものである。

単一因子実験のやり方を具体的に説明しよう。収率向上の目的で，因子として，

温　度　80°，100°，120° … 3水準
反応時間　5分，10分，15分 … 3水準

の2因子をとり，それぞれを3水準に変化させて，単一因子実験により最適条件を求めることを考えてみる。まず，時間を5分に固定し，温度の水準を3段階に変えて実験を行ったところ，温度は120°にしたときに収率が最適であった。そのため，つぎは温度を120°に固定し，時間を3水準に変化させて実験を行ったところ，収率は時間を10分にしたときに最高であった。したがって，最適条件は温度が120°，時間が10分となる。このようにして因子の最適組み合わせを求めるのが単一因子実験である。

単一因子実験では，2つの因子の組み合わせ効果(すなわち交互作用)がなければ，最適条件を求めることができる。しかし，温度により時間の効果が異なる場合には，正しい最適条件が求められないので，このときには，要因実験を行う必要がある。

関連ページ
因子　20
要因実験　396
要因効果　394

B \ A	A_1(80°)	A_2(100°)	A_3(120°)
B_1(5分)	30	50	60
B_2(10分)	50	70	80
B_3(15分)	60	80	90

表1 各水準組み合わせにおける収率(%)
(交互作用が存在しないとき)

B \ A	A_1(80°)	A_2(100°)	A_3(120°)
B_1(5分)	30	50	60
B_2(10分)	60	85	80
B_3(15分)	65	70	65

表2 各水準組み合わせにおける収率(%)
(交互作用が存在するとき)

■交互作用が存在しない場合の単一因子実験 左ページのA(温度)，B(時間)の各水準組み合わせにおける収率(%)が表1のようであったとする。BがB_1のとき，AをA_1, A_2, A_3と順次変えていくと，収率は20，10(%)ずつ順次増加していく。これはBがB_2, B_3のときにも成立する。したがってAの水準効果は，Bの水準によって影響を受けないことになる。このことは，Aの水準を固定してBの水準を変えても成立するので，AとBの組み合わせ効果(交互作用)は存在しない。交互作用が存在しない場合には，単一因子実験により最適水準を求めることができる。表1の場合には，BをB_1に固定してAの各水準で実験するとA_3が最適条件になる。つぎに，AをA_3に固定してBの水準を変えて実験するとB_3が最適条件になる。したがって，最適条件はA_3B_3となる。

■交互作用が存在する場合の単一因子実験 A，Bの交互作用が存在する場合には単一因子実験では最適条件が求められない場合がある。表2のように収率が与えられているとする。BをB_1に固定して，AをA_1，A_2，A_3に変えていくと収率は順次20，10(%)ずつ増加するが，この関係はBがB_2, B_3のときには成立しない。したがってこの場合には，AとBの交互作用が存在する。交互作用が存在する場合に，前と同じ方法で最適条件を求めるとA_3B_2になる。しかし表2からわかるように，真の最適条件はA_2B_2となる。このように交互作用が存在する場合には，すべての各水準の組み合わせについて実験を行わないと，正しく最適条件が求められないことがある。したがって，要因実験を行う必要がある。また交互作用が存在しない場合でも，要因実験を行うことにより，必要な情報を単一因子実験よりも少ない実験回数で得ることができる。

■単一因子実験の用い方 技術的知識が十分でない場合には，因子を1つずつとりあげてこれの特性値への影響を調べることが必要になるし，逆に十分知識がある場合には最適条件を天下りに与えて実験する場合もある。(宮村)

単体法

線形計画問題の基本的解法。シンプレックス法ともよばれている。線形計画法においては，この解法が確立しているため，原理的には必ず解けることが保証されている。大規模な線形計画問題は，計算機を利用して解かれるが，その際にも単体法に基づくソフトウェアが用いられる。1947年にダンツィーク(Dantzig, G.)によって開発された。

★解説

線形計画(LP)の問題は必ず**標準問題**(制約式が等式)に帰着されるので，単体法も標準問題から出発する。標準問題で変数が n 個，等式の制約式が m 本，非負条件が n 個あったとする。これらの制約式をみたす n 個の変数の組 $(x_1, x_2, \cdots, x_n)$ を**許容解**という。

また，$x_1, \cdots, x_n$ のうち，正のものがちょうど m 個あるとき(つまり，残りの $n-m$ 個は0に等しいとき)，その組を**基底解**とよぶ。以下では許容解でありかつ基底解であるものを単に基底解とよぶ。単体法は，基底解に対して，**シンプレックス基準**とよばれる量を計算し，それが正である限りは，目的関数の値をもっと小さくする基底解があると判断して，次の基底解を求めるという手順から成る。

基底解において正の値をとっている変数を**基底変数**とよぶが，基底解の基底変数のうちのどれか1つを0とし，代わりに非基底変数の1つを基底変数に組み入れたものが，上述の「次の基底解」に当たる。単体法においては，次の基底解を求める手続きは**ピボット演算**とよばれるが，これは連立1次方程式における消去演算に相当する。

次の基底解に対するシンプレックス基準はピボット演算で自然に得られ，正のものがあれば次の基底を求める演算を繰り返し，シンプレックス基準がすべて負または0となれば計算を打ち切る。そのときの基底解が求める解である。

関連ページ
線形計画法　218

■具体例　具体的なLP問題を解くことを通して単体法の説明をしよう。

$$
\begin{aligned}
&\text{目的：} 2x_1-3x_2-x_3 \quad \to \text{最小化}\\
&\text{条件：} x_1+x_2-x_3=4\\
&\qquad 2x_1+x_2+x_4=12\\
&\qquad 2x_1-x_2-x_5=2\\
&\qquad x_j \geqq 0 \ (j=1,\ 2,\ \cdots,\ 5)
\end{aligned}
$$

を解くことを考える。

■最初の基底解　単体法は基底解を改良する手段であるから，まず基底解を求めなければならない。上の問題で，もし，条件の x_3，x_5 の係数が + であれば，x_1，x_2 を 0 に等しいとして，$x_3=4$，$x_4=12$，$x_5=2$ という自明な基底解が得られるから，それを最初の基底解と考えればよい。

しかし，この問題の条件はそのようになっていないので，このままでは，自明な基底解が得られない。そこで，**人工変数** x_6，x_7 を導入する。条件の第1式，第3式を，

$$
\left.\begin{aligned}
x_1+x_2-x_3+x_6=4\\
2x_1-x_2-x_5+x_7=2
\end{aligned}\right\} \qquad (1)
$$

と修正する。x_6，x_7 の係数は見かけ上 +1 なので，$x_6=4$，$x_4=12$，$x_7=2$ は自明な基底解になっている。ただ，人工変数は無理に制約式の中に入れたのであるから，LP問題の最小解の基底変数に含まれているのはおかしい。その段階では $x_6=x_7=0$ となっていて，制約条件(1)はLP問題の制約条件に一致していなくてはならない。

■フェーズⅠ　単体法の第1段階では，必要なら人工変数を導入して，自明な基底解がすぐに求まるようにし，目的関数を人工変数の和 x_6+x_7 として，これの最小化を図る。人工変数も非負条件があるので，目的関数の最小値が 0 となれば，人工変数の値はすべて 0 に等しくなり，もとの問題の基底解が得られる。

■フェーズⅡ　この基底解から始め，次々によりよい基底解を求める。

■単体法の演算　2つのフェーズとも，演算はほとんど同じで，それは，

①シンプレックス基準が正の非基底変数を1つ見つけ，これを次の段階で基底変数に入れることにする。

②その代わりに基底変数から追い出す変数を定める。そのため，θ と記される量(後出)を計算し，その最小値を与える変数を追い出す。

③①と②で見つけた変数の入れ替えのためにピボット演算を行う。

④新しいシンプレックス基準を見て，正のものがなければ計算を止め，正のものがあれば①に戻る。

という計算手順である。

■単体表　この計算を進めるには，**単体表**(シンプレックス表ともいう)を作るとわかりやすい。その際，目的関数を

$$z_0=2x_1-3x_2-x_3$$

と置き，この式を移項して，

$$z_0-2x_1+3x_2+x_3=0$$

を新たな制約式とし，目的関数 $z_0 \to$ 最小化　という形をとると便利である。

フェーズⅠの目的関数を z_1 とおき，フェーズⅠでは $z_1 \to$ 最小化　を目的と考える。そして，目的関数を変形して，最初の基底解の基底変数をその中に含まないようにするのが，第一の仕事である。すなわち，制約式は，

$$\begin{array}{llllll}
z_1 & & & -x_6-x_7 & =0 \\
z_0-2x_1+3x_2+x_3 & & & & =0 \\
\quad x_1+2x_2-x_3 & & +x_6 & & =4 \\
\quad 2x_1+x_2 & +x_4 & & & =12 \\
\quad 2x_1-x_2 & & -x_5 & +x_7 & =2
\end{array}$$

であるので，第1式に第3，第5式を加え人工変数 x_6，x_7 を消去すると，

$$z_1+3x_1+x_2\quad -x_3\quad -x_5\qquad =6$$

と書ける。

■最初の単体表　単体表とは上の制約式の各係数を表の形にまとめたもので，最初の表は次のようになる。

	1	z_1	z_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
z_1	6	1	0	3	1	−1	0	−1	0	0
z_0	0	0	1	−2	3	1	0	0	0	0
x_6	4	0	0	1	2	−1	0	0	1	0
x_4	12	0	0	2	1	0	1	0	0	0
x_7	2	0	0	2	−1	0	0	−1	0	1

表1

ここで，左端と上端の欄に書いた文字はそれぞれ，基底変数及び全変数を表しており，今後 z_1 行とか x_3 列とかよぶことにする。1と書いた列は定数項を表している。この列は z_0 行との交点の値以外は正でなければならない。

段階①　さて，この表で，z_1 行の各非基底変数に対応する個所の値がシンプレックス基準とよばれる量を表している。ここでは x_1 の下の値が3で正である。これは x_1 を基底変数に入れることで z_1 の値をより小さくすることが可能であることを示している。それで，x_1 を基底変数に入れる。

段階②　次に，x_1 列の z_1，z_0 行以外で正のものを探す。ここでは x_6，x_4，x_7 の各行とも合格である。この値で対応する1の列の値を割ったものを θ とする。この例では，

$$\frac{4}{1}=4,\quad \frac{12}{2}=6,\quad \frac{2}{2}=1$$

がそれぞれの θ の値となる。この最小値は最後の1であるから，基底変数から追い出すのは x_7 である。

ところで，x_1 列と x_7 行の交点にある数を**ピボット**とよぶ。単体表を見るとわかるように，基底変数になっている列は，自身の行との交点の要素だけが1で，他はすべて0になっている。たとえば，x_4 列では下から2行目の x_4 行との交点の要素は1で，他の各行に対応する値はすべて0である。

段階③　段階①で，x_1 が基底変数にはいり，段階②で x_7 が基底変数から出ることが定まったので，新しい単体表では，最下段の行が x_1 に変わる。そこで，x_1 列の最下段の要素が，新しい単体表では1になっており，他の x_1 列の要素はすべて0になっているはずである。

ピボット演算は，そのような変換を

実行する。すなわち，ピボットを1とし，他の x_1 列の要素を0に変える。そのため，まずピボットが1になるよう，ピボットの値でその行全部を割る。いまの例では，x_7 行のすべての値を割るわけである。すると，それぞれの値は

1 0 0 1 −.5 0 0 −.5 0 .5

となる。これらの値の3倍を z_1 行から引くと，それぞれ

3 1 0 0 2.5 −1 0 0.5 0 −1.5

である。このように，x_1 列は最下段が1で他は0であるようにするため，各行ごとに同様の演算を実行する。その結果は表2のようになる。

	1	z_1	z_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
z_1	3	1	0	0	2.5	−1	0	.5	0	−1.5
z_0	2	0	1	0	2	1	0	−1	0	1
x_6	3	0	0	0	2.5	−1	0	.5	1	−1
x_4	10	0	0	0	2	0	1	1	0	−1
x_1	1	0	0	1	−.5	0	0	−.5	0	.5

表2

段階④と繰り返し これで段階③が終わり，z_1 行を見ると x_2 の下に2.5があるので，再び段階①へ戻って，x_2 を基底変数に入れる。②③の段階の演算を実行すると，次の表3が得られる。

	1	z_1	z_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7
z_1	0	1	0	0	0	0	0	0	−1	−.5
z_0	−.4	0	1	0	0	1.8	0	-1.4	−.8	1.8
x_2	1.2	0	0	0	1	−.4	0	.2	.4	−.4
x_4	7.6	0	0	0	0	.8	1	.6	−.8	−.2
x_1	1.6	0	0	1	0	−.2	0	−.6	.2	.3

表3

ここで1の列の x_1 行との交点の値が0になった。これは $z_1=0$ となり，フェーズⅠが終わったことを意味する。そこで，z_1 行と人工変数 x_6，x_7 を除いた単体表によりフェーズⅡを進める。ピボット演算の仕方はいままでと同じであるから，計算結果だけを示す。

	1	z_0	x_1	x_2	x_3	x_4	x_5
z_0	−.4	1	0	0	1.8	0	-1.4
x_2	1.2	0	0	1	−.4	0	.2
x_4	7.6	0	0	0	.8	1	.6
x_1	1.6	0	1	0	−.2	0	−.6
z_0	−17.5	1	0	0	0	−2.25	-2.75
x_2	5	0	0	1	0	.5	2.24
x_3	9.5	0	0	0	1	1.25	.75
x_1	3.5	0	1	0	0	.25	−.45

x_3 を基底に入れ，x_4 を出す。

表4

■LPの解 この結果，シンプレックス基準はすべて負になったので，計算はこれで止め，$x_1=3.5$，$x_2=5$，$x_3=9.5$ のとき，目的関数 z_0 は最小値 −17.5 をとることがわかる。これが，はじめのLP問題の解である。

■単体法の改善 表2の x_7 列，表3の x_6，x_7 列，表4の x_5 列，第2段目の x_4，x_5 列などは，計算にはかなりの手間がかかっているのに，結局は解に何の影響もしていない。この特徴を上手に利用して，これらの計算をしないですませると計算の手間を大幅に縮小できる。これが**改訂単体法**とよばれる方法で，実用化されている計算機プログラムはこれを用いている。(森村)

直交表

任意の2因子について，その水準のすべての組み合わせが，同じ回数ずつ現れるように実験を割り付けるときに用いられる表。2水準系と3水準系の直交表がよく用いられる。

★解説

表1に2水準系の直交表の一例を示す。これは次の性質を満たす。

(1) 各列に1と2が同じ回数ずつ現れる。

(2) どの2列をとっても(1, 1), (1, 2), (2, 1), (2, 2)が同じ回数ずつ現れる。

No. ＼ 列番	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2
基本表示	a	b	ab	c	ac	bc	abc
群番号	1	2	2	3	3	3	3

表1　2水準系の直交表

この2つの性質を満足する表を一般に直交表とよび，$L_N(p^k)$という記号で表す。Lは直交表の起源となった**ラテン方格法**(*Latin Square*)の頭文字, Nは実験の大きさ(行の数), pは**因子**の水準数，kは列の数を示している。表1の直交表は $L_8(2^7)$ である。

基本表示の下には群番号が示されているが，これは次の性質をもつ。

(1) 第1群の列は基本表示aのみを含み，数字1と2が4つずつまとまっている。

(2) 第2群の列は基本表示a，bのみを含み，1，2が2つずつまとまっている。

(3) 第3群の列は基本表示a，b，cを含み，1，2が原則的に1つずつ変わる。

この性質は**分割法**のときに利用される。基本表示は任意の2列の積がどの列に現れるかをみるときに用いられる。1列と3列の積は，$a \times ab = a^2b \equiv b$ より第2列となる。

関連ページ
ラテン方格法　400
因子　20
分割法　330

■直交表による実験　4つの因子 A, B, C, D をとりあげ，これらはすべて2水準とする。もし4因子要因実験を行えば，すべての水準の組み合わせ $2^4=16$ 回の実験をしなければならない。ところで，因子間の交互作用がないものと仮定すれば，直交表 $L_8(2^7)$ を用いて表2のような8回の実験を行うことにより，A, B, C, D の効果を検出できる。ここでは，実験を次のように割り付けて行っている。

(1)　$L_8(2^7)$ には7つの列があるが，A を1列，B を2列，C を4列，D を7列に割り付けている。

(2)　直交表の各列の数字1，2を，その列に割り付けられた因子の水準に対応させると，8つの水準の組み合わせが決まる。

(3)　この8つの実験の順序を無作為化して実験を行って，データを得る。

いまのように交互作用が存在しない場合には，因子の水準の組み合わせすべてを実験しなくても，その1/2実施にあたる8回の実験で A, B, C, D の効果を調べることができる。たとえば，A の2水準 A_1 と A_2 の比較を行うには，

（A_1 水準でのデータの合計）

−（A_2 水準でのデータの合計）

$=(x_1+x_2+x_3+x_4)-(x_5+x_6+x_7+x_8)$

の比較を行えばよい。これは，直交表の第2の性質から，A_1, A_2 で B_1, B_2 がそれぞれ2回ずつ実験されており，

列番	1	2	4	7	水準の組み合わせ	実験順序	データ
No. ＼ 要因	A	B	C	D			
1	1	1	1	1	$A_1\ B_1\ C_1\ D_1$	5	x_1
2	1	1	2	2	$A_1\ B_1\ C_2\ D_2$	1	x_2
3	1	2	1	2	$A_1\ B_2\ C_1\ D_2$	3	x_3
4	1	2	2	1	$A_1\ B_2\ C_2\ D_2$	6	x_4
5	2	1	1	2	$A_2\ B_1\ C_1\ D_2$	2	x_5
6	2	1	2	1	$A_2\ B_1\ C_2\ D_1$	8	x_6
7	2	2	1	1	$A_2\ B_2\ C_1\ D_1$	7	x_7
8	2	2	2	2	$A_2\ B_2\ C_2\ D_2$	4	x_8

表2　直交表による実験

また C, D についても同様なことがいえるので，B, C, D の影響が入らないことによる。B_1 と B_2, C_1 と C_2, D_1 と D_2 の比較もいま述べたのと全く同じようにできる。

■各要因の平方和の計算　2水準系での i 列の平方和 $S_{(i)}$ は，上に述べた考え方より，

$S_{(i)}=(T_1-T_2)^2/$(データの総数)

$T_1=i$ 列が1であるデータの和

$T_2=i$ 列が2であるデータの和

と求められる。このことより，A, B, C, D の平方和は列番1，2，4，7の平方和より，また誤差要因の平方和は残りの3列の3，5，6列より求めることができる。なお，各列の平方和の自由度が2水準系の場合は1であることに注意して不偏分散を求めれば，各要因効果の有無を分散分析によって調べることができる。

交互作用がある場合，3水準系の直交表については〔12〕〔61〕を参照。（宮村）

追加効率指標の使い方

追加効率指標は，排反案(→ 116 ページ)の比較に使うための指標である。2 つの排反案(代替案)があるときに，両者の「効果の差」と資源の「投入量の差」の比率で与えられる。

$$追加効率指標 = \frac{効果の差}{投入量の差}$$

★解説

目的地に行くための 2 つのルート A，B があるとしよう。ルート A を通ると所要時間は 30 分で交通費は 1500 円かかり，ルート B では 20 分で 2000 円かかる。経済的なルートはどちらになるかを判断するのに，むやみに「率」の値を計算してはいけない。たとえば，1 分当たりのコストを求めると，ルート A は 50 円/分，ルート B は 100 円/分になり，ルート A が効率的だといってよいかというと決してそうではない。

判断のねらいによって使う指標も違ってくる。出費に関係なく，とにかく速く行きたいときは所要時間の短いルート B がよく，時間に余裕があり，安く行きたいのなら，交通費の安いルート A がよいことになる。もし，時間が貴重で忙しい人が限られた時間のなかでなるべく安く行くことが問題なら，つぎのような考え方をしなければならない。

この人はとにかく目的地に行かねばならないので，A か B かのいずれかのルートを通ることになる。そこで，所要時間の短い B を通ると仮に決めてみて，A に乗りかえることを考える。B から A に乗りかえると，余分に 10 分かかるわりに 500 円安くなり，1 分当たり 50 円の効果が得られることになる。もし，この人が 1 分当たり 50 円以上の率でかせぐ機会をほかにもっているのなら，A に乗りかえるのは損になるだろう。これが追加効率の考え方である。

関連ページ

効率指標の使い方
116

■**追加効率の適用例**　ある下請け工場に包装設備Aがある。この設備で包装する際に，包装する製品と包装材の供給作業と，包装した完成品をパレットに積んで運搬できるような荷姿をつくる作業が必要になる。

これらの作業を1人の作業者で行うと，作業ペースが遅くなり，設備の能力に遊びが生じて週当たり1000個しか包装できない。

投入する作業者の数を2人, 3人と増加させていくと,設備の稼動率が上がり,処理量はそれぞれ1600個,1900個と増すことが見積もられている。

いま，この工場に1個につき100円の利益が得られる注文の引き合いがきたとしよう。ちょうどこの時は別の作業に人がとられ，忙しくて人手不足になっている。注文量はたくさんあり，もし受けるとすると1人につき週当たり4万円のアルバイト作業者を雇わねばならない。何人の作業者を雇うのが利益を最大にするだろうか。

この問題は追加効率の考え方を用いて解くことができる。人数を0人から1人に増やすと1人当たりの追加効率(追加利益率)は10万円/人になり，1人当たりの人件費4万円/人より大きいので，1人追加するのは有利になる。次に1人から2人に増やすと追加効率は6万円/人で同様に有利になる。2人から3人に増やすと追加効率は3万円/人で，1人当たりの人件費4万円/

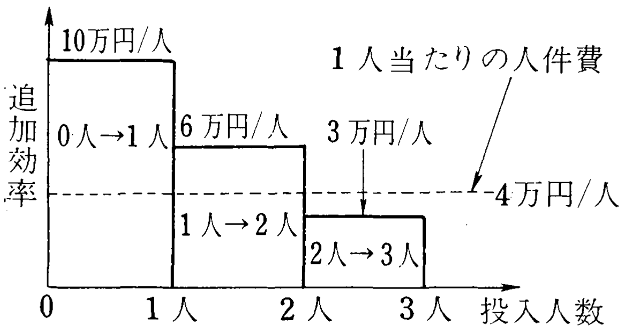


図1　設備Aでの追加効率

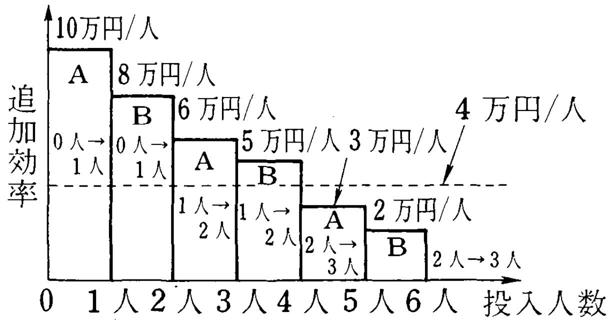


図2　設備AとBでの追加効率

人より高くつくので不利になる。したがって，2人雇って1600個の注文をこなすのがよいことがわかる。

追加効率指標はつぎのような問題状況のときに適用すると，その真価を発揮することになる。前記の工場にはもう1台の包装設備Bもあり，その設備を使うと，次のような追加効率の値が得られたとしよう。

0人→1人	8万円／人
1人→2人	5万円／人
2人→3人	2万円／人

この設備でも作業者を2人雇うのがよいことになる。そこで，雇える作業者が3人しかいない場合はどうなるか。

図2のように設備AとBの追加効率を大きい順にならべてみることにより，設備Aに2人，設備Bに1人雇うのがよいことが容易にわかる。(中村)

手余り状態と手不足状態

生産能力が需要量を上回っている状態を「手余り状態」といい，逆に需要量が生産能力を上回っている状態を「手不足状態」という。手余り状態は暇なときを，手不足状態は忙しいときをいう。

手余りか手不足かにより，方策の有利さは全く異なってくるので注意しなければならない。

★解説

生産能力を構成する要素には，設備や人員のように，需要量の変動に即応して能力を増減できないものが含まれている。そのために生産能力が需要量にぴったり一致するときはそれまで，次の2つの状態のどちらかであることになる。

(1) 需要が少なくて生産(供給)能力が余っている状態。

(2) 需要が多くて生産(供給)能力が不足している状態。

このように，企業や工場や工程が前者の状態にあるときを手余り状態にあるといい，後者の状態にあるときを手不足状態にあるという。

いずれの状態にあるかで，方策の損得の内容は全く異なることになる。一般には，手余り状態のとき供給能力を増しても効果は生じない。販売量を増す方策とか供給能力を他に転用する方策，供給能力を減らす方策が効果を生むことになる。また，手不足状態のときは供給能力を増す方策が大きな効果をもたらすことになる。

不良の減少，設備の稼動率の増大，材料原単位の減少，方法改善による作業能率の向上，内外製区分…といった方策を評価する際には，手余りか手不足かによりコストの減少の仕方，収益の増加の仕方が異なり，方策の効果が大幅に違ってくるので注意しなければならない。

関連ページ

経済性の比較の原則　98

機会損失　70

■**不良低減の効果**　ある工場で製品Aをつくっている。最近の生産量は月当たり約2万個で，そのうち平均して10％が不良になっている(良品が1万8000個)。製品1個当たりの売り値と製造原価は，次のように求まっている。

売　価	1000円
材料費	300円
変動加工費	150円
人件費	200円
減価償却費	100円
固定経費	70円
利　益	180円

ここで，人件費，設備の減価償却費，固定経費は月々の総額を1個当たりに割り掛けたものである。

いま，この工場で，不良率を半減する方策を検討することになった。その改善のために，コストをいくらまでかけられるかをみてみよう。

(1)**手余り状態の場合**　良品を1万8000個以上つくっても売れないので，不良率を低減しても収益の増にはならない。つぎに，1万8000個の良品を得るのに必要な生産量を求めてみよう。

18000÷(1－0.05)＝18948個

したがって，良品率が90％から95％にあがると，生産量は2万個から1万8948個になり，1052個減少する。この分だけ材料費と変動加工費を節減できる。

	手余り	手不足
収　益	×	○
材料費	○	×
変動加工費	○	×
人件費	×	×
減価償却費	×	×
固定経費	×	×

○印：変動する　×印：変動せず

表1　変動の仕方

(300＋150)×1052＝473400 円

この値が改善のためにかけてよいコストの上限になる。なお，人件費，減価償却費，固定経費は不良率が向上しても変化しない費用で計算の対象にのぼってこないことに注意しておこう(経済性の比較の原則→98ページ)。

(2)**手不足状態の場合**　フル生産して良品数を増し収益の増大をはかることができる。生産量は2万個で良品数は1000個増加することになる。したがって，不良率を低減しても，費用は変わらず，収益が良品数の増加分だけ増すことになる。

1000個×1000円/個＝100万個

100万円以下の改善コストであればペイすることになる。

このように，不良率の低減効果も，手余り状態と手不足状態とでは大きく異なってくる。その理由は，手余りと手不足で比較の対象(この例では不良率10％と5％の場合)間で変動要素が，表1のように異なるからである。

(中村)

定期発注方式

発注間隔をあらかじめ定めておき，発注量をその都度定める在庫管理方式。次の発注時点までの需要量の予測が大切になる。有限の品切れ損失を考慮する場合には，新聞売り子の問題とよばれるタイプが基本的なものとなる。

★解説

■定期発注法の特徴　発注時期にいちいち予測をして需要量を定めるので，対象品目が多くなりすぎるとやりにくくなる。また，細かな情報を加味することも可能なので，金額の張るもの，需要変動の大きいものにはこの方式が適している。

■需要予測　品目ごとに過去の需要変動を常に追跡・管理し，できれば変動の理由をできるだけ明確にしておくことが望ましい。他の品目と独立でなく使用されるものには特に注意が必要である。指数平滑法が利用される場合もある。

■発注量　つぎの発注時までと調達期間内との需要をまかなう必要量から，手持ち分を差し引いた量を発注することになる。ただし，これに安全余裕をつけ加えるのがふつうである。したがって，

発注量＝[(次の発注までの期間＋調達期間)内の需要量の推定値] − [現在発注してあって，まだ納入されていない量] − [現在の在庫量] ＋ [安全余裕]

となる。

■安全余裕　発注点法と同じ方式も考えられるが，もっと簡単には，上の発注量の計算式の第１項の推定値とその実績との差を記録しておき，その差の最大値を用いることが考えられる。あるいは，この差から，標準偏差を推定し，発注点法のときと同様の安全係数を乗じて求める方法も利用されている。

関連ページ
新聞売り子の問題　182
指数平滑法　158

■発注量の計算例 (1)需要が確定している場合。

表1のようなデータが与えられたとする。

調達期間	2ヵ月
発注間隔	1ヵ月
現在在庫	30 箱
発注残	120 箱
今月の需要	55 箱
来月の需要	65 箱
再来月の需要	70 箱

表1

調達期間が2カ月であるから，いま発注する品物は来月と再来月には間に合わない。その間の需要は，既に発注済みの品物で手当てされなければならない。既に発注済みで，まだ納入されていないものを**発注残**とよぶ。これと現在の在庫が向こう2カ月間の需要に当てられるから，2カ月後の在庫は,

$$\text{現在の在庫} + \text{発注残} - \text{調達期間中の需要}$$

として求められる。したがって，いま発注すべき量は，第3月目の需要量からこの2カ月後の在庫を引いた値である。結局，左ページに示す発注量の式で，安全余裕0の計算をすればよい。すなわち,

$$(55+65+70)-120-30=40$$

が今期の発注量となる。

(2) 需要に偶然変動がある場合。

将来の需要が完全に確定的であるとは考えにくいのがふつうであろう。そのとき，調達期間と発注間隔内の需要量を推定しなければならない。それには需要変動の特徴をつかむことが望ましい。もちろん，需要量について特別の情報があればそれを活用しなくてはならないが，ここでは簡単のため，毎月の需要量は独立に変動するものと仮定する。過去7カ月の需要が表2のようであったとすると，月平均需要量は61，範囲は16である。

月	需要量
1	60
2	52
3	68
4	66
5	58
6	64
7	59

表2

したがって，表1の調達期間から発注残までは同じとすると，3カ月間の平均需要量は183箱である。安全余裕の計算法は発注点方式のときと同じ(→296ページ)方式によると,

$$1.65\times\sqrt{3}\times16\times0.37=16.9$$

となるから，発注量は

$$183-120-30+17=50\ \text{箱}$$

と計算される。この場合は，過去の推定値がわからないので，推定値と実績の差を利用することは考えていない。

■品切れ損失を考える場合 上の計算例では，品切れを起こさないということだけを目的とし，その範囲で最小の在庫を狙っていた。もし，品切れが起きたとき，若干の損失を見込むことができれば，保管費との釣り合い上最適な初期在庫量を定めることができる。これは，発注点方式においてEOQを定めたのと同様の考えに基づく。

この型の在庫モデルで典型的なものは，新聞売り子の問題と (s, S) 方式である。 (森村)

データベース

コンピュータ処理が可能な形で，ある特定の目的のために一定の形式で集めたデータをファイルという。1つの組織体には，このような個別ファイルが多数あり，同一データ項目が複数ファイルに重複することやある処理のために複数のファイルを必要とすることが多い。そこでデータを整理・統合して1つの集合体として，利用者が必要な時に必要なデータを利用できるようにしたものをデータベースという。

★解説

■データベース・サービス　データベース・サービスを行う組織・機関などをデータバンクという。データベース・サービスには，データベースの構築，流通，検索，解析などがある。データベースを構築する業者をプロデューサーという。多額の費用と情報の更新作業を必要とする。構築された情報を，磁気テープやコンピュータ・システムにより提供する業者をディストリビューターという。国際的な情報ネットワークにより欧米のデータベースを日本で利用できる機会も増えたが，米国などでは一部の情報を海外に提供しない事態も発生している。ユーザーに代わってデータベースの情報を検索・解析する業者をブローカーという。文献検索などにこの種のブローカーが多い。なお，データバンクをデータベースと同意義に用いることもある。

■データベースの種類　データベースには文献・記事に関するもの，数値に関するもの，画像に関するものがある。文献・記事データベースは，書誌(ビブリオグラフィック)，抄録(アブストラクト)，原報(フル・テキスト)と分類され，論文などの科学技術文献や新聞記事などが対象とされている。数値データベースには，科学技術データ，物性データ，経済データなどがある。画像データベースはこれからの課題とされている。

関連ページ

統計パッケージのファイル管理
264

ディストリビューター	サービス(システム)名	提供される主なデータベース	料金体系
日本科学技術情報センター	JOIS-Ⅱ	科学技術の文献データベース	従量制
日本特許情報センター	PATOLIS	国内特許，実用新案や米国特許など（文献データベース）	従量制
日本経済新聞社	NEEDS-TS	日本のマクロ経済，ミクロ経済，米国のマクロ経済，株価（ファクト・データベース）	会員制
	NEEDS-IR	新聞・雑誌記事（文献データベース）	従量制
市況情報センター	QUICK	株価（国内，海外）市況，外国為替（ファクトデータベース）	会員制
日本SDC	SEARCH-J	世界の特許，医薬，化学などの文献データベース	会員制
G.E.Information Service Co.(電通国際情報サービス)	MARK Ⅲ	国内外のマクロ経済，産業統計，株価などファクト・データベース	従量制
DIALOG Information Service, Inc.(丸善，紀伊国屋)	DIALOG	約100種の科学技術，経済，経営，社会・人文科学，学際的な文献データベース	従量制
SDC(日本SDC)	ORBIT	約60の科学技術，経済・経営，社会・人文科学，学際的な文献データベース	従量制
The N.Y.Times Information Service Co.(日本経済新聞社)	The N.Y.Times Information Bank (The Information Bank)	N.Y.Times紙他約90種の新聞，雑誌の文献データベース	従量制
CAS(Chemical Abstracts Service)(化学情報協会)	CAS ONLINE	CASが収集した数百万件の化学物質の構造などのファクト・データベース	従量制
日本電子計算	JIP/BRS	世界の医学，電気工学などの文献データベース	従量制

表1　日本で利用可能な主なオンライン・サービス

■データベース利用に際して　世界のデータベース数は約1400，オンラインサービスされているものが約960と推定されている。そのうち日本で利用可能なものが約250である。現在，オンラインは主として抄録に限られているようであるが，将来は原報もバッチ処理からオンライン処理に代わると思われる。日本において利用可能な主なオンライン・サービスのデータベースには，表1に示すようなものがある。これは1982年の段階の資料で，次の文献から転載したものである。

長田洋：「データベース・サービスの現状」(『オペレーションズ・リサーチ』No4，Vol.27，182/186 (1982))（新村）

適合度検定

観測されたn個の資料をk個の場合に分類して，これらの資料がある特定の母集団分布からとられたものであるとみなしてよいかどうかを検定する。これを適合度検定という。適合度検定には χ^2 (カイ2乗)分布表が用いられる。

★解説

n個のデータをk個の場合に分類したところ，表1の観測値の欄に示すように，F_1，F_2，…，F_k になったとする。「もし，データがある特定の分布からとられたものであるとすれば，それらのマス目に入る数の期待値がいくらになるか」を計算する。このような数のことを**理論値**という。観測値 F_i ($i=1, 2, \cdots, k$) に対応して，理論値f_iが得られたとすると，統計量

$$\chi^2=\frac{(F_1-f_1)^2}{f_1}+\frac{(F_2-f_2)^2}{f_2}+\cdots+\frac{(F_k--f_k)^2}{f_k}$$

は近似的に **χ^2 分布**にしたがう。上式により算出した実現値のことを **χ^2 値**とよぶことがある。観測値と理論値とのくい違いが大きいほど，χ^2 値が大きくなる。そこで，観測値と理論値との間に**有意差**があるかどうかを，χ^2 分布を用いて検定することになる。そのためには，自由度がいくらになるかがわからないといけない。自由度とは，大まかに「理論値の個数から，それを算出するのに用いたデータからの制約の個数を減じた数」と考えればよい。表1では，理論値がk個あるが，これを計算するにあたって，合計がnであるという情報をデータから得ている。それ以外には，データに頼らないで理論値が算出されているならば，自由度は ($k-1$) ということになる。しかし,さらにデータの平均を用いているならば,自由度は($k-2$)になる。なお,理論値が5未満のところは，5以上になるようにプールしておかないと，χ^2分布を使うときの近似がわるくなる。

関連ページ

χ^2 分布　314

有意差　65

場合	1	2	……………	k	計
観測値	F_1	F_2	……………	F_k	n
理論値	f_1	f_2	……………	f_k	n

5未満のところはプールする。

$$\chi^2=\frac{(F_1-f_1)^2}{f_1}+\frac{(F_2-f_2)^2}{f_2}+\cdots+\frac{(F_k-f_k)^2}{f_k}$$

この値とχ^2分布表の値とをくらべる。

表1　適合度検定の考え

サイコロの目	1の目	2の目	3の目	4の目	5の目	6の目	計
観測値	9	11	13	10	8	9	60
理論値	10	10	10	10	10	10	60

表2　サイコロの目の適合度検定

<table>
<tr><td>死亡者数</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>計</td></tr>
<tr><td rowspan="2">観測値</td><td rowspan="2">109</td><td rowspan="2">65</td><td>22</td><td>3</td><td>1</td><td rowspan="2">200</td></tr>
<tr><td colspan="3">まとめて 26</td></tr>
<tr><td rowspan="2">理論値</td><td rowspan="2">108.7</td><td rowspan="2">66.3</td><td>20.2</td><td>4.1</td><td>0.7</td><td rowspan="2">200</td></tr>
<tr><td colspan="3">まとめて 25.0</td></tr>
</table>

表3　馬にけられて死んだ人

■サイコロの目の適合度検定　表2はサイコロを60回投げたときの，出た目の数の回数である。もし，サイコロが正当なものであれば，どの目も10回実現することが期待される。これが理論値である。理論値は6個あるが，これを算出するにあたって，データからうける制約は，合計が60であるということだけである。したがって，自由度は$6-1=5$になる。表2から，観測値と理論値とのくい違いを調べると，

$$\chi^2=\frac{(9-10)^2}{10}+\frac{(11-10)^2}{10}+\cdots+\frac{(9-10)^2}{10}=1.6$$

である。一方，自由度5のχ^2分布での片側5％点をχ^2分布表から読むと，

$$\chi^2_5(0.05)=11.07$$

よって「サイコロは正しくつくられている」との仮説は棄却されない。

■馬にけられて死んだ人　プロシャの軍隊の話である。Bortkiewiczという人が，10個の隊で20年間に，馬にけられて死んだ人数を調べ，1隊1年を1個の区画と考えて，死亡件数の分類を行い，表3のようにまとめた。死亡者数が平均λのポアソン分布(→352ページ)にしたがうという仮説をたて，観測値と理論値とのくい違いを調べ，χ^2分布表を用いて検定してみると，仮説は棄却されないことがわかる。ここで，理論値は値の小さいところをプールして3個。これを算出するには，合計が200であることと，データの平均$\lambda=0.61$を用いるので，自由度は$3-2=1$になる。　(牧野)

デザイン行列(配置行列)

デザイン行列は，重回帰分析の説明変数の値(連続変数)や分散分析などの実験の配置(1，0，−1などで表す)をデータ行列としてまとめたものである。

デザイン行列の作成方法を知れば，重回帰分析，分散分析，共分散分析，多項式回帰，反応曲面モデル，数量化Ⅰ類，2群の場合の判別分析と数量化Ⅱ類などの諸手法を，一般線形モデル(統計的線形モデル)として統一的に扱うことができる。

★解説

■一般線形モデル　yを外的基準(目的変数)とし，p個の説明変数(因子) x_1，…，x_pの線形結合で表される式(1)を一般線形モデルとよぶ。

$$y_i=\beta_0+\beta_1x_{1i}+\beta_2x_{2i}+\cdots\cdots+\beta_px_{pi}+\varepsilon_i \qquad (1)$$

$$(i=1,\ \cdots,\ n)$$

ここで，添字iはnサンプル中のi番目のデータを意味し，ε_iは誤差項を表す。これを行列で表せば式(2)になる。

$$\underset{\boldsymbol{y}}{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}=\underset{\boldsymbol{X}}{\begin{bmatrix} 1 & x_{11} & \cdots & x_{p1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{1n} & \cdots & x_{pn} \end{bmatrix}}\underset{\boldsymbol{\beta}}{\begin{bmatrix} \beta_0 \\ \vdots \\ \beta_p \end{bmatrix}}+\underset{\boldsymbol{\varepsilon}}{\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}} \qquad (2)$$

ここで，行列Xを**デザイン行列**とよぶ。1列目は定数項であり，2列目以降は連続変数または因子水準の1，0，−1などである。$\boldsymbol{y}$とXからパラメータβの値を推定する。

関連ページ

質的データの解析　162

FUNCATによる判別　320

林の数量化理論　298

数量化Ⅰ類　188

数量化Ⅱ類　190

■デザイン行列の構築　定数項β_0を推定するため，1列目の要素をすべて1にする。

連続変数x_iを**回帰効果**とよび，そのままの値がデザイン行列の1列になり，β_iが推定される。x_i^2などは，x_iの2乗値を新しく回帰効果と考え，デザイン行列の1列とする。

x_iが因子の場合，因子水準に応じて以下の規則で何列かに展開される。

■主効果　因子(カテゴリカル変数)が水準mならば，m列に展開される。各水準は各1列に対応する。AとBがそれぞれ2水準と3水準のデータの場合，その主効果は2列と3列の0/1データになる。

data		A		B		
A	B	A_1	A_2	B_1	B_2	B_3
1	1	1	0	1	0	0
1	2	1	0	0	1	0
1	3	1	0	0	0	1
2	1	0	1	1	0	0
2	1	0	1	0	1	0
2	3	0	1	0	0	1

■交互作用効果　前出の主効果に対し，次の$A*B$を交互作用効果という。

AB	A_1B_1	A_1B_2	A_1B_3	A_2B_1	A_2B_2	A_2B_3
11	1	0	0	0	0	0
12	0	1	0	0	0	0
13	0	0	1	0	0	0
21	0	0	0	1	0	0
22	0	0	0	0	1	0
23	0	0	0	0	0	1

交互作用効果の各列は，主効果の列A_iとB_jの積A_iB_jになる。すなわち，2×3の6列に展開される。容易に分かる通り，主効果の5列は交互作用効果の6列と線形独立ではない。

■ネスト効果　主効果Aとネスト効果$B(A)$を指定した場合，主効果Aの2列と交互作用効果$A*B$の6列，計8列が作られる。すなわち，主効果B(ネスト変数とよぶ)をモデルに入れない点に特徴がある。

■連続変数に対するネスト効果　連続変数XをAでネストした効果$X(A)$は次のように展開される。

		A		$X(A)$	
X	A	A_1	A_2	$X(A_1)$	$X(A_2)$
21	1	1	0	21	0
24	1	1	0	24	0
22	1	1	0	22	0
28	2	0	1	0	28
19	2	0	1	0	19
23	2	0	1	0	23

すなわち，主効果Aの各列の要素に対応する連続変数の値を掛ければよい。これにより，Aの各水準ごとにXの傾きを計算できる。

■デザイン行列の構成2　水準kの主効果Aのj列を論理式$A(j)=(A=A_j)-(A=A_k)$で与える。この方法によって作られたデザイン行列はフルランクになる。ただし，$(A=A_j)$はAが水

A	A_1	B	B_1	B_2
1	1	1	1	0
2	−1	2	0	1
		3	−1	−1

準 j の時1，それ以外は0とする。

■各種モデル いま，A，B，C をカテゴリー変数，そして X_i を連続変数を表すものとする。

次に，一般線形モデルのデザイン行列を生成し，その推定値を求める仮空の命令 MODEL を考える。

MODEL　目的変数＝説明変数

たとえば，次の命令は，

MODEL　$Y=X_1$

単回帰モデル

$$y=\beta_0+\beta_1 x_1+\varepsilon$$

を表す。デザイン行列 X の1列目は定数項に対応し1の値をとり，2列目には X_1 の値がくる。

以下にモデル式とモデル名を記す。

$Y=X_1$	単回帰モデル
$Y=X_1\quad X_2$	重回帰モデル
$Y=X_1\quad X_1*X_2$	多項式回帰モデル
$Y=A$	1元配置
$Y=A\ B\ C$	3元配置で主効果のみ
$Y=A\quad B\quad A*B$	2元配置で交互作用あり
$Y=A\quad B(A)$ $=A\quad A*B$	階層モデル
$Y=A\quad X_1$	分散共分散分析
$Y=A\quad X_1(A)$	傾斜が A の水準ごとに異なる
$Y=A\quad X_1\quad X_1*A$	傾斜が等しいかどうか検定できる

■反応曲面モデル 多項式回帰モデルでは定数項が1列目，X_1 の値が2列目，X_1 と X_2 の積が3列目に入る。この拡張として次のモデルを考える。

MODEL　$Y=X_1\quad X_2\quad X_1*X_2$
$X_1*X_1\quad X_2*X_2$

このモデルは X_1 と X_2 の2次曲面になる。X_1 を温度，X_2 を触媒量としYを生産量とすれば，最大値(最小値)を与える X_1 と X_2 の値で操業すればよいことになる。

■自己回帰モデル x_t を時刻 t の時系列の値とし，変数名を X_0 とする。m 時点前(タイムラグ m)の値を $X_m=x_{t-m}$ として，次の重回帰モデルを考える。これを自己回帰モデルという。予測の一手法である。

MODEL　$Y=X_0\quad X_1\quad X_2\quad X_3$

■非線形モデル 線形モデルにならない，$y=\beta_0 e^{\beta x}\varepsilon$ のようなモデルを非線形モデルという。しかし，両辺の対数をとれば次式になる。$\log y$ を y，$\log\beta_0$ を β_0，$\log\varepsilon$ を ε，

$$\log y=\log\beta_0+\beta_1 x+\log\varepsilon$$

と置き換えれば，単回帰モデルになる。このように置き換えのできないものは，非線形回帰推定を行わなければいけない。

■数量化Ⅰ類と多元配置 n 元(多元)配置で主効果のみを考えたものが数量化Ⅰ類である。分散分析法のプログラムで，アルゴリズムを拡張してパラメータの推定値も計算すれば，数量化Ⅰ類になる。ただし，たとえば主効果 B の3列を足せば $B_1+B_2+B_2=1$ と

いう制約(デザイン行列Xはフルランクでない)があり，最小2乗推定値$\boldsymbol{B}$が求まらない。この場合，一般最小2乗法($\boldsymbol{g}_2$逆行列)を用いるか，各主効果の1列をデザイン行列から省いて最小2乗推定を行えばよい。

3水準をもつ因子Bは3列に展開され，B_1列を省いて重回帰分析を行い，$\boldsymbol{\beta}_2$(B_2列)と$\boldsymbol{\beta}_3$(B_3列)の推定値を得た。$\beta_2\times0+\beta_3\times0=0$ であるから，

$\boldsymbol{B}$	$\boldsymbol{B}_1$	$\boldsymbol{B}_2$	$\boldsymbol{B}_3$	$\boldsymbol{\beta}_2$	$\boldsymbol{\beta}_3$	値
1	1	0	0	0	0	0
2	0	1	0	1	0	$\boldsymbol{\beta}_2$
3	0	0	1	0	1	$\boldsymbol{\beta}_3$

Bの第1水準の推定値は形式的に0とみなせる。逆に,始めから係数0がB_1列に与えられたものと考えてもよい。

以上の作為的に求めた推定値は，主効果内での相対的意味しかもたない。定数cを足して，$0\to c$，$\beta_2\to\beta_2+c$，$\beta_3\to\beta_3+c$ と変換してもよい。林の数量化では，データ平均がゼロになるように変換されている。**デザイン行列の構成2**に従えば，推定値の和はゼロ $(\beta_3=-\beta_1-\beta_2)$ になる。

■因子効果の意味　次の共分散モデルを考える。Aの推定値をa_1とa_2, Bの

MODEL　$Y=A\quad B\quad X$

推定値を$\boldsymbol{b}_1\boldsymbol{b}_2\boldsymbol{b}_3$として，定数項を$c$とする。$A=A_1$ で $B=B_2$ の場合のモデルY_1と $A=A_2$ で $B=B_3$ のモデルY_2は，次の別々のモデルを考えたことになる。

$$Y_1=(c+a_1+b_2)+\beta x+\varepsilon$$

$$Y_2=(c+a_2+b_3)+\beta x+\varepsilon$$

すなわち，データをAとBの6通りの水準組み合わせで層別し，連続変数Xの推定値は一定で定数項のみ異なる6組のモデルを考えたことに等しい。

■判別分析との関係　2群判別の場合，G_1群に1，G_2群に0の外的基準を与えて重回帰分析を行えば，判別分析が行える。このため，判別分析の変数選択も重回帰分析のアルゴリズムを利用できる。

■一般逆行列　次の条件のうち，

1．$AGA=A$
2．$GAG=G$
3．$(AG)'=AG$
4．$(GA)'=GA$

1を満たすGを一般逆行列，1と2を満たすものをg_2逆行列，1から4を満たすものをg_4逆行列という。

■最小2乗法　一般線形モデルのβは(一般)最小2乗法で求まる。行列表記を用いれば，誤差平方和 $\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}=(y-X\beta)'(y-X\beta)$ は,正規方程式 $X'X\boldsymbol{\beta}=X'\boldsymbol{y}$ の解$\boldsymbol{\beta}$で最小値をとる。Xがフルランクの場合，$(X'X)$ の逆行列が求まり，解は$\boldsymbol{\beta}=(X'X)^{-1}X'\boldsymbol{y}$ になる。ランク落ちする場合，たとえばg_2逆行列 $(X'X)^-$ を用いて，$\boldsymbol{\beta}=(X'X)^-X'\boldsymbol{y}$ を計算すればよい。連続変数であっても，変数間に制約条件(多重共線性)があれば，Xはランク落ちする。　(新村)

点推定

統計的推定の仕方に2通りある。1つは点推定であり，もう1つが区間推定である。得られた標本に基づいて，母集団分布の母数(パラメータ)が，この値であろうと推定しようとするのが点推定である。

★解説

母集団の**平均・分散**を，それぞれ母平均・母分散とよぶが，これらはいずれも母数である。母集団から得られた，大きさnの任意標本(X_1，X_2，…，X_n)に基づいて，たとえば母平均を，そのものズバリ推定しようとするのが点推定であるが，その推定の仕方にもいろいろある。母平均μの推定量のことを$\hat{\mu}$と書くが，標本から計算された**メジアン**をもって$\hat{\mu}$とすることも考えられるし，いちばん大きな値と小さな値に注目して，それらを加えて2で割った値を$\hat{\mu}$とすることもあろう。しかし，ふつうは**標本平均**$\overline{X}$を計算して，これを$\hat{\mu}$とする。すなわち，

$$\hat{\mu}=\overline{X}$$

それでは，なぜこのようにするのがよいか。それには，**推定量**のもつ望ましい性質は何か，について考えておく必要がある。この問題に関して，ふつういわれていることは，推定量が次の性質をもつことである。

(1) 不偏性 (2) 有効性 (3) 十分性 (4) 一致性

これらの性質をもつ推定量のことを，それぞれ

不偏推定量，有効推定量，十分推定量(充足推定量ともいう)，一致推定量

という。なお，上のX_i ($i=1, 2, \cdots, n$)や$\overline{X}$は，すべて**確率変数**であり，$\overline{X}=\hat{\mu}$とすることは，$\overline{X}$をもって，μの推定量とするということである。しかし，実際には，X_iの実現値(データ)として，x_iがとらえられるわけであって，データに基づく標本平均$\overline{x}$によりμの推定をするようなとき，$\overline{x}$のことをμの**推定値**とよぶ。

関連ページ
母集団 260
平均 344
分散 332
メジアン 272
標本平均 261
推定量 260
確率変数 56

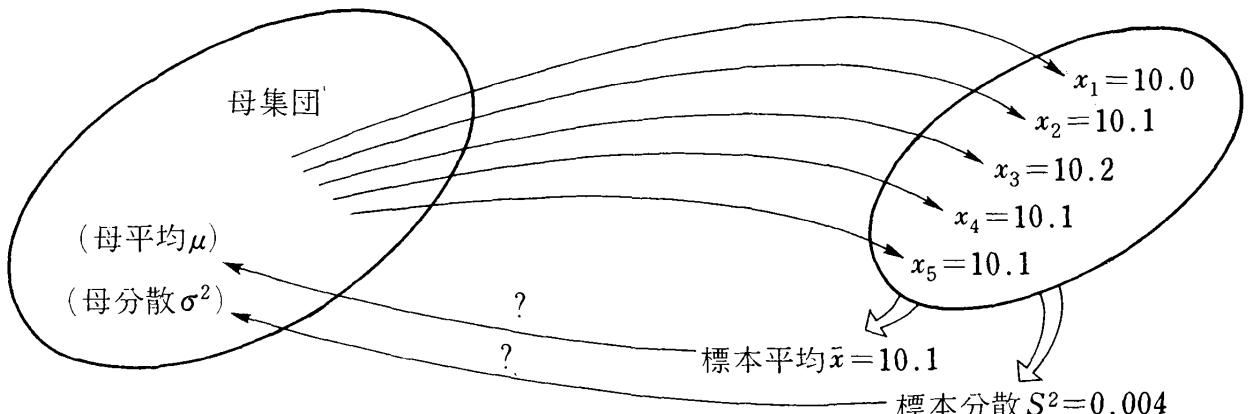


図1　平均・分散の点推定

■不偏推定量　いくつかの推定量が考えられるとき，その推定量の期待値(→58ページ)が，母数の値に等しくなることが望ましい。このような性質をもつ推定量を，不偏推定量という。たとえば，平均がμ，分散がσ^2である母集団からの，大きさnの標本$(X_1, X_2, \cdots, X_n)$に基づく標本平均$\overline{X}$については，$E(\overline{X})=\mu$が成り立つので，$\overline{X}$は母平均μの不偏推定量である。しかし，標本分散S^2については，

$$E(S^2)=\{(n-1)/n\}\sigma^2$$

になってしまうので，S^2はσ^2の不偏推定量にならない。これを修正して，

$$\{n/(n-1)\}S^2$$

を用いることにすれば，これがσ^2の不偏推定量になる。そこで，この統計量のことを不偏分散とよんで，標本分散と区別する。なお，ある母数の不偏推定量は1つとは限らないで，多数あるのがふつうである。たとえば，母平均の推定量として荷重平均

$$Y=\omega_1X_1+\omega_2X_2+\cdots+\omega_nX_n$$

$$\left(\begin{array}{l}\text{ただし，}\omega_i\geqq 0\ (i=1, \cdots, n)\text{ で}\\ \omega_1+\omega_2+\cdots+\omega_n=1\end{array}\right)$$

を考えると，$E(Y)=\mu$となるので，Yもμの不偏推定量である。

図1は，製品の山から5個とり出して寸法を測った例(単位；cm)を示している。$\bar{x}=10.1$は，母平均μの不偏推定値であるが，$S^2=0.004$は母分散σ^2の不偏推定値にならない。$n=5$であるから，$\{5/(15-1)\}\times 0.004=0.005$が$\sigma^2$の不偏推定値である。

■有効推定量　不偏性をもつ推定量のうち，分散最小のものが望ましい。このような推定量のことを，有効推定量という。上の荷重平均の例で，Yがμの有効推定量になるためには，$\omega_i=\dfrac{1}{n}$ $(i=1, 2, \cdots, n)$，つまり$Y=\overline{X}$でなければならない。

■十分推定量　母数を推定するのに，ある推定量があって，それ以外の統計量は，母数の推定に無関係であるとき，その推定量を十分推定量という。たとえば，$\overline{X}$はμの十分推定量である。

■一致推定量　標本数を大きくすることにより，いかほどでも母数の真の値に近いものが得られるような推定量のことを一致推定量という。　(牧野)

統計シミュレーション

統計分野でモンテカルロ法を用いて数表などを作成することを，統計シミュレーションという。コンピュータ処理能力の向上に助けられて，今後は実用的な問題解決のために威力を発揮すると考えられる。

★解説

■統計シミュレーション 気象·土木·医学分野などでは，分布の平均値等の代表値よりも，n サンプル中の大きい方から第 k 番目の順位にある値の分布とか，大きい方から k 番目と l 番目の値の差の分布とか，最大値のような極値の分布などが重要となることが多い。しかし，これらの分布は個別問題ごとに異なってくることが多いので，理論的に未検討なことが多く，コンピュータシミュレーションにより数表などを作成した方が実際的であることが多い。

上で述べた極値に関しては，グンベルらが極値統計学として理論展開しているが，その内容は一般的にいって難しい。シミュレーションはマシン時間がかかるが，このような理論的に難しいものを簡単に取り扱うことが可能である。

■ブート・ストラップ統計学 手もとにある n 個のサンプルから最大限の情報を引き出そうとするのが，**boot-strap** 統計学である。この n サンプルを **bootstrap sample** とよび，これを十分なだけコピーして増やして母集団とみなす。この母集団から乱数を用いて n 個のサンプリングを多数回行い，その分布特性を調べる。たとえば，10 個のサンプルを 100 回コピーすれば，1000 個データが作られる。〔0, 1〕の一様乱数で小数第 3 位までもとめて，区間〔0.000, 0.001〕の値であれば一番目のデータをサンプリングすればよい。名称の由来は，コンピュータの小さなブート・ストラップルーチンが次々に大きなローダ(ファイル)をロードし，増えていく様子からつけられた。

関連ページ

シミュレーション 166
シミュレーション言語 168
モンテカルロ法 388
SAS 136

■極値統計学　世の中の現象には，必ずしも平均値や最頻値などの代表値でとらえきれないものが多い。たとえば，ダム工事や河川の提防建設に際しては，最大雨量や最大流量を予測して対策をたてなければならない。極値統計学はこのとき問題となる最大値や，最小値などの分布を論じる。

グンベルらにより理論展開はされているが，実用的な個別の数表などの作成は統計シミュレーションを利用すればよい。

■正規分布の最大値や範囲の性質　平均0，分散1の正規分布 $N(0, 1)$ から100個のデータをサンプリングし，最大値,最小値,範囲幅,平均値,標準偏差値の分布を調べる。正規乱数を1万個発生させ，100個ごとに上記の統計量を計算し，その値の100個の組の分布を調べる。SASのDATAステップで次のようにプログラミングできる。

```
① DATA  a ;
② DO  j=1  TO  100 ;
③   DO  i=1  TO  100 ;
④   ARRAY  x(i)  x1-x100 ;
⑤   x=NORMAL (1 2 3 5 7) ;
⑥   END ;
⑦   h=MAX (OF  x1-x100) ;
⑧   l=MIN (OF  x1-x100) ;
⑨   r=RANGE (OF  x1-x100) ;
⑩   m=MEAN (OF  x1-x100) ;
⑪   s=STD (OF  x1-x100) ;
⑫ OUTPUT ;  END ;
⑬ PROC  MEANS ;
```

④は100個の変数X1からX100に配列名Xを与えている。⑤の1 2 3 5 7は正規乱数の初期値である。③から⑥のループ内で100個の正規乱数が配列XすなわちX1からX100に割り当てられる。⑦から⑪でX1からX100までの値の最大値，最小値，範囲，平均値，標準偏差値が，変数H, L, R, M, Sに割り当てられる。②から⑫のループにより100回繰り返し計算が行われ，⑫のOUTPUT文でX1からX100と上記の5変数がSASデータセットに100レコード出力される。⑬では,この100サンプルの基礎統計量を求めている。

次に結果を示す。

	平均	標準偏差	最小値	最大値	平均の標準誤差
H	2.516	0.381	1.805	3.719	0.038
L	−2.510	0.421	−3.811	−1.592	0.042
R	5.026	0.547	3.697	6.213	0.055
M	0.015	0.101	−0.225	0.279	0.010
S	1.002	0.062	0.835	1.145	0.006

表1　正規分布の極値などの諸統計量（100サンプル）

以上の統計量はヒストグラムを描けば一層明確になるが，最大値と最小値は最大または最小の方向へ偏った分布になる。一方，範囲と平均と標準偏差は当然予知できたことだが，左右対称の分布になっている。

100個の正規乱数の最大値は平均2.516で，その最大値は3.719である。

（新村）

統計調査

集団について，その特徴を数量的に表す統計を求めるための調査のことを統計調査という。統計調査は国勢調査を典型として始まったものであるが，統計調査が簡便な方法で行われる場合，アンケートというよび方もされている。

★解説

■全数調査と標本調査　研究の対象として考えられている集団の単位全体の集合である**母集団**を構成しているすべての単位について調査することを**全数調査**という。これに対し，母集団の一部を抽出して調査することを標本調査という。統計では，多くの資料を調査し，現状の把握や将来の推測に役立てているが，そのためには統計資料は信頼性の高いものであることが大切である。しかし，調査には時間や費用がかかる。そこで，調査の目的・調査項目・調査対象・期間・調査人員・経費などに応じて，全数調査によるか標本調査によるかを，きめてかかる必要がある。ここで，とくに注意しなければならないことがある。それは，調査には誤差がつきものだということである。たとえ全数調査であっても，ある程度，誤差が入り込むことは避けられない。まして標本調査ということになると，

[標本調査における誤差]
＝[標本抽出にともなう誤差]＋[標本抽出以外の誤差]

であるから，標本抽出にともなう誤差が加重される。それでは，調査誤差を小さくするためには，必ず全数調査を実施したほうがよいかというと，そうとは限らない。「標本抽出誤差」よりも「抽出以外の誤差」のほうが，はるかに大きいのがふつうであって，標本調査では，この「抽出以外の誤差」を，かえって小さくすることができる，という場合も少なくないからである。

関連ページ
母集団　260

- 統計調査
  - 全数調査
  - 標本調査
    - 確率抽出法
    - 有意抽出法

表1　統計調査

調査方法	長　所	短　所
全数調査	正確な回答が期待できるならば完全な調査である。	費用がかさむ。 集計時間が長びく。
標本調査	あらかじめよく検討しておいて調査を行えば，調査や集計が手軽にできて，かなりよい結果が得られる。	全数ではなしに，一部を抜き出して調査するので，標本抽出にともなう誤差が入りこむ。

表2　全数調査と標本調査の比較

■調査誤差　主要な誤差として，次のようなものが考えられる。

(1) 調査対象や調査項目の選び方，定義の仕方などが，調査の目的と一致しないために生ずる誤差。

(2) 多くの調査では，あらかじめ調査対象をリストし，これに従って調査するが，調査すべき対象がリストから漏れていたり，対象にならないものがリストに記載されていたりする。このように，リストの不完全さに起因して誤差が生ずる。

(3) 被調査者は，いつも正確な解答をしてくれるとは限らない。また，たとえ実測調査であっても，測定には必ず誤差をともなう。

(4) 集計方法のミスや誤記による誤り。

■標本のとり方　標本調査における標本のとり方に，**有意抽出法**と**確率抽出法**とがある。有意抽出法では，調査担当者が意図的に標本を選ぶ。その際，できるだけ典型的と思われる標本を抽出して調べようとする方法を**典型法**という。この方法による調査は，客観性に欠けるきらいがある。この欠点を，いくぶん除去する方法が**割当法**である。割当法では，母集団について事前にわかっている特性に関して，母集団での比率と同じになるように標本を構成していこうとする。しかし実際には，そのように近似させていくことは困難なので，あまりよい抽出法とはいえない。確率抽出法には，そのような難点がない。

(牧野)

統計的推測

集団から一部のものを抜きとって調べ，それに基づいて客観的に，集団全体の特徴を知ろうとする。これが統計的推測であり，その理論や方法を研究するのが推測統計学である。

★解説

■標本と母集団 上で，抜きとられる一部のものを標本(サンプル)といい，集団全体のことを母集団という。推測統計学の特徴は，標本と母集団とを明確に区別することにある。たとえ全数調査であったとしても，母集団の一般的性質を推測するために，それは観念的には，やはり標本として扱われる。

■無作為抽出 標本の選びかたは，ふつう無作為抽出(任意抽出，ランダム・サンプリングともいう)による。それは，母集団の抽出単位の大きさがNであるとき，その中からn個の標本を，等しい確率で抽出する方法である。そのために，たとえば母集団の各抽出単位に，1からNまでの番号をつけておき，n個の**乱数**に対応する抽出単位を標本にとる，という具合にすればよい。

■標本分布 一般に，標本 $(X_1, X_2, \cdots, X_n)$ をn個の**確率変数**の組とみなしたとき，その関数を統計量とよび，統計量の**確率分布**を標本分布という。標本が無作為抽出によって選ばれているならば，そこに標本分布の知識を適用して，客観的な統計的推測を行うことができる。

■統計的推測の仕方 大きく2つに分かれる。1つは**統計的推定**であり，もう1つは**仮説検定**である。母集団の**特性値**や母集団分布が何であるかを推しはかるのが推定であり，その仕方はまた，**点推定**と**区間推定**の2つに分かれる。これに対し，母集団の特徴が，あらかじめ考えられている仮説に従うとみなしてよいかどうかを調べるのが，仮説検定である。

関連ページ
乱数 388
標本分布 314
確率変数 56
確率分布 56
仮説検定 62
特性値 272
点推定 254
区間推定 82

■**統計量と標本分布**　大きさ n の標本 $(X_1, X_2, \cdots, X_n)$ は，n 個の確率変数の組と考えられるが，これをもとにして，たとえば，

標本平均 $\overline{X}=(X_1+X_2+\cdots+X_n)/n$

標本分散 $S^2=\{(X_1-\overline{X})^2+(X_2-\overline{X})^2+\cdots+(X_n-\overline{X})^2\}/n$

などの**統計量**がつくられる。**標本平均**や**標本分散**は，**推定**や**検定**に用いられる基本的な統計量であるが，ORの分野などでは，i 番目の大きさのものを考えたりすることもある。これを**順序統計量**という。

いずれにしても，統計量は確率変数の関数なので，これもやはり確率変数である。これらを推定や検定に用いるためには，それらの確率変数の従う確率分布を調べておく必要がある。このような確率分布のことを標本分布という。上の例でいえば，標本平均の分布，標本の分散の分布，順序統計量の分布など，いずれも標本分布である。標本分布に対応する理論分布も標本分布とよばれるが，それらの中で，正規分布がはたす役割は大きい。そのほかによく用いられる標本分布としては，χ^2 分布(カイ2乗分布→314ページ)，t 分布(→314ページ)，F 分布(→315ページ)などがある。

■**推測統計学を創った人々**　記述統計学を創った研究者の1人であるK.ピアソンは，ピアソン系分布曲線・積率法・χ^2 検定法などを考案し，また有力な統計学雑誌であるBiometricaを創刊するなど，その後の推測統計学の確立・普及にあずかって力があった。

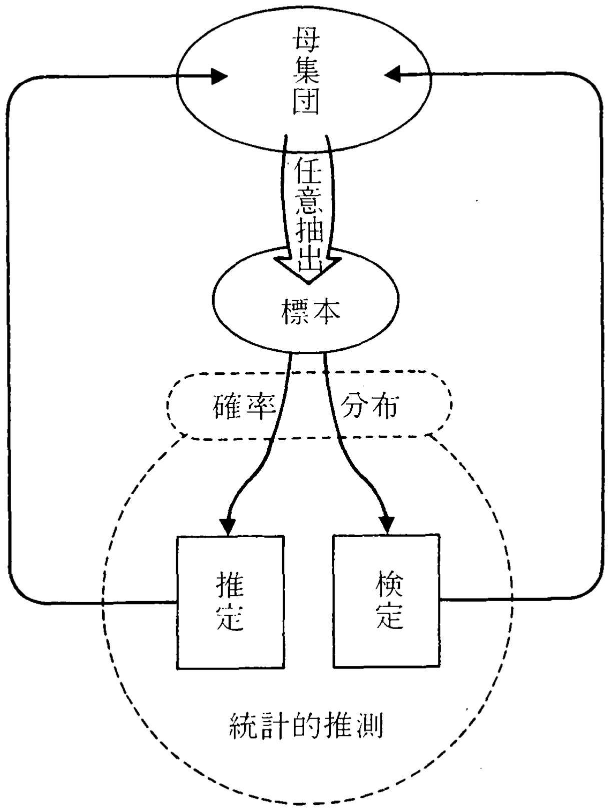


図1　統計的推測の考え

実際に推測統計学の発展をみるのは20世紀にはいってからであるが，これにはロンドン郊外にあるローザムステッドの農事試験場が大いに貢献した。圃場試験では，地力の変化などのために，小標本しか得られない。このことに着目したゴセットが，小標本の統計的推測に欠くことのできない t 分布を考案した。これに興味をもったフィッシャー(R. A. Fisher)は，さらに進んで F 分布を考案し，圃場試験での誤差処理に関して考えられていた実験配置法を，実験計画法などの形で確立した。　(牧野)

統計的品質管理

統計的品質管理はSQCと略称されることが多い。JISでは，これを次のように要約している。

「近代的な品質管理は，統計的な手段を採用しているので，とくに統計的品質管理ということがある」

統計的品質管理は，1930年前後にシュハートが初めて統計学を品質管理の目的に活用して品質管理を体系化して以来，長い間，広く役立てられている。

★解説

統計的品質管理(以下SQCという)は決して難解な統計学を用いることを考えてはおらず，むしろ，なるべく平易な手法を用いて，技術的に解決困難な問題を分析することを目的としている。

このため，工場のデータを収集して**ヒストグラム**を作り，製品の品質特性の分布の状態を把握することから始まり，さらに，**特性要因図**を作り，これに対して**パレート図**を作ることにより，特性要因図中の各項目に重みをつけることなどが行われる。また，工程を安定せしめる方法としては**管理図**が用いられ，品質の保証を確実にするためには**抜取検査**が適用される。

一方，SQCでは，上のような簡単な手法だけではなく，技術的に高度な問題に対しては，**実験計画法**や**多変量解析**なども応用されることが少なくない。また，最近は信頼性データの解析を目的として**ワイブル確率紙**を用いた手法もよく用いられる。

反面，**FMEA**，**FTA**，および**品質展開**などはほとんど統計的手法を用いない。SQCとは対照的な手法といえよう。

関連ページ

ヒストグラム 274

パレート図 300

管理図 68

抜取検査 108

実験計画法 160

ワイブル確率紙 419

FMEA，FTA 36

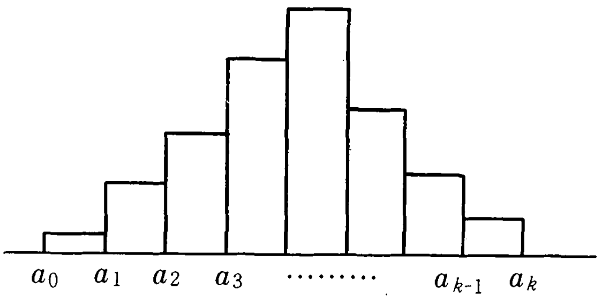


図1　ヒストグラム

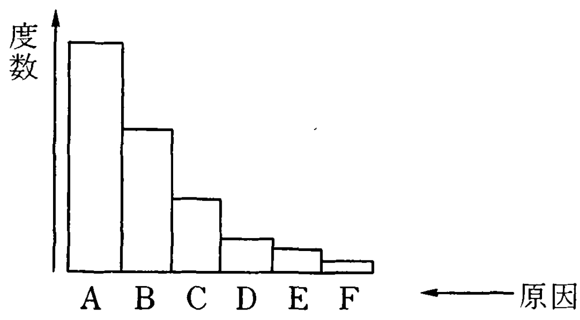


図3　パレート図

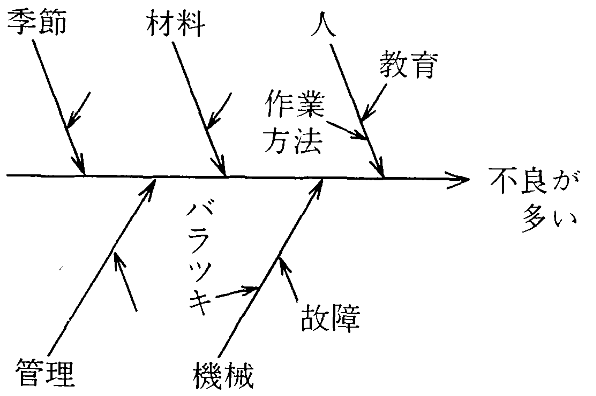


図2　特性要因図

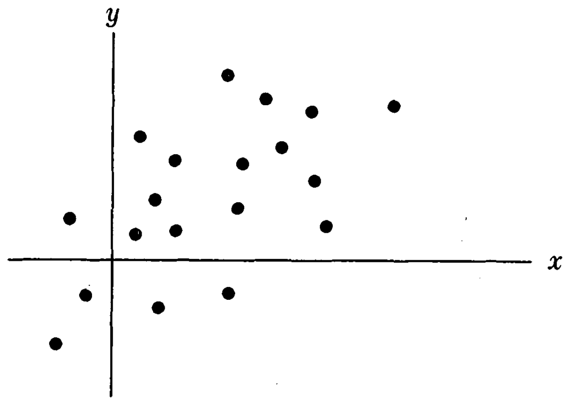


図4　散布図

■**統計的品質管理の進め方**　このSQCを進めるときに必要な手段を以下に列記して説明する。

(1)**ヒストグラム**　図1のようにデータを柱状の図表としたもので，データ分布の形を見るのに適している。

(2)**特性要因図**　図2のように魚の骨状にした問題項目に対する原因を整理することを目的としたグラフである。

(3)**パレート図**　図3のようにトラブルの数の大きさの順に原因を並べたグラフをいう。

パレート図はイタリアの経済学者パレートの名に由来して名づけられたという。一部，経済・経営の分野ではABC分析ともいう。

(4)**散布図**　相関や回帰を論ずるときに，2次元のグラフ上の2次元データ(x_i, y_i)をプロットし，xとyの関係や各データの特性を検討するのに用いる。

(5)**二項確率紙**　主として不良率などの解析を容易ならしめるために工夫された確率紙であって，xが二項分布$B(x; n, p)$にしたがうとき，これを逆正弦変換し，

$$\sqrt{n}\,\theta=\sqrt{n}\sin^{-1}\sqrt{\frac{x}{n}}$$

とすると，この$\sqrt{n}\,\theta$は近似的に正規分布$N(\sqrt{n}\sin^{-1}\sqrt{p},\ \frac{1}{4})$となることを利用したものである。これを用いると，不良率の区間推定および簡易相関分析などが，この確率紙の上で簡単に行うことができる。〔4〕〔5〕〔41〕

(真壁)

統計パッケージのファイル管理(SASを例として)

統計パッケージの多くは，データ管理を処理部分から切り離して標準的なファイルとして扱うことが多い。これにより，ファイルの運用管理が1元的に行え，また処理部分のデータの入出力の標準化が計れる。SASでは，SASデータセットとよばれる表形式のファイルが処理の単位であり，SASデータベースのメンバーとして1元管理できる。

★解説

■外部ファイルとシステムの標準ファイル　カード・MT・ディスクなどに任意の形式で登録されたデータを外部ファイルという。普通はシーケンシャル・ファイルのことが多いが，階層型ファイルなどの形式をとることもある。これらの外部媒体上のデータを，右ページの表形式のファイルとして一定の形式で作成したものをシステムの標準ファイルという。SASではSASデータセット，BMDPではSAVEファイルなどという。汎用化を計るためSASでは，実数と整数データをすべて倍精度データに標準化しているため，データ量も増えて無駄も生じるが，処理モジュールの入出力が標準化できる。

■SASデータベース(SAS DB)　各種のファイルを，1元管理する目的で集めたものをDBという。DBのタイプには**木構造**，**網構造**，**関係型**がある。統計処理は右ページの表1に示す表形式ファイルを処理単位としているため，関係型が適している。SASでは表形式のファイルをSASデータセットとよび，DBのメンバーとして1元管理できる。管理の内容は，DBと外部ファイルの入出力管理(レポートライティングを含む)，メンバーの内容変更，メンバー間の連結・併合・更新，メンバーの属性変更などである。

関連ページ
汎用統計パッケージ　306
SAS　136
データベース　246

var. / obs.	x_1	x_2	$\cdots$	x_p
1	x_{11}	x_{12}	$\cdots$	x_{1p}
2	x_{21}	x_{22}	$\cdots$	x_{2p}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
n	x_{n1}	x_{n2}	$\cdots$	x_{np}

表１　データの構造

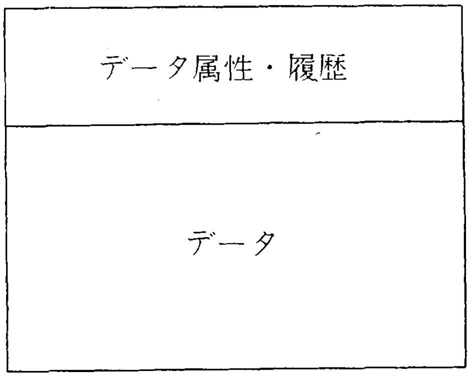


図１　SASデータセットの構造

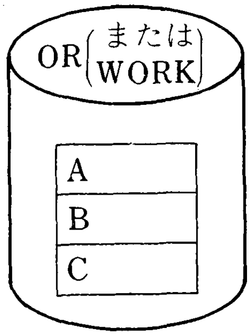


図２　SASデータベース

■データの構造と変数のタイプ　統計処理の扱うデータは，表1のような表形式を原則としている。行側は，標本または観測対象とよび，列側は変数または項目を表す。

外部ファイルでは，表の各行は，カードやMTの1レコード(複数のこともある)に対応することが多い。変数の値は，文字タイプ・整数タイプ・実数タイプに分けられる。質的データは，文字タイプと整数タイプのいずれでも表される。

SASでは，この表形式データをデータセットとよぶ。データセットには，図1に示すように変数名や書式等のデータの属性と履歴が付加されている。このため，以下の処理ではデータセット名だけで引用できるので，変数の書式指定のわずらわしさから解放される。数字項目はすべて倍精度に変換されるので，計算精度に問題が起こらない。文字項目は200文字まで扱える。変数 p は4000までと制限があるが，サンプル数 n には制限はない。欠測値の取り扱いも柔軟である。

■DB　図2に示すように，データセットA, B, Cの集合をDBという。この例ではDB名がORであり，OR.AでもってDBのメンバーであるデータセットが参照できる。データを保管する必要のないときは，その処理の間だけ通用するDB(WORK)を用いればよい。WORK.AはAと省略できる。

■ファイル管理　外部ファイルを入力しSASデータセットの生成や，データセットの内容をレポートライティングしたり，外部ファイルに出力する以外に，データセットの間でリレーショナルDBに似た次の処理が行える。

データセットAに変数を追加したり，条件に合致するデータを抽出し，新しくデータセットBとする。データセットCと同じ項目をもつデータセットDを後にくっつけてデータセットEとする。同じ対象の異なった項目がデータセットFとGに分かれている場合，合併してデータセットHとする。データセットM1の内容の一部を修正ファイルTで修正し，データセットM2とする。　(新村)

投資案の感度分析

不確実な状況下での投資案の経済性を検討するのに頼りになるのが，感度分析をしてみることである。不確実な要因の数値を動かして，利益に対する感度を調べる方法が感度分析である。とくに，予測がわるいほうにはずれた場合，利益がマイナスになったり，目標利益額を割るおそれがないか，という一種の安全性(抵抗力)を調べておくことが有用となる。

★解説

一般に，投資案の**キャッシュ・フロー**をきめる要因として次のものをあげることができる。

(1) 需要量
(2) 価格
(3) 投資案(製品)の寿命
(4) 操業費用
(5) 投資額
(6) 利子率

これらの要因の値を推定して投資案の利益(正味の得)を計算する際に，各要因の推定値のズレが利益に及ぼす影響も調べておくと，意思決定の参考になる。一定額(あるいは率)の利益の変動をもたらす特定の要因の変動率の逆数のことを一般にその要因の感度とよんでいる。すなわち，ある要因の値が少し変動するだけで，利益の額が大きく変わるとき，その要因の感度は高いことになる。

感度の高い要因がわかれば，その要因についてさらに精度の高い予測をしてみるとか，要因の値がわるい方向に動かないような手を打つ準備をあらかじめしておくことが可能になる。

一般的にいうと，製品の価格や需要量の感度は高くなることがわかっている。したがって，製品市場の動向にたえず注意を払わねばならない。

関連ページ
方策のキャッシュ・フロー　358

投資案の感度分析の例を示す。ある鉱山を買収して精製工場を建設し，毎年10万トン以上の売り上げが見込まれている事業計画がある。初期投資額は鉱山の買収，工場建設，設備の購入など合わせて500億円かかる。操業費用はトン当たり平均2万円で，このほかに年間固定経費が10億円かかると予想されている。製品の価格はトン当たり10万円，投資の寿命は約20年が見積もられている。資本の利率は10%である。

もし予定通りの売り上げが生じるとすると，年当たりの収益は次式になる。

R=10万トン×10万円/トン=100億円

また初期投資の年平均値を求めると，

$$C=500\text{億円}\times[P\rightarrow M]_{20}^{10\%}$$
$$=500\times0.11746$$
$$=58.73\ \text{億円}$$

になる。年当たりの平均利益(年価)は次式になる。

$$\text{利益}=100^{\text{億円}}-2^{\text{万円/トン}}\times10^{\text{万トン}}$$
$$-10^{\text{億円}}-58.73^{\text{億円}}$$
$$=11.27\ \text{億円/年}$$

■感度の計算　このような投資案の感度を求めるのに，おのおのの要因の損益分岐点を計算してみるのが有効である。表1に示す記号を用いて，各要因の損益分岐点の計算式をあげておこう。

$$\text{価格の分岐点}=v+\frac{F}{D}+\frac{I}{D}\times[P\rightarrow M]_n^i$$

要因と記号	損益分岐点	感度
価格(p)	8.87万円/トン	8.85
販売量(D)	8.59万トン	7.89
操業費(V)	3.13万円/トン	1.77
固定経費(F)	21.27億円	0.89
投資額(I)	595.95億円	5.21
寿命(n)	14　年	3.33

表1　損益分岐点と感度

$$\text{販売量の分岐点}=\frac{1}{p-v}\times\{F+I[P\rightarrow M]_n^i\}$$

トン当たり操業費の分岐点

$$=p-\frac{F}{D}-\frac{I}{D}\times[P\rightarrow M]_n^i$$

$$\text{固定経費の分岐点}=(p-v)\times D-I\times[P\rightarrow M]_n^i$$

初期投資額の分岐点

$$=\{pD-vD-F\}/[P\rightarrow M]_n^i$$

寿命の分岐点はいわゆる回収期間であり，利子率の分岐点は利回りになる。この例では回収期間は14年，利回りは12.7%になる。

計算結果は表1の通りである。

ここで感度は次式で求めた値である。

$$\text{要因の感度}=\frac{\text{推定値}}{|\text{推定値}-\text{分岐点}|}$$

表1をみると，価格，販売量，投資額の感度が大きくなっている。これらの要因の推定値を再検討して，不確実性にそなえた方がよいことがわかる。　(中村)

投資案の経済性指標

投資案の経済性を判断するのによく用いられる指標は，大別してつぎの３種類ある。

(1) 投資案のリターンから投資額を差し引いて得られる「正味の得」の金額を表す指標。つぎの３つの表し方がある。

①正味現価　②正味終価　③正味年価

(2) 投資案のリターンの投資額に対する「率」の指標。投資効率を表すもの——利回り(利益率)。

(3) 投資案のリターンが投資を回収する「期間」を表す指標——回収期間。

★解説

投資案の初期投資額をC，投資の効果が及ぶ期間(投資案の寿命)をn，おのおのの期末のリターン額をR_1，R_2，…，R_n，計算利率をiとして，それぞれの指標を説明する。

(1)**正味現価**(net present value)

$$P=-C+\frac{R_1}{(1+i)}+\frac{R_2}{(1+i)^2}+\cdots+\frac{R_n}{(1+i)^n}$$

(2)**正味終価**(net final value)

$$S=-C\times(1+i)^n+R_1\times(1+i)^{n-1}+R_2(1+i)^{n-2}+\cdots+R_{n-1}\times(1+i)+R_n$$

(3)**正味年価**(net adjusted annual value)

$M=P\times$[資本回収係数]

(4)**利回り**(rate of return on investment)　投資を回収しきる利率rすなわち，

$$C=\frac{R_1}{(1+r)}+\frac{R_2}{(1+r)^2}+\cdots+\frac{R_n}{(1+r)^n}$$

(5)**回収期間**(pay-back period)　回収がすむ，すなわち，

$$C\leqq\frac{R_1}{(1+i)}+\frac{R_2}{(1+i)^2}+\cdots+\frac{R_k}{(1+i)^k}$$

となる期間N。

関連ページ

資金の時間的価値　144

資金の時間的価値の換算係数　146

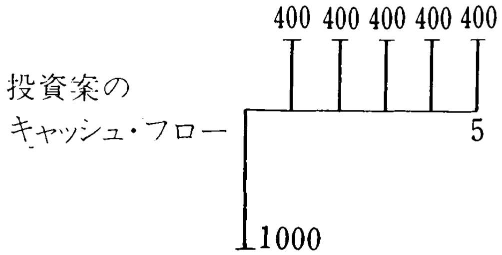
年価法の役立ち　286

利回り法　406

回収期間法　52

単一の投資案をおのおのの指標で判断する方法を数値例で示す。

■数値例 投資額が1000万円で，5年にわたり毎年末に400万円のリターンがあがる投資案を考える。この投資案はリターンが毎年等額で与えられているので，換算係数(→146ページ)を使って計算できる。$i=10\%$ とする。



(1)正味現価P

$=400\times[M\rightarrow P]_5^{10\%}-1000$

$=400\times3.791-1000$

$=516.3$ (万円)

(2)正味終価S

$=400\times[M\rightarrow S]_5^{10\%}-1000\times[P\rightarrow S]_5^{10\%}$

$=400\times6.105-1000\times1.611$

$=831.5$ (万円)

(3)正味年価M

$=400-1000\times[P\rightarrow M]_5^{10\%}$

$=400-1000\times0.264$

$=136.2$ (万円/年)

これらはすべて正味の得を表しており，正の値になっているのでペイする投資案になる。実際には，この3種類の指標のうち1つについて計算すればよい。

(4)利回り(r)

$$1000=400\times[M\rightarrow P]_5^r$$

を満たすrが利回りになる。この式をつぎのように変形する。

$$[P\rightarrow M]_5^r=\frac{400}{1000}=0.4$$

資本回収係数［$P\rightarrow M$］の数表で $n=5$ のところをみると，$i\fallingdotseq29\%$ でほぼ2.5の値になっている。したがって，この投資案の利回り r は29%になる。

$$r=29\%>i=10\%$$

となり，ペイすることがわかる(→438ページ)。

以上に示した4つの指標は，投資案がペイするか否かについて同一の結果を導くので，どの指標を用いてもよいことになる。

(5)回収期間(N)

各年度での回収額の現価を計算すると次のようになる。

年度	回収額
1	$400\times[M\rightarrow P]_1^{10\%}=$ 363.6
2	$400\times[M\rightarrow P]_2^{10\%}=$ 694.4
3	$400\times[M\rightarrow P]_3^{10\%}=$ 994.8
4	$400\times[M\rightarrow P]_4^{10\%}=$ 1268.0
5	$400\times[M\rightarrow P]_5^{10\%}=$ 1516.4

この結果，4年度には1000万円の投資を回収しきることになり，回収期間は4年になる。

もし基準になる回収年数が3年と設定されていれば，この投資案は不合格，5年と設定されていれば合格になる(→52ページ)。 (中村)

動的計画法

多段階決定問題を解くときに有力なひとつの算法である。最適性の原理を利用して，問題を各段ごとの関数方程式を解く問題に帰着させる。生産在庫管理，広告費配分問題，探索問題など，いろいろな分野で応用される。また理論的には変分法とも関連があり，マルコフ連鎖と結びついてマルコフ決定過程も生み出している。ベルマン(R. Bellman)の創始による(1957)。

★解説

■多段階決定過程　第 t 期の状態が x_t のときに決定 d_t をとると，第 $t+1$ 期の状態は $x_{t+1}=h_t(x_t,\ d_t)$ となり，$z_t=g_t(x_t,\ d_t)$ の利得が得られるものとする。このとき N 期間における総利得 $z=z_1+\cdots+z_N$ を最大にするような各期の決定を求めるのが**多段階決定問題**である。ある期における決定はそれ以後の状態に影響を及ぼすので，そのことを考慮の上，各期の決定はなされなければならない。

■動的計画関数方程式　多段階決定問題は，つぎのような関数を導入することによって，N 個の関数方程式に帰着させることができる。

$$f_t(x)=\left(\begin{array}{l}\text{第 } t \text{ 期の状態が } x \text{ であったときに，} t \text{ 期}\\ \text{以降ずっと最適決定をとることにより得}\\ \text{られる第 } t \text{ 期から第 } N \text{ 期までの利得の和}\end{array}\right)$$

第 t 期の状態が x であったとして決定 d をとると，第 $t+1$ 期の状態は $h_t(x,\ d)$ となる。したがって，第 $t+1$ 期から第 N 期までに得ることができる利得の最大値は，上の定義から $f_{t+1}(h_t(x,\ d))$ となる。したがって，$f_t(x)$ はつぎの関数方程式をみたす。

$$f_t(x)=\max_d \{g_t(x,\ d)+f_{t+1}(h_t(x,\ d))\}$$

右辺で最大値をとる d が状態 x での最適決定である。

関連ページ
数理計画法　186
マルコフ決定過程
　380

■最適性の原理　動的計画関数方程式を導く際に，つぎのような最適性の原理を暗黙のうちに使っている。ここで**政策**といっているのは，第 t 期にどの状態にいればどの決定をとるか，を一意的に定めた規則のことである。

> N期間問題の最適政策は，問題を第 t 期から第N期までの $N-t+1$ 期間問題に縮小しても，やはり最適である。

この最適性の原理が満たされるならば，目的関数は z_t の和である必要は必ずしもなく，積であったり，ほかの単調性をもった関数であっても差し支えない。

■生産在庫管理(生産平滑化)の例　動的計画法の応用例の1つとして，生産在庫管理の問題をとりあげる。ある工場では，今後N期間の生産計画をたてようとしている。各期の需要 D_t $(t=1,\ \cdots,\ N)$ は既知ではあるが，変動は激しい。一方，各期の生産量 s_t もある程度変えることができるが，人員の関係などから変動は少ないほうが望ましい。このことは第 t 期の生産費用 $p_t(s_t,\ s_{t-1})$ が s_t ばかりでなく s_{t-1} にも依存することで表される。

各期の需要に対して商品が不足することは許されないが，前もってその一部を生産しておくことはできる。このときの在庫保管費は，第 t 期末の在庫量を u_t とすると，$c_t(u_t)$ で与えられる。第0期の生産量 s_0 と期末在庫量 u_0 は既知であり，第N期末に残った商品 u_N 個には $v(u_N)$ だけの価値が見込まれる。

N期間の総費用を最小とする生産計画を立てるにはどうしたらよいか。

■動的計画法による定式化　この問題では，決定変数は各期の生産量 s_t で，状態変数は前期の生産量と期末在庫量の組 $(s_{t-1},\ u_{t-1})$ である。このように，状態変数がベクトルとなることもある。$u_t=u_{t-1}+s_t-D_t$ であるから，状態変換関数 h_t は，

$$h_t(s_{t-1},\ u_{t-1},\ s_t)$$
$$=(s_t,\ u_{t-1}+s_t-D_t)$$

となる。また，利得関数は費用関数と考えたほうが都合がよく，

$$p_t(s_t,\ s_{t-1})+c_t(u_{t-1}+s_t-D_t)$$

となる。

$f_t(s,\ u)$ を，第 t 期の状態が $(s,\ u)$ であったときにかかる第 t 期から第N期までの総費用の最小値とすると，

$$f_t(s,\ u)=\min_{s'}\{p_t(s',\ s) + c_t(u+s'-D_t) + f_{t+1}(s',\ u+s'-D_t)\}$$

という関数方程式が得られる。ただし，$t=N$ のときは右辺の{ }内は，

$$p_N(s',\ s)+v(u+s'-D_N)$$

となる。

f_Nは既知関数なので，$t=N,\ N-1,\ \cdots,\ 1$ と逆順に関数方程式を解けば，解が求められる。参考文献〔29〕〔47〕。　(高橋)

特性値

統計資料は，基本的には分布型として表現されるが，実際的な扱いとしては，分布型の特徴を表す測度である特性値(母数ともいう)に注目して分析するのがふつうである。このため，資料全体の中心的な位置を示す量である代表値，広がりを示す量である散布度が用いられる。代表値としてよく用いられる量は平均値であるが，メジアン・モードなどを用いることもある。散布度としては分散・標準偏差がよく用いられるが，その他，範囲や四分偏差なども用いられる。

★解説

■平均値　データの値すべてを加えて，それをデータの個数で割ったものが平均値である。**推測統計学**における論理構成の上からは，平均値が最も重要な代表値である。

■メジアン　データを大きさの順にならべたとき，ちょうど中位にあるデータの値を**メジアン**(中位数)または**中央値**という。

■モード　データの中で，いちばんひんぱんに現れるものを**モード**という。モードのことを**最頻値，並み数**ともいう。

■分散と標準偏差　データが,平均値のまわりにどの程度集中しているか,いいかえれば,バラツキの程度がいくらか,をみるのに,**分散**や**標準偏差**が用いられる。n個のデータ

$$x_1,\ x_2,\ \cdots,\ x_n$$

から分散を計算するには，平均値$\overline{x}$（エックス・バーと読む)を計算しておいてから，$\dfrac{1}{n}\{(x_1-\overline{x})^2+\cdots+(x_n-\overline{x})^2\}$を求めればよい。標準偏差は，このようにして求められた分散の平方根である。

■範囲　データの中のいちばん大きな値から，いちばん小さい値をひいたものを**範囲**という。

■四分偏差　データを大きさの順にならべたとき，下位から4分の3にあたる値(第3四分位数)と4分の1にあたる値(第1四分位数)との差の1/2を**四分偏差**という。

関連ページ
平均値　344
推測統計学　260
分散　332

資料が階級に分けられていて，x_1の値のものがf_1個, x_2がf_2個，……，x_kがf_k個であるとし，資料全体の個数は，

$$f_1+f_2+\cdots\cdots+f_k=N$$

とする。このとき，

代表値
- 平均値；$\bar{x}=\dfrac{1}{N}(x_1f_1+\cdots\cdots+x_kf_k)$
- メジアン；資料を大きさの順にならべたとき，ちょうど中央に位置する資料の値。
- モード；資料の中で，いちばん頻度の高い値。

散布度
- 分散；$S^2=\dfrac{1}{N}\{(x_1-\bar{x})^2f_1+\cdots\cdots+(x_k-\bar{x})^2f_k\}$
- 標準偏差；$S=\sqrt{\text{分散}}$
- 四分偏差；$(Q_3-Q_1)/2$

（ただし，Q_1は第1四分位数，Q_3は第3四分位数）

■計算上の注意　代表値のなかで，もっとも重要な位置を占めるのは平均値であり，散布度のなかでは分散，及びその平方根である標準偏差である。これらの算法については，それぞれの項目を参照されたい。ここでは，そのほかの代表値，散布度を計算する際の注意を記しておく。

■メジアン　データの個数nが奇数のときには，データを大きさの順にならべて，$(n+1)/2$ 番目がちょうど中位になるので，その値をもって，メジアンとすればよい。しかし，nが偶数で，$2m$ となるような場合には，m番目の値と $(m+1)$ 番目の値の平均値をメジアンとする。メジアンは，「それより大きな値をもつデータの数と，それより小さな値をもつデータの数とが同じ」という点で，代表値として受けいれやすい量である。

■モード　上の例のように，データが階級分けされているとき，頻度 $f_i(i=1, 2, \cdots, k)$ のもっとも大きい階級の(上限値＋下限値)/2 をモードと定めるのがふつうであるが，これは階級の分け方で多少，左右されてしまうことを注意しなければならない。

■四分偏差　メジアンの場合と同様に，第1四分位数 Q_1，第3四分位数 Q_3 を算出する。Q_1 についていえば，それ以下の値をもつデータが全体の 1/4 であるということであり，Q_3 については，それ以上の値をもつデータが 3/4 あるということである。(牧野)

度数表

統計調査によって得られた結果は，集団の要素に共通の性質(たとえば，性別・年齢・所得金額)によって分類され，集計されて統計表にまとめられる。統計表には，性別のように，質的性質によって分類されたものと，所得金額のように，量的性質によって階級に分類されたものとがある。前者を構造統計表といい，後者を度数分布表という。度数分布表から，累積度数表をつくることができる。これらをまとめて，度数表という。

★解説

統計データの解釈を容易にするために，**度数表**がつくられる。よく用いられる度数表は，**度数分布表**と**累積度数表**である。しかし，累積度数表をつくるには，度数分布表を必要とするので，ふつうそれらを1つの度数表にまとめて利用する。このような度数表から，**ヒストグラム**(柱状図表)や**累積図表**が書ける。度数表で，階級の間隔を**級間隔**，または**階級の幅**という。ヒストグラムで，階級の幅を1にとって，これに高さをかけると，階級の度数に等しくなる。

ヒストグラムや累積図表を，**相対度数**や**累積相対度数**を用いてグラフに示すこともできる。たとえば，階級の数が k で，それぞれの階級にはいる度数を f_i $(i=1,\ 2,\ \cdots,\ k)$ とするとき，これを全度数 N で割り，かつ階級の間隔(これらを**級間隔**という)を1にとれば，相対度数 f_i/N が得られる。このとき，

$$\frac{f_1}{N}+\frac{f_2}{N}+\cdots+\frac{f_k}{N}=1$$

である。

なおグラフは，視覚に訴えて，データの性質の理解をはかるためにつくられるので，読解が容易なものでなければならない。

関連ページ

度数分布表　74

所得金額 (百万円)	人数 (人)	累積人数 (人)
以上　未満		
10～ 12	675	675
12～ 14	406	1,081
14～ 16	283	1,364
16～ 18	217	1,581
18～ 20	177	1,758
20～ 30	438	2,196
30～ 40	247	2,443
40～ 50	117	2,560
50～ 60	56	2,616
60～ 70	36	2,652
70～ 80	21	2,673
80～ 90	10	2,683
90～100	11	2,694
100～150	14	2,708
150～200	3	2,711

表1　度数表

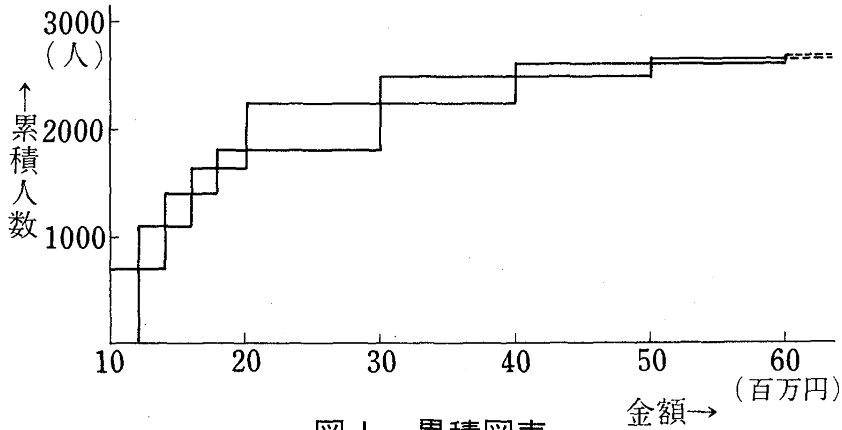


図1　累積図表

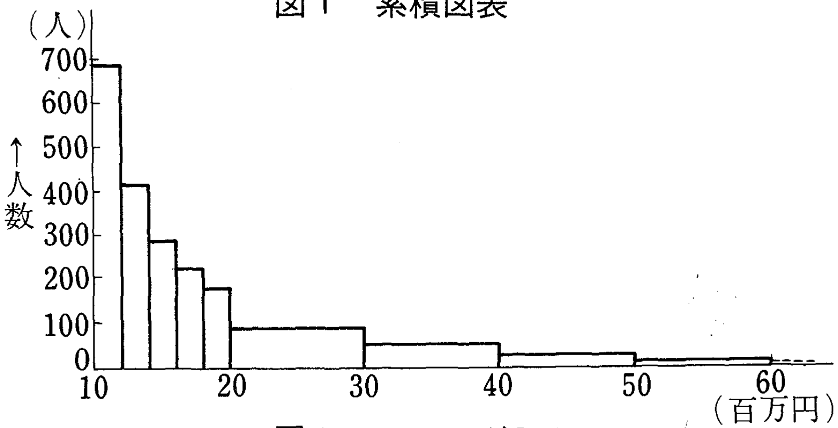


図2　ヒストグラム

■ヒストグラムと累積図表　表1は，ある税務署で公示した昭和56年の年間所得金額1000万円以上の申告所得データからつくった度数表である。表では，階級の数が15になっている。これをもとにしてヒストグラムをつくるには，次のような注意を払う必要がある。それは，上の表で，1000万円から2000万円までは階級の幅が200万円になっているが，200万円から1億円までは1000万円きざみになっている，などということについてである。これをヒストグラムに書くには，次のようにする。いま，1000万円から2000万円までの各階級の度数を，そのまま縦軸の高さとするならば，2000万円から1億円までは，実際の度数の1/5倍を高さにとる。これは，度数折れ線の下の面積が，なるべく不変にたもたれるようにするための措置である。しかし，そのようにしてヒストグラムをつくっても，折れ線で結ぶとき，不等間隔の処理で，また困ってしまう。図2には，紙面の都合で，横軸6000万円までしか目盛ってないが，図で，点線で書いた部分が，不等間隔の処理についての1つの方法である。ただし，そのように書くと，折れ線の下の面積が不変でなくなってしまうので困る。これに対し，累積度数表から累積図表を書くには，そのような困難はおこらない(図1)。その意味では，度数分布表よりも累積度数表のほうが都合がよい。

(牧野)

二項分布

コインをn回投げて表のでる回数を数える，という実験を考えると，①1つの試行(この場合“コインを一回投げる”)の結果が他の試行の結果に影響を与えない，②注目する事象E(この場合“表がでる”)の生ずる確率が試行ごとに変化することなく一定している。このような場合，同じ試行をn回繰り返したとき，事象Eの生起する回数の分布は二項分布(binomial distribution)となる。

★解説

上のような試行の列を**ベルヌイ試行**(Bernoulli trial)という。①は試行の**独立性**をいい，②は事象Eが生ずるか否かという結果の**2値性**と，確率が試行回数によらず一定という**定常性**を述べている。各試行においてEの生ずる確率をp,生じない確率を $q=1-p$ とする。このとき，n回の試行中，事象Eの生起する回数をXとおくと，

$$P\{X=k\}={}_nC_kp^kq^{n-k}=\frac{n!}{k!(n-k)!}p^kq^{n-k}$$

$$(k=0,\ 1,\ 2,\ \cdots,\ n)$$

平均　$E(X)=np$　　　分散　$\mathrm{Var}(X)=npq$

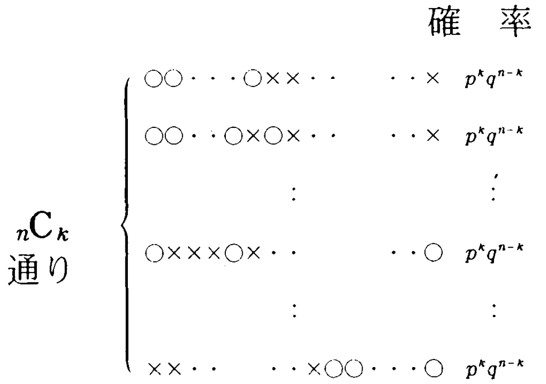
となる。これをパラメータ$(n,\ p)$の二項分布とよぶ。

これは次のように求まる。図1のどれかの系列が生ずれば，$\{X=k\}$となる。いずれの系列の生ずる確率も，①②の性質より p^kq^{n-k} となり等しい。このような系列の総数は，n個の中からk個を選んで○印をつけるときの組み合わせの数 ${}_nC_k=n!/k!(n-k)!$ だけある。したがって，上のように求まる。組み合わせ数 ${}_nC_k$ を二項係数とよび，$\binom{n}{k}$ とも書き表す。これは，$(p+q)^n$ を二項展開したときの p^kq^{n-k} の項の係数となっているからである。

このベルヌイ試行において，Eがr回生ずるまでの試行回数の分布は，**負の二項分布**($r=1$ のとき**幾何分布**)となる。

関連ページ

負の二項分布　404
幾何分布　404



(いずれの系列も○印 k 個，×印 $n-k$ 個ある。○印は事象 E の生起を表す。)

図 1　二項分布の求め方

■不良品の個数の分布　ある製品を生産するとき，不良率が 2 %であるという。その製品 50 個詰の箱に不良品が 2 個以上混じっている確率を求めてみよう。不良品数 X の分布は $n=50,\ p=0.02$ の二項分布と考えてよい。したがって，2 個以上不良品の含まれる確率は，

$$\mathrm{P}\{X\geqq 2\}=\sum_{k=2}^{50}{}_nC_k(0.02)^k(0.98)^{n-k}$$

となる。これをこのまま計算するより余事象の確率を利用する方が簡単で，

$$\mathrm{P}\{X\geqq 2\}=1-[\mathrm{P}\{X=0\}+\mathrm{P}\{X=1\}]$$

$$=1-\frac{50!}{0!50!}(0.02)^0(0.98)^{50}$$

$$-\frac{50!}{1!49!}(0.02)^1(0.98)^{49}$$

$$=1-0.364-0.372=0.264$$

と電卓で求まる($0!=1$ と約束する)。

この箱詰をたくさん購入するとき受け入れ検査をする。いま 1 箱を調べたとき，その中に 4 個以上不良品があったとする。$p=0.02$ と仮定すると，その確率は 0.017 となる。不良率 2 %の

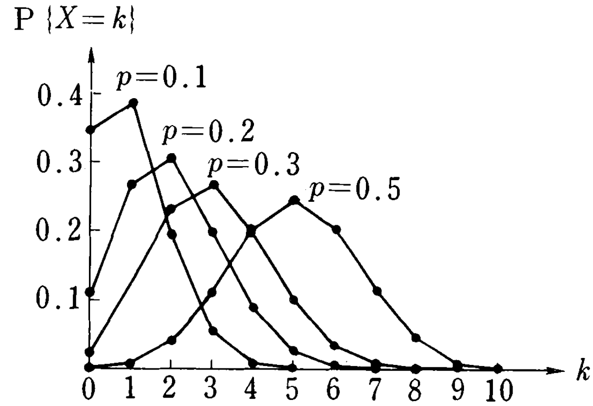


図 2　二項分布（$n=10$ のとき）

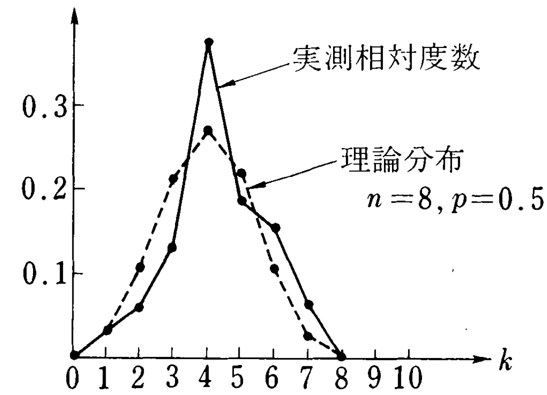


図 3　ねずみの毛状の遺伝

製品の箱詰としてはありそうもないことが起きたわけで，本当は不良率がもっと大きい規格外れのものではなかろうかと考えるのが自然だろう。この考えを発展させて，抜取検査(→ 108 ページ)の方式を設計できる。

■ねずみの毛状の遺伝　ねずみの毛状を決める遺伝子があり，W は直毛，w は巻き毛をつくる。W が優性である。雌は ww 型，雄は Ww 型のつがいで，一腹 8 匹を産んだ 32 組について，直毛(Ww 型) の子の数の分布を観測した(R. A. Fisher ほか)。メンデルの法則によれば，Ww 型も ww 型も等確率で生まれるので，理論的には，$n=8,\ p=0.5$ の二項分布に従うはずである(図 3)。

(森)

2重クロス表の独立性と属性相関係数

2重クロス集計では，χ^2 検定により2項目間の独立性が検定できる。また，項目が2値カテゴリーの場合やカテゴリー間に順序が考えられる場合には，連続変数の相関係数に対応する種々の属性相関係数が計算できる。これを用いれば，項目間の関連性が分かる。

★解説

■独立性　表1に示すように項目XとYのカテゴリーの組ごとの頻度表を $m\times n$ 分割表という。頻度 f_{ij} の代わりに，確率(比率) $p_{ij}=f_{ij}/f_{..}$ が用いられることもある。$m=n=2$ の場合を四分表，それ以外を多重分割表ともいう。

$X \backslash Y$	$Y_1\ Y_2 \cdots Y_n$	計
X_1	$f_{11}\ f_{12} \cdots f_{1n}$	$f_{1\cdot}$
X_2	$f_{21}\ f_{22} \cdots f_{2n}$	$f_{2\cdot}$
$\vdots$	$\vdots\ \vdots\ \vdots$	$\vdots$
X_m	$f_{m1}\ f_{m2} \cdots f_{mn}$	$f_{m\cdot}$
計	$f_{\cdot 1}\ f_{\cdot 2} \cdots f_{\cdot n}$	$f_{..}$

表1　$m\times n$ 分割表

項目XとYが独立とは，$\overline{f}_{ij}=f_{i\cdot}f_{\cdot j}/f_{..}$ として，$f_{ij}=\overline{f}_{ij}$ $(i=1, \cdots, m\quad j=1, \cdots, n)$ のときをいう。独立性の検定には，四分表の関連係数 δ を拡張した次の χ^2 統計量が用いられる。

$$\chi^2=\sum_{i=1}^{m}\sum_{j=1}^{n}(f_{ij}-\overline{f}_{ij})^2/f_{ij}$$

■属性相関係数　四分表で2項目が順序尺度，すなわち $X_1<X_2$，$Y_1<Y_2$ とする。このとき，$f_{11}>\overline{f}_{11}$ または $f_{22}>\overline{f}_{22}$ のときを正の関連(相関)，$f_{11}<\overline{f}_{11}$ または $f_{22}<\overline{f}_{22}$ のときを負の関連という。関連は連続変数の相関係数に対応する概念である。$f_{12}=f_{21}=0$ の場合を完全関連といい，$f_{12}=0$ または $f_{21}=0$ の場合を最大関連という。多重分割表でも，項目が順序尺度であれば関連を考えることができる。右ページに関連性の尺度を表す種々の属性相関係数を示す。

関連ページ

χ^2 検定　314
クロス集計　90
決定行列　104
ROC 曲線　14

■**例題**　表2に示す分割表のデータを用いて属性相関係数の紹介を行う。Y_1 と Y_2, Y_3 と Y_4 を1つのカテゴリーとみなした場合，四分表になる。

$X \backslash Y$	Y_1	Y_2	Y_3	Y_4	計
X_1	4	4	2	3	13
X_2	1	1	3	2	7
計	5	5	5	5	20

表2　2×4分割表

■**四分表の場合**　次の δ は独立性からの乖離（かいり）度を表す。

$$\delta = f_{11} - \overline{f}_{11} = \frac{1}{f_{..}}(f_{11} \cdot f_{22} - f_{12} \cdot f_{21})$$

$$= (8 \cdot 5 - 2 \cdot 5)/20 = 1.5$$

これを基に，次の**ユールの関連係数 Q** が得られる。

$$Q = \frac{f_{11}f_{22} - f_{12}f_{21}}{f_{11}f_{22} + f_{12}f_{21}} = \frac{f_{..}\delta}{f_{11}f_{22} + f_{12}f_{21}}$$

$$= \frac{20 \times 1.5}{8 \times 5 + 2 \times 5} = \frac{30}{50} = 0.6$$

X と Y が独立の場合には $Q=0$，正負の最大関連と完全関連の場合は，$Q = \pm 1$ になる。

四分点相関係数 r は，ピアソンの相関係数を四分表へ拡張したものである。

$$r = \frac{f_{..}\delta}{\sqrt{f_{1.}f_{2.}f_{.1}f_{.2}}} = \frac{20 \times 1.5}{\sqrt{10 \times 10 \times 13 \times 7}} = 0.314$$

完全関連で $r = \pm 1$ になる。

ロジスティック相関係数 V は，$(-\infty, \infty)$ の範囲の比率を表す。

$$V = \log\{f_{11}f_{22}/f_{12}f_{21}\} = \log 4$$

■**多重分割表の場合**　多重分割表の属性相関係数は，χ^2 統計量を修正したものである。

$$\chi^2 = \frac{2 \times (4-3.25)^2 + (2-3.25)^2 + (3-3.25)^2}{3.25} + \frac{2 \times (1-1.75)^2 + (3-1.75)^2 + (2-1.75)^2}{1.75}$$

$$= 2.418$$

χ^2 の値は $f_{..}$ とともに大きくなるので，次の ϕ^2 を考える。

$$\phi^2 = \chi^2 / f_{..} = 0.121$$

この平方根0.348を **ϕ 係数**とよぶ。カテゴリー数を $n > m$ として，完全関連で $\phi^2 = m-1$ になる。

ピアソンのコンティンジェンシー係数 C は，ϕ 係数からカテゴリー数 m の影響を省いたものである。

$$C = \sqrt{\chi^2/(f_{..} + \chi^2)} = \sqrt{\phi^2/(1+\phi^2)}$$

$$= 0.328$$

ただし，$\max C = \sqrt{(m-1)/m}$ で $m = \infty$ のとき1になるが，m の小さな値では1にならないので，$m \leq 5$ では使わないほうがよいとされている。

チェプロウのコンティンジェンシー係数 T^2 は，ϕ^2 と C の欠点を若干修正しているが，$m \neq n$ の場合には $|T^2| < 1$ である。

$$T^2 = \phi^2 / \sqrt{(m-1)(n-1)} = 0.070$$

クラマーのコンティンジェンシー係数 Cr は，$0 \leq Cr \leq 1$ になる。

$$Cr = \phi^2/(m-1) = 0.121 \quad (n > m)$$

（新村）

日程計画

指定期日，手持ち資源などの制約のもとで，計画達成に必要な作業の日程を決めることが日程計画である。古くから使われているガント図が役立つ場もあるが，建設計画などにはPERTやCPMが使われる。

このほか，ジョブショップ・スケジューリング(または順序づけ)やラインバランシングなども，日程計画のために研究されている分野であるが，厳密には整数計画法に帰着される難問であるため，なかなか実用にされない。

★解説

■ガント図　横線工程表ともよばれる。例示は右ページで行うが，作業進度を横線棒グラフで示すものである。簡単に作れる点は便利であるが，工程表を作る担当者の勘に頼るだけであるから，ごく簡単なものにしか向いていない。

■PERT　292ページでその内容について説明するが，ポラリス・ミサイル開発計画に当たって考案された方法で，その手法の使いやすさと結果の簡明さのために，プロジェクト型の日程計画，たとえば大規模な工事において広く利用されるようになった。線形計画法やシミュレーションとともに，ORの3大手法といわれたこともある。

■GERT　PERTでは，仕事の順序は確定的であったが，不確定な要素のはいったときに用いられうるよう，計算機シミュレーションの一手法として考え出された。

■ジョブショップ・スケジューリング　多数の受注品があり，その製作にはいくつかの機械にかける必要がある。受注品ごとに機械にかける順序が定まっている，おのおのの機械にかける品物の順序をくふうして，全品目の加工をできるだけ早く終わらせる問題である。**順序づけ**ともいう。

■フローショップ・スケジューリング　上と似た問題で，全品目とも，加工順序が同一の場合をいう。右ページ参照。

関連ページ

PERT　292

線形計画法　218

シミュレーション　166

GERT　50

■**ガント図の例**　印刷工場で2台の機械に仕事をかけるが，仕事によってはⅠしかかからない(図1では太線で示している)。仕事には段取り時間30分が必要であるとしよう。図1では段取り時間を点線，仕事時間を実線で示している。仕事が始められる時刻も考慮して図1のようなガント図が作られる。

機械の能力に差があって，機械Ⅱは機械Ⅰの1.5倍速いものとする。この場合には，図2のように，能力に応じて時間の目盛りを変えておくと便利である。このようなくふうを時に応じ，場に応じてするとよい。

■**2段階工程の順序づけ**　2種類の機械A，Bをこの順に用いてn個の品物を加工する。おのおのの加工時間がわかっているとき，最も早く全体の仕事を終えるためには，どの順に仕事をかければよいか。これは**ジョンソン・ルール**とよばる方法で簡単にできる。

仕事	A	B
a	10	12
b	8	15
c	13	9
d	11	11

たとえば，右の表のような仕事が与えられていたとする。表中の8つの数のうちの最小は8である。これはAの列中にあるので，bの仕事を最初に行う。bを除いて，次の最小値を探すと9で，これはBの列中にある。そこで，cの仕事を最後に行う。以下この方法を繰り返す。すなわち，残りの2つの仕事a，dのうちの最小値はaの10でAの列にあるから，aを先に行う。結局，図3に示すような仕事の順序で加工するのがよいことがわかる。

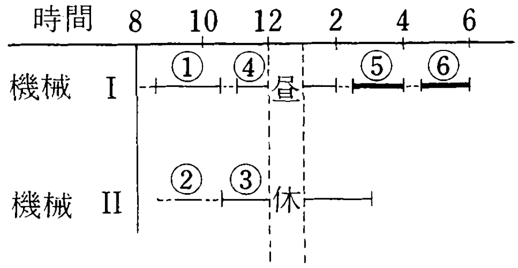


図1　ガント図(横線工程表)

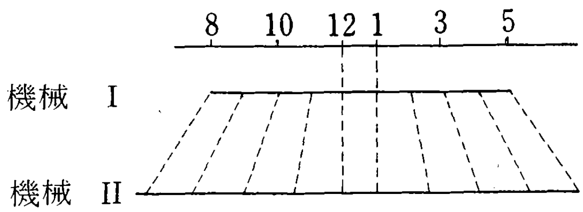


図2　ガント図の拡張

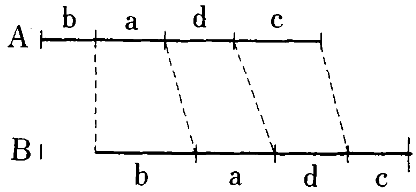


図3　ジョンソン・ルール

すべての仕事が終了する時刻は55となる。たとえば，a，b，c，dの順で仕事にかけると57が終了時刻になって図3の順序より悪い。この方法は特別な3段階の場合にまで拡張される。　　　　(森村)

日本の統計

一国の統計機構は，その国の行政全般の考え方や統計発展の歴史によっていろいろな型がある。わが国の統計機構は，統計事務を各省に分散させた分散型をとっており，これらの統計を相互に調整するための機関として総務庁統計局に統計基準部が設けられている。

地方公共団体の場合は，一部集中型に近い統計機構となっている。

統計に関する法令には，①統計機関の組織と権限を定めた統計組織法(総務庁設置法など)，②統計の真実性の確保，統計調査の重複の除去，統計体系の整備及び統計制度の改善発達を定めた統計手続法(統計法など)，③個別調査の実施に関する統計実体法(各省の統計調査規制など)があって，統計の発展と真実性・効率性の保持に寄与している。

★解説

18世紀までの統計は，国家の政治上注目すべきことがらを叙述的に記述するにとどまっていた。19世紀に入ってから，国家の情報収集活動が活発になるにつれて統計資料も増加し，その結果をより簡潔明快に示す方法として数字による表現が一般化して，近代的統計が誕生した。

わが国では，1871年7月，大蔵省に統計司を設けたのが官庁において統計の名称が使われるようになった最初で，1881年太政官の中に統計院(総務庁統計局の前身)が置かれるに及んで統計の用語が確定するとともに，官庁統計の基盤が確立した。現在では，国の主な統計は9省庁において実施されるようになっており，それらの結果は，経済・社会の実態と動向の判断資料として使われている。さらに，経済・社会のメカニズムを数式的，数量的にとらえる計量経済モデルを通して，経済政策，社会政策の効果判断(たとえば所得税引き上げの経済効果，大学・高校授業料補助の社会各層への影響把握)ができるようになり，的確な立法，行政を行うための判断資料として統計の役割は増大している。

関連ページ

国民経済計算 120
景気動向統計 92
生産統計 204
貿易統計 354
労働統計 414
家計統計 60
企業経営統計 72

(1) 国の主な統計機関と調査名はつぎのとおりである。

統計機関	実施している主な統計調査
総務庁統計局	国勢調査，事業所統計調査，住宅統計調査，社会生活基本統計調査などの基本的・定期的なセンサス。 労働力調査，家計調査などの経常的な毎月調査。
経済企画庁	景気動向調査と国民経済計算推計。
大蔵省	貿易統計，税務統計，企業経営統計など。
文部省	学校基本統計調査などの教育統計。
厚生省	厚生行政基礎調査などの厚生統計。
農林水産省	農業センサスなどの農林水産統計。
通商産業省	工業統計調査，商業統計の2つのセンサス。 生産，商業に関する各種の動態調査や鉱工業生産指数などの鉱工業関連指標。
労働省	毎月勤労統計調査，賃金構造基本統計調査などの労働統計。
建設省	建築着工統計，建設工事統計などの建設関連統計。

(2) 国が実施する調査は，それぞれ速報・報告書によって公表されているほか，主な結果はつぎの刊行物に収録されている。

刊行物などの名称	主管部	公表周期·時期
日本統計年鑑	総務庁統計局	毎年4月
日本経済指標	経済企画庁調査局	毎月
国民所得統計年報	経済企画庁経済研究所国民所得部	毎年2月
国民経済計算	〃	年4回
財政金融統計月報	大蔵省大臣官房	毎月
文部統計要覧	文部省大臣官房	毎年6月
厚生統計要覧	厚生大臣官房統計情報部	毎年10月
ポケット農林水産統計	農林水産省経済局統計情報部	毎年3月
通商産業統計要覧	通商産業大臣官房調査統計部	毎年11月
鉱工業指数年報	〃	毎年7月
運輸経済統計要覧	運輸省官房統計情報管理部	毎年3月
労働統計要覧	労働大臣官房政策調査部	毎年3月
建設統計要覧	建設省計画局	毎年10月ごろ
鉄道要覧	日本国有鉄道情報システム部	毎年3月
物価指数月報	日本銀行調査統計局	毎月
〃　年報	〃	毎年5月

(市野)

ネットワーク・モデル

いくつかの都市を結ぶ道路網のように，ノード(節(ふし))の集合と，そのうちのいくつかの対を結ぶアークの集合を一緒に考えたものをネットワーク(網)という。アークには向きがついている場合も多い。ネットワークの上を流れる物があるとき，最小費用，最短経路，最大流量などの流し方を求める問題は一般にネットワーク最適化問題とよばれる。案外な問題がこの問題に帰着されたり，その特殊性を生かした算法が得られているなどのことのため，重要視されている。

★解説

■輸送問題　いくつかの倉庫からいくつかの販売店に品物を輸送するのに，各倉庫の供給量を上回らない範囲で各販売店の需要をすべて満たし，しかも総費用を最小にする輸送計画を求める問題で，特殊な算法が考えられている。これは，倉庫と販売店をノードとし，各倉庫から各販売店のすべてへ有向アークが付与されているネットワークになっている(図1)。倉庫でも販売店でもないノードが介在すれば，積み替えのある輸送問題になる(図2)。

■最短路問題　各アークに長さが与えられているネットワークにおいて，あるノードから，他のどこかのノードに至る最短路を求める問題である。この問題は設備更新の問題にも利用される(右ページ参照)。

■最長路問題　上で，最も長い路を求めることは，PERTやCPMにおける，クリティカル・パスを求めることと同じになる。PERTは大規模な工事の日程計画などに広く利用されている手法である。

■最大流問題　各アークに容量が与えられているとする。どのアークも容量以下の物を流して，かつ全体として最大の流量を流すような流し方を求める問題である。道路網の基本設計などに利用できよう。

関連ページ
輸送問題　392
PERT　292
CPM　142
最大流問題　130

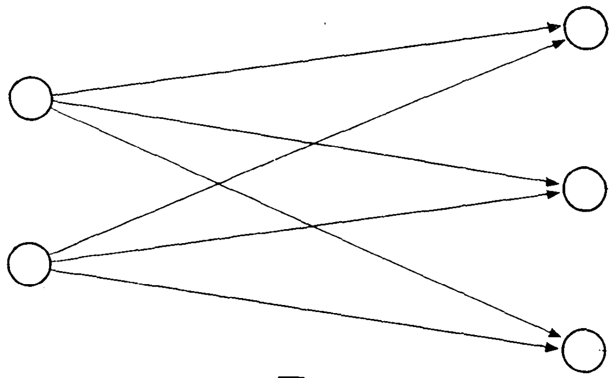


図 1

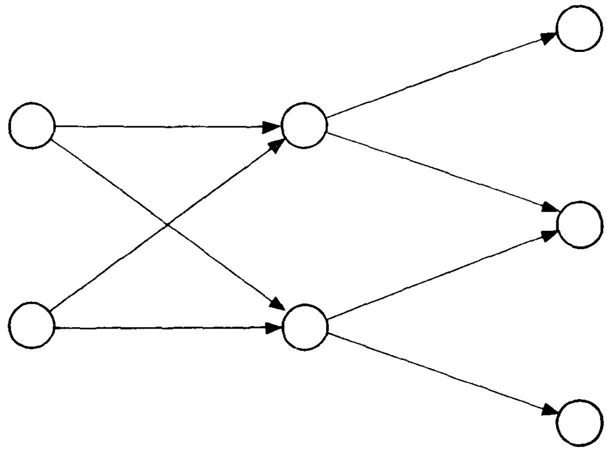


図 2

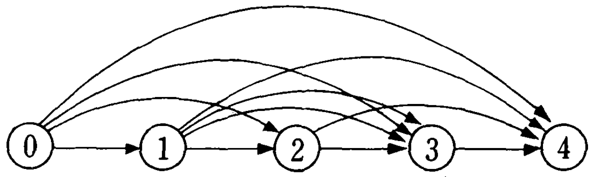


図 3

■設備更新の問題　自動車を何年目に更新するのがよいか，という問題を考えてみる。ノード0で表す時点で1台の自動車を購入する。ノード1，2，……はその後1年後，2年後，…の時点を表している。第i年目に購入した自動車をj年目まで使用するとしたとき，購入価格と売却価格との差，維持費，及びその車に乗るときの気分の評価も必要なら含めて，総費用C_{ij}を算定しておく。これをそれぞれのアークに付与すると図3のようなネットワークができる。このネットワークで，ノード0からノード4に至る最短路を見つければ，それが，4年間の経費を最小にする更新政策を与える。たとえば，0→2→4という路が最小費用を与えるのであれば，2年後に自動車を買い替えるのがよいということになる。なお，この問題は動的計画法(→270ページ)を利用して定式化することもできる。

■ネットワーク問題の解法　輸送問題には簡単で有効な計算法が見出されており，「輸送問題」の項で述べている(→392ページ)。図2のような「積み替えのある輸送問題」は，各積み替えノードに仮想の予備在庫があって，送り出すにはそれを使い，受け取ったときはそれに加える，という考え方を取るとき，標準的な輸送問題に帰着できる。

輸送問題の特別な場合とも見られる割当問題は，上で予備在庫1としたときの積み替えのあるモデルと同じになるし，この特殊な場合が最短路問題であって，割当問題よりずっと容易に解が求まる。

「割当問題」の古典的な解法の1つについては別項で述べる(→420ページ)。また，「最大流問題」についても，有用な解法がいくつか工夫されている。その1つについては別項に記す。

（森村）

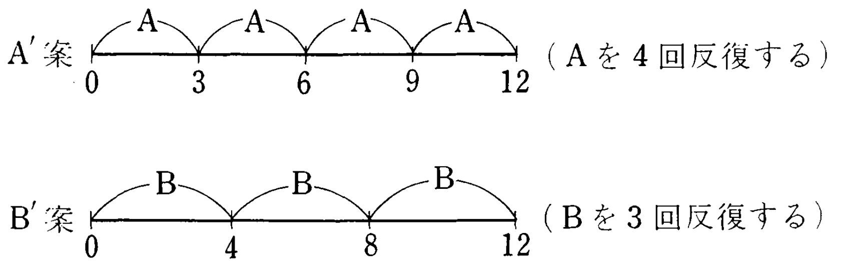
年価法の役立ち

正味年価，すなわち年当たりの平均利益の指標(→268ページ)で方策の経済性を判断する方法である。年価法は，寿命の違う複数の代替案(相互に排反的な案)を比較するときに使える便利な方法である。

★解説

たとえば，家屋を建設する場合，木造にするか鉄筋コンクリートにするかという代替案があるとき，各案は投資額や毎年にかかる費用が異なるだけでなく寿命も違ってくる。工場での機械を専用機にするか汎用機にするかという問題や，製品設計での材料や設計案の選択の問題でも寿命の違いを正当に考慮する必要が生じる。

このように，寿命の違う案の経済性を判定するには，案の間の寿命の不揃いを調整しなければならない。そのやり方は，同一設備を反復して使うこと(**like-for-like replacement**)を仮定する方法である。たとえば，A案の寿命が3年でB案の寿命が4年のときは，A案，B案を下図のように反復して使うA′案，B′案におきかえる。この結果，A′案とB′案は寿命が同じになり，直接に比較できることになる。



A′案とB′案は現価，終価，年価のいずれを用いても比較できる。その際，A′案，B′案のような反復案の年価は，A案，B案の単一案の年価と一致するので，単一の年価で比較してよいことがいえる。

関連ページ

投資案の経済性指標　268

年価　144

■反復案の年価は単一案の年価に一致する　下表に示すような寿命の違う2つの設備を比較してみよう。$i=10\%$。

案	初期投資	年間操業費	寿命
A	1000万円	600万円	4年
B	2000万円	600万円	6年

ここで設備Aを3回反復し，12年間の案A′の年価 $M_{A'}$ を求めてみる。

$$M_{A'}=\Big[\{1000+600[M\to P]_4^{10\%}\}+\{1000+600[M\to P]_4^{10\%}\}\times[S\to P]_4^{10\%}+\{1000+600[M\to P]_4^{10\%}\}\times[S\to P]_8^{10\%}\Big]\times[P\to M]_{12}^{10\%}$$
$$=(2902+2902\times0.683+2902\times0.467)\times0.1467$$
$$\fallingdotseq 915\ (\text{万円})$$

つぎに，設備A単一でみたときの年価 M_A を計算する。

$$M_A=1000[P\to M]_4^{10\%}+600$$
$$\fallingdotseq 915\ (\text{万円})$$

このように，反復案の年価は単一案の年価に一致することがわかる。

■寿命が違う案は単一案の年価で比較する　上の数値例のように，寿命が違う案の比較では，反復案A′(3回反復)と反復案B′(2回反復)の比較をせざるをえないが，A′の年価とAの年価，B′の年価とBの年価はそれぞれ一致するので，単一案Aと単一案Bについて年価で比較してよいことになる。

$$M_A=1000[P\to M]_4^{10\%}+600$$
$$\fallingdotseq 915\ (\text{万円})$$
$$M_B=2000[P\to M]_4^{10\%}+300$$
$$\fallingdotseq 759\ (\text{万円})$$

この結果，Bの設備が安くつくことになる。

ちなみにA，B案を現価で比較してみよう。

$$P_A=1000+600[M\to P]_4^{10\%}$$
$$\fallingdotseq 2902\ (\text{万円})$$
$$P_B=2000+300[M\to P]_6^{10\%}$$
$$\fallingdotseq 3307\ (\text{万円})$$

現価ではAの設備が安くなるが，この計算法は正しい方法ではない。

■寿命を長くする効果　いま100万円の設備を用いたときの寿命が3年と推定された。もし寿命が6年と倍増したとき，設備の値段は，いくらまでなら高くなってもペイすることになるかを調べてみよう。$i=12\%$ とする。

寿命が6年の設備の値段を X として，両者の年当たりのコストが一致する X の値を求めてみる。

$$100[P\to M]_3^{12\%}=X[P\to M]_6^{12\%}$$
$$X=100[P\to M]_3^{12\%}/[P\to M]_6^{12\%}$$
$$=100\times1.71=171\ (\text{万円})$$

寿命が2倍になったとき，設備の値段は約1.7倍までは引き合うことがわかる。　　（中村）

ノンパラメトリック検定

母集団分布についての情報がほとんどなくて，標本数もそれほど大きくない場合，特定の母集団分布に依存しない検定法が望まれる。ノンパラメトリック検定は，このようなときに用いられる検定法であって，母数によらない検定ともよばれている。その方法は，きわめて多種にわたり，たとえばランダム性の検定，独立性の検定，適合度検定，母集団分布についての検定などにおいて，さまざまな方法が開発されている。

★解説

2つの母集団の平均値の間に有意差があるかどうかを検定するのに，よく t 分布が用いられる。しかし，そのためには，母集団分布が正規分布であることが前提条件となり，さらに母分散が等しいことが要請される。このような仮定を設けることが適当でないとき，たとえば右ページに示すウイルコクソン検定が用いられる。これは，2つの母集団において，その分布の型は同じであるが，位置に違いがあるかもしれない，ということがわかっているとき，位置の差の検定をしようとするものである。この検定に用いられる統計量は，2つの母集団からの標本をこみにして大きさの順にならべて得られる，一方の母集団の標本の**順位和**である。ノンパラメトリック検定法では，このように，観測値そのものでなく，大きさの順に並べたときの情報を利用するものが多い。ところで，ウイルコクソン検定は，2つの母集団の位置についての検定であるが，これとは別に，2つの母集団の分布関数の間に差があるかどうかを検定する方法の1つとして，コルモゴロフ・スミルノフ検定がある。また，よく用いられるノンパラメトリック検定として，**順位相関**を計算したり，**一致の度合い**を調べたりする方法もある。

関連ページ
平均値の差の検定 324
t 分布 314
正規分布 200
統計量 314

■ウイルコクソン検定　これは，2つの母集団分布について，その分布関数を，一方は $F(x)$，他方を $F(x-\theta)$ とおけるとき，

帰無仮説　$\theta=0$

対立仮説　$\theta \neq 0$

を検定するノンパラメトリック検定である。具体的には，次の手順で検定を行う。いま，母集団分布 $F(x)$ からの標本を

$$x_1, \ x_2, \ \cdots, \ x_m$$

とし，$F(x-\theta)$ からの標本を

$$y_1, \ y_2, \ \cdots, \ y_n$$

とする。このとき，まず $N=m+n$ (個)の標本をこみにして，小さい方から順に，1からNまでの順位をつける。つぎに，$F(x)$ からのデータ x につけられている順位の和 W を求める。m，n がともに10以上のとき，W は，近似的に，

$$平均 \quad a=\frac{1}{2}n(n+m+1)$$

$$標準偏差 \quad b=\sqrt{\frac{nm(n+m+1)}{12}}$$

の正規分布に従うことがわかっているので，これを用いて上の仮説を検定する。対立仮説を，上のように

$\theta \neq 0$

とすれば，両側検定を行うことになるが，$\theta>0$ や $\theta<0$ とおけば，片側検定となることはいうまでもない。

■ウイルコクソン検定の例　2つの食餌療法が，体重減少の効果において差があるかどうかを調べたい。年齢，健康状態，実験を始める前の体重などがよく似ている40人を抽出し，無作為に20人ずつの2つのグループに分けた。ある期間の実験後，体重減少を示す40個のデータを，小さい順に並べて，始めのグループの順位和 W_0 を計算したところ，

$$W_0=329$$

であった。

この場合，$m=n=20$ であるから，順位和 W が近似的に

$$平均 \quad a=\frac{1}{2}\times 20\times 41=410$$

$$標準偏差 \quad b=\sqrt{\frac{20\times 20\times 41}{12}}$$

$$=36.97$$

の正規分布にしたがうと考えてよい。したがって，

$$Z=\frac{W-410}{36.97}$$

が標準正規分布に近い分布をする。いまの場合，Z の値が，

$$Z_0=\frac{329-410}{36.97}=-2.19$$

になる。この値は標準正規分布の両側95％点である -1.96 より小さいので，有意水準5％で帰無仮説を棄却する。

■ケンドールの順位相関係数　表1は，ある高校で卒業就職希望者に対して行った調査からとってきた，2人の生徒による希望職業群の比較である。ただし表の職業群について，たとえば事務というのは，「専門的，技術的，管理的，事務的職業」をさし，販売は「販売的職業，サービスの職業」のことをいう，などとなっている。2人の生徒は，いずれも建設を第1順位にあげているが，2人の生徒による順位についての相関はいくらぐらいか，を知るために，よく**ケンドールの順位相関係数** τ を算出する。これは，n 個の項目から任意の2個を取り出したとき，それに対する2人の判定者間の評価の順位が同じならば $+1$，逆順ならば -1 という評点を与え，それらの総和 S を用いて，

$$\tau=\frac{S}{\frac{1}{2}n(n-1)}$$

により算出される。

τ を手軽に計算するには，負の評点の和，つまり逆順の個数 Q を数え，

$$\tau=1-\frac{2Q}{\frac{1}{2}n(n-1)}$$

によるのがよい。

たとえば，表1については $n=7$ であるが，逆順の数がいくらになるかを，次のようにして数える。

はじめに，Aのつけた順位を番号順に並べ，それに対応するBの順位を書き込んだ表2をつくる。表2で，逆順数が0，3，1，……とあるが，これは1より高い順位のものが右方にないので0個，5より高い順位のものが右方に3個，3より高い順位のものが1個，……という具合に数えたものである。これらの数を合計して，逆順の総和が $Q=6$ になることがわかる。したがって，ケンドールの順位相関係数は，

$$\tau=1-\frac{2\times 6}{\frac{1}{2}\times 7\times 6}=0.429$$

になる。

職業群 ＼ 生徒	事務	販売	運輸	金属	機械	繊維	建設
生徒A	7	3	2	6	4	5	1
生徒B	4	3	5	7	2	6	1

表1　職業選択の順位表

Aによる順位	1	2	3	4	5	6	7
Bによる順位	1	5	3	2	6	7	4
逆順の数	0	3	1	0	1	1	0

表2　逆順の数

■スピアマンの順位相関係数　2人の判定者間の一致の度合いを知るのに，

$$R=1-\frac{6\Sigma d^2}{n^3-n}$$

もよく用いられる。ただしn, dは,

n; 対象となった項目の数

d; 順位の差

である。上で算出されるRのことを,スピアマンの順位相関係数という。これは,2人の順位づけが全く一致していれば1,完全に正反対ならば-1になるようにつくられている。また,nがある程度大きければ,統計量

$$T=R\sqrt{\frac{n-2}{1-R^2}}$$

が,近似的に自由度$(n-2)$のt分布にしたがうことを用いて,仮説検定をすることもできる。

表2について,Rを計算するには,次のようにする。まず,順位差dが

0, −3, 0, 2, −1, −1, 3

になるので,

$$\Sigma d^2=0^2+(-3)^2+\cdots\cdots+3^2=24$$

したがって,

$$R=1-\frac{6\times24}{7^3-7}=0.571$$

となる。

■一致性の係数 W 左ページの表1は,2人の生徒A,Bによる職業選択の順位であったが,さらに生徒Cの選択順位も合わせて,表3のような,生徒全体の順位の一致性をみるには,一致性の係数Wが役立つ。これは,次の手順で計算される。一般に,n個の項目を,m人が判定して,順位をつけるものとする。このとき,

職業群 / 生徒	事務	販売	運輸	金属	機械	繊維	建設
生徒A	7	3	2	6	4	5	1
生徒B	4	3	5	7	2	6	1
生徒C	7	6	1	4	3	5	2
順位合計	18	12	8	17	9	16	4

表3　3人の生徒による順位

$$S=\Sigma\left\{(\text{順位合計})-\frac{m(n+1)}{2}\right\}^2$$

を計算して,

$$W=\frac{12S}{m^2(n^3-n)}$$

を求めればよい。

評価が完全にランダムになされていれば$W=0$,一致していれば$W=1$になる。また,統計量Wが近似的に自由度

$$\nu_1=(n-1)-\frac{2}{m},$$

$$\nu_2=(m-1)\left\{(n-1)-\frac{2}{m}\right\}$$

のF分布にしたがうことを用いて,仮説を検定することもできる。

上の表3について,Wを計算してみよう。$m=3$, $n=7$で,

$$S=(18-12)^2+\cdots\cdots+(4-12)^2=16$$

よって,

$$W=\frac{12\times166}{3^2\times(7^3-7)}=0.659$$

(牧野)

PERT

Program Evaluation and Review Technique の頭文字を並べたことばであるが，ふつうパートとよばれて親しまれている。プロジェクトの日程計画を立てるときなどによく用いられる方法で，仕事の順序関係に主眼を置き，クリティカル・パスとよばれる一連の仕事を重点的に管理することによって，納期内の作業完了を目指す。CPM も似た方法で，今は PERT（パート）に含めて取り扱われている。また，PERT には，PERT/COSTのようないくつかの拡張型もある。

★解説

■矢線図　PERT による仕事の第一歩は矢線図(アローダイアグラム)を作ることである。複雑な開発計画や建設計画となると，この矢線図を作るだけでも大変で，電算機を利用することになるが，ここでは，ごく簡単な例を用いて矢線図を作る原理を説明する(右ページ参照)。

■仕事所要時間の見積もり　おのおのの仕事の所要時間が一意的に見積もられていると話は簡単になる。しかし，天候に左右される仕事などには特に多いことであるが，所要時間が偶然変動を持っているとみなさざるをえないこともある。また，ある種の仕事については，過去の経験がないため，見積もりに自信の持てないこともあろう。このような場合には，最大所要時間，最小所要時間，及び最も自信の持てる(すなわち最も起こりそうな)所要時間の 3 つを見積もる。これらをそれぞれ，**悲観値**，**楽観値**，**最尤値**とよぶ。これら 3 つの値から所要時間のしたがう確率分布をベータ分布によって推定して利用することも多い。

■クリティカル・パス　矢線図をネットワークと見ると，その上の最長路は，全く遊び時間のない一連の仕事の集まりを表す。これをクリティカル・パスとよび，これを求めることが PERT の中心的な作業になる。

関連ページ
CPM　142

クリティカル・パス　88
ネットワーク・モデル　284

部屋の改修工事を例にとる。仕事は次の表に示されるように分割される。

仕事	内　容	先行作業
A	家具類運び出し	なし
B	床の補修	A
C	窓枠取り替え	A
D	サッシ取り付け	C
E	ガラス入れ	D
F	壁紙貼り替え	B
G	家具類据え付け	E，F

従来の木枠の窓は傷みが激しいので，アルミ・サッシに取り替え，一部の床が壊れているので補修し，ついでに壁紙も明るいものに貼り替えようというのが全体の作業の内容である。

このとき，床の補修や窓枠の取り替えは並行して作業してもよいが，窓枠の取り替えがすまないうちにサッシを取り付けることはできない，というように，おのおのの仕事の間には，おのずから定まる順序関係がある。仕事Dの前に仕事Cを終えていなければならないとき，仕事Cは仕事Dに**先行する**という。表の「先行作業」の欄に，それぞれの仕事に先行する仕事が書き込んである。

さて，各仕事は2つのノードとその間を矢じるしで結んで表す(図1(a))。左側のノードは，その仕事を始める前の状態を表し，右側のノードはそれが終わった状態を表す。ノードには適当に番号をつけ，それを書き込む(図2)。

仕事Cが仕事Dに先行するという関係は，図1(b)のようにしてごく自然に表現することができる。仕事Eには仕事Dが先行するから，この右方に仕事Eを示す矢じるし(アロー)をつなげればよい。

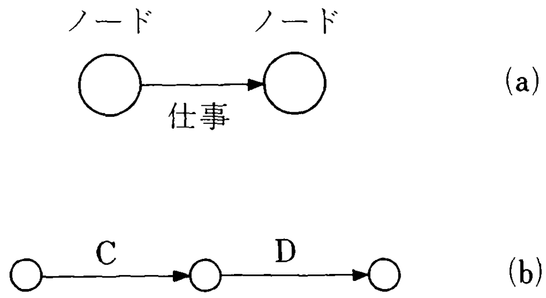


図1

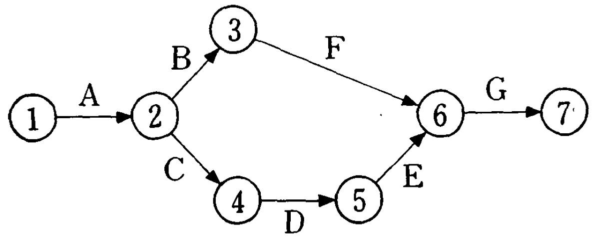


図2

仕事Bと仕事Cは，仕事Aの終了したあとならば並行的に進めることができる。この関係を示すため，図2で，ノード2から，BとCを示す2つの矢じるしが分かれている。

仕事Eと仕事Fが終わると仕事Gが始められるので，ノード6は仕事EもFも終わっているという状態を示している。こうして，例にあげた改修工事の仕事は，図2のような矢線図に書き上げることができた。

各仕事の所要時間をここに書き込んで，クリティカル・パスを見つける次の段階の作業に移る。その解説は89ページに記す。　(森村)

発想的 OR (heuristic OR)

数式を用いないで，考え方や論理の道筋をはっきりさせる OR 的作業の総称。発想法や論理展開の明示法及び質的要因の分析法などから成り，発見的 OR ともいう。OR の問題発見から問題設定に至る手順とも考えられ，定式化に先立って十分活用されることが望まれる。ときには，数式化を経ない問題解決の方法となることもある。

★解説

■集団発想法　どこに問題があり，どこに焦点を当てて解析すべきか，といった点について，チームの全員から着想を引き出すことが大切である。このため，**ブレーン・ストーミング**や**デルファイ法**などが利用される(右ページ参照)。

■問題の構造化手法　問題を構造化して全体象を把握するための手法で，個人でも集団でも利用される。2つの方向，すなわち機能展開型と事実収束型とが考えられている。前者の代表的方法として**ワークデザイン**や**系統図法**，後者のそれとして **KJ 法**がある。前者は機能や目的を「何のために？」と考えながら次々と展開していって，まずは理想的なシステムを描き出そうという方向であり，後者は一見無関係とも見える個々の事実を積み重ねる中から全体の構想を探り出そうという方向である。

■決定の視覚化　将来どのように展開するか不分明な過程において，時折下す決定の影響がどのようになるかを視覚的に示すために，**決定の木**や **PDPC** が用いられる。

■要因の分析手法　メリット・デメリット分析，±インパクト分析，マトリックス分析，あるいは円形や星形のグラフのパターンによる分析やチェックリストによる分析など，時と場合に応じていろいろ利用されている。

■総合化　量的データを基に総合判断をするため，**フェース法**などが利用される。

関連ページ
ブレーン・ストーミング　295
デルファイ法　295
ワークデザイン　416
KJ 法　416
決定の木　106
PDPC　106
メリット・デメリット分析　386
マトリックス分析　386
フェース法　322

左ページにあげた手法のうち，ここでは集団発想の項に掲げた，ブレーン・ストーミングとデルファイ法について解説する。その他の手法については，それぞれ別項目とし解説するので，そちらを参照していただきたい。

■ブレーン・ストーミング（brain storming） 数人，多くとも10人内外の参加者を集めた会議形式によって行われる。参加者は着席順に主題についての着想を単語かごく短かい文で述べる。この段階で他の者は一切批評はしない。司会者はそれを黒板や紙に書き並べ，一巡したら更に繰り返す。新しい発想のない者は「パス」と言って，次の順の者に発言を任す。何巡かして，新しい発言がなくなればやめる。

その後，黒板に書かれた語句を適当にグループ分けする。この段階でKJ法が使われたりする。整理が終わったら，グループごとに優劣や可能性の検討を行うが，その際必要があれば発言者のより詳しい説明を求めたり，要因の分析法などを利用するのもよい。

要は，たとえどのように細かなものでも，すべての発想を拾い上げることで，OR活動を正しい方向に向けさせることが目的であるから，参加者をできるだけリラックスさせ，自由な雰囲気を作って発言をしやすくすることが大切である。このためには司会者の役割が重要で，発言を引き出す間は，もっぱら書記役と進行役とに徹する心掛けが必要である。少し慣れると，この方法は気軽に行え，その効果も大きいので広く実施されている。

■デルファイ法 意見を聞くべき人が多くなってしまうと，それらを一堂に集めることも，たとえ集めてもブレーンストーミングを行うことは極めて困難になる。代わりにアンケートなどによる意見の集約が考えられる。この方法では，はじめはアンケートや自由討議によって得られた質問表を回答者に配布して，項目ごとに賛否や意見を聞く。

集められた回答を集計し，否定的意見の集中した項目は除いて，意見分布が見やすいように整理して質問表を作り直し，再度回答を求める。回答者は他人の意見を知った上で回答をし直す。

このようにフィードバックを繰り返しながら何回か回答を求めると，意見の収束が比較的得られやすいという。

もともと，デルファイ法は，超長期的な将来の技術の実現を予測するための手法として1967年にランド研究所で開発された。その後その有効性が認められて，技術予測のみならず，社会現象などの予測にも広く用いられるようになった。ここでは，予測手法というより意見集約の手段として取り上げた。

デルファイ法の名は，デルファイ神殿で，アポロがゼウスの予言を聞いたというギリシャ神話にちなんでつけられたとのことである。

参考文献 近藤次郎[22]など。（森村）

発注点方式

発注間隔をとくに規定せず，在庫量が発注点とよぶ水準にまで減少したとき，一定の数量を発注する在庫管理方式(JIS－OR用語)。発注点は，調達期間内の需要量に見合うものとして定められる。発注点法による在庫管理は，主として常備材に対して適用され，需要変動の幅が比較的小さいものに適している方式である。

★解説

■対象の選定　発注点方式による在庫管理は，日々の需要が数量的に大きいが金額的には比較的小さいような，いわゆる常備材に適すると考えられている。たとえば，ボルトなどのように，多くの部署で使用し，その数も多いものは他の部品とは無関係に在庫管理をしてもよい。需要変動は毎日同一の確率分布にしたがうとみなせる，という立場でそのモデルは作られる。この仮定が多少怪しいときでも，金額的にあまり高くなければこの方式を用いることも多い。そのため，ABC分析で対象を選ぶことも行われる。

■発注点　在庫モデルの項で説明したように，調達期間内の需要量しか在庫していなくなった時点で発注すればよいが，調達期間も需要量も確率的に変動するから，

発注点＝調達期間内の平均需要量＋安全余裕

として発注点は定められる。

■安全余裕　もし，ある特定の期間の需要変動について特別の知識があればそれを利用すべきであるが，通常のモデルではそのような状況を考えていない。そこで，ふつう，

$$\text{安全余裕} = \underset{(\text{通常 } 1.65)}{\text{安全係数}} \times \sqrt{\text{最大調達期間}} \times \left(\begin{array}{l}\text{単位期間内需要}\\\text{量の標準偏差}\end{array}\right)$$

として計算する。ここで，**最大調達期間**とは，目標の調達期間に許容される遅れを見込んだもので，納入側との交渉や過去の実績により定められる(具体例は右ページ参照)。

関連ページ

ABC分析　300
在庫モデル　128

■発注点の計算例　右の表1のような需要量の部品がある。最大調達期間を2週として，発注点を計算してみる。

週	需要量
1	223
2	205
3	187
4	251
5	214
6	196

表1

ここにはわずか6個のデータしかないので，これから平均や標準偏差を求めても，かなり粗い推定値しか得られない。平均は当然

$$\frac{1}{6}(223+205+187+251+214+196)$$

$$=212.7$$

という計算で求められるが，標準偏差は定義どおり分散(→332ページ)を求めるよりは，範囲(つまり最大値から最小値を引いたもの)に表2に示す係数をかけて求めるほうが,このように個数の少ないデータのときには向いている。この計算法によると標準偏差は，

$$(251-187)\times0.39=25.0$$

である。

それで，安全係数を1.65とすると，安全余裕は，

$$1.65\times\sqrt{2}\times25=58.3$$

と計算される。

調達期間は2週間なので，この間の平均需要量は $212.7\times2=425.4$ となり，左ページの式から，発注点は，

$$425.4+58.3=483.7$$

と計算される。

実際には，このような半端な量では管理しにくいかもしれない。たとえば，100個単位で箱詰めにされているような場合である。そのときは適当に数値を丸め，たとえば

データ	係数	データ	係数
2	.89	7	.37
3	.59	8	.35
4	.49	9	.34
5	.43	10	.32
6	.39	12	.31

表2

500を発注点と考えるのが現実的であろう。要は，発注という行動の指針を与えておくことである。

■発注量の計算　発注点は，発注という行動を起こさせる指針であったが，そのとき，いくら発注すべきかが分からないと現実の行動には結びつかない。

発注量は輸送形態，梱包形態なども考慮して定めるべきであろうが，基礎的な値としては　EOQ(→129ページ)をとればよい。いまの例で，1年を51週とみて，$212.7\times51=10850$ が年間需要量 r，1回の発注費 K が5,000円，保管費(変動分)15円/年とすると，EOQの公式 $x^*=\sqrt{2Kr/h}$ に入れて $x^*=\sqrt{2\times5000\times10850/15}=2689$ を得る。それで1回に27箱を発注すると，ほぼ3カ月に1回注文することになろう。ただし，発注量は多少狂っても大差はないと考えられている。

■費用の見積もり　上の例で K は発注に費される人件費(検収や経理事務も含むが，発注をしなくてもかかる場合は除く)や連絡費が主なもの，h は在庫のために要する利子が大きなものであろう。

（森村）

林の数量化理論

日本において数量化理論といえば，林知己夫博士(統計数理研究所)により体系づけられた質的データの解析手法を総称することが多く，外的基準がある場合とない場合に大別できる。前者の手法としては，数量化Ⅰ類とⅡ類がある。後者の手法としては，数量化Ⅲ類とⅣ類，さらには多次元尺度構成法(MDS)と共通点の多い各手法がある。

★解説

■記号 δ の説明　説明変数が質的データの場合，各アイテムカテゴリーに反応する次のダミー変数を考える。

$$\delta_i(jk)=\begin{cases} 1\cdots i\ \text{サンプルが}\ j\ \text{アイテム}\ k\ \text{カテゴリーの値をもつとき} \\ 0\cdots i\ \text{サンプルが}\ j\ \text{アイテム}\ k\ \text{カテゴリーの値をもたないとき} \end{cases}$$

各ダミー変数に x_{jk} という数量を与えて，その線形和をとったものを i サンプルのサンプルスコア $\hat{y}_i$ とする。

$$\hat{y}_i=\sum_{j=1}^{m}\sum_{k=1}^{\gamma(j)} x_{jk}\delta_i(jk) \qquad (i=1,\ \cdots,\ n)$$

ただし，$\gamma(j)$ は j アイテムのカテゴリー数。

m はアイテム数。n はサンプル数。

■外的基準が数量　i 番目のサンプルの外的基準が数量 y_i で与えられている時，誤差平方和 $\Sigma(y_i-\hat{y}_i)^2$ を最小にするような，カテゴリースコア x_{jk} を求める手法が数量化Ⅰ類である。また，y_i と $\hat{y}_i$ の相関係数最大基準でもよい。

■外的基準が分類　n 個のサンプルが m 個の群に分類されている場合，相関比 η^2 (＝群間の分散／全体の分散)を最大にするようにサンプルスコア x_{jk} を求める手法が数量化Ⅱ類である。

■外的基準がない場合　サンプルスコア y_i (＝$\hat{y}_i$)とカテゴリースコア x_{jk} を同時に推測することにより，似た反応パターンをもつサンプルとカテゴリーが近くに集まるようにしたパターン分類手法を数量化Ⅲ類という。

関連ページ
質的データの解析　162
デザイン行列　250
数量化Ⅰ類　188
数量化Ⅱ類　190
数量化Ⅲ類　194
数量化Ⅳ類　196
MDS　42

■**数量化の系譜**　林氏が数量化理論を展開するきっかけは，「質的なものに任意に0，1，2，3などの数値を与えるリッカート・スケールに疑問を持ったときに始まる」とされている。すなわち，解析者が勝手に質的データに数値を与えて間隔尺度のように扱い，既存の多変量解析手法を適用していることに対する反省を出発点としている。

これに対し，同氏はデータのもつ情報から，何らかの基準を設けて数量を推測することを考えた。

外的基準のある場合には，外的基準を最もよく表現するように説明変数の質的データに数量を与える。これを**カテゴリー・スコア**とよび，これから各データごとに求まる外的基準の推定値を**サンプル・スコア**という。

外的基準のない場合には，距離や近さというものの定義を総合的立場から解析者の目的に応じて規定し，似たものどうしをできるだけ近くに集めるように内部操作を行い数量を決める。

同氏のグループによる啓蒙普及活動は特筆に値する。このため日本においては，質的データの解析が日常茶飯事となった。このように，林の数量化は日本におけるデータ解析を豊かなものとしたが，現場技術者の間に多少の混乱が認められる。すなわち，一部では数量化といえば林の数量化を指しているきらいがあり，他の数量化に無関心である傾向がみられる。

■**ダミー変数と数量化**　歴史的にみれば，林の数量化の方がダミー変数の多変量解析への導入よりも早い。しかし，今日では数量化Ⅰ類とⅡ類はダミー変数を用いた重回帰分析と判別分析と考えたほうが，理解が容易な場合もある。すなわち，重回帰分析や判別分析の知識が援用できる。たとえば，説明変数の選択も，重回帰分析の逐次変数選択法を参考にすればよい。

数量化Ⅲ類はアルゴリズムが異なるが，ダミー変数を用いて主成分分析や因子分析を行っても似たような目的が達成できる。

データ解析は，各手法のアルゴリズムの差異による違いに注意し，似たような目的を達成する手法を組み合わせて比較評価することが望ましい。

■**外的基準のない場合の林の数量化**

外的基準のないデータで，個体iとjの間の関係を表す尺度e_{ij}が与えられているとする。e_{ij}が頻度の形で与えられた時，数量化Ⅲ類が用いられる。同様にして，e_{ij}が①漠然とした類似性(数量化Ⅳ類)，②類似性および非類似性(K-L 型数量化)，③大小関係(一対比較)，④ランク・オーダのある群分け(MDA-OR)，⑤ランクオーダのない群分け(MDA-UO)に分かれる。

この分野は，多次元尺度法や双対尺度法などという名のもと他にも多くの手法が提案されている。　（新村）

パレート図

累積度数グラフの応用の１つである。たとえば在庫問題において，使用金額の大きい順に品目を並べ，累積品目数百分率，累積金額百分率を計算し，それらを横座標，縦座標にして，方眼紙上に打点して曲線(折れ線)で結ぶ。このようにしてかかれた曲線のことを**パレート曲線**，または**パレート図**という。

★解説

パレート図の生みの親は，アメリカの経営コンサルタントで，**品質管理**の推進者でもあるジュラン博士であるが，博士は，**ローレンツ曲線**の考えを，次のように品質管理の分野に導入した。すなわち，部品の生産管理において，不良原因を不良項目の大きさの順に並べ，これを横軸にとる。縦軸には，不良品数や損失金額の**累積百分率**をとる。このようにして，パレート図と名づけられたグラフをかく。すると，不良品数や損失金額が，ごくわずかの不良項目によって占められている場合が多く，そのことがよくわかる。そこで，この項目を重点管理項目として対策を講ずれば，大きな効果が期待されよう。これが，ジュラン博士の提唱である。**在庫管理**の分野でよく用いられている **ABC 分析**は，パレート図の応用の１つである。昭和 42 年に制定された**日本工業規格**オペレーションズ・リサーチ用語(JISZ8121)の中で，ABC 分析とは，「在庫品目が非常に多いとき，それを使用金額の大きさの順に並べ，A，B，C の３種類に分類し，能率的に重点管理を行うやり方をいう」と意味づけされている。しかし，このような努力配分を考える必要性は，何も在庫問題に限ってのことではない。**格差**とか**バラツキ**といったことばで表現される現象を正しく認識し，分析し，対策を講ずるのに，パレート図や ABC 分析の考えが役立つ。

関連ページ
品質管理　318
ローレンツ曲線　327
在庫管理　128

(1) 売上高		(2) セールスマン人数	(3)累積人数百分率(%)	(4)累積金額百分率(%)
階級	金額(万円)			
1	以上 5000～6000 未満	5	5	26.8
2	4000～5000	5	10	48.8
3	3000～4000	7	17	72.7
4	2000～3000	3	20	80.0
5	1000～2000	4	24	85.9
6	900～1000	1	25	86.8
7	800～ 900	1	26	87.6
8	700～ 800	1	27	88.3
9	600～ 700	1	28	89.0
10	500～ 600	2	30	90.0
11	400～ 500	3	33	91.4
12	300～ 400	3	36	92.4
13	200～ 300	12	48	95.3
14	100～ 200	22	70	98.5
15	0～ 100	30	100	100.0

表1　累積百分率の表(セールスマンの売上高)

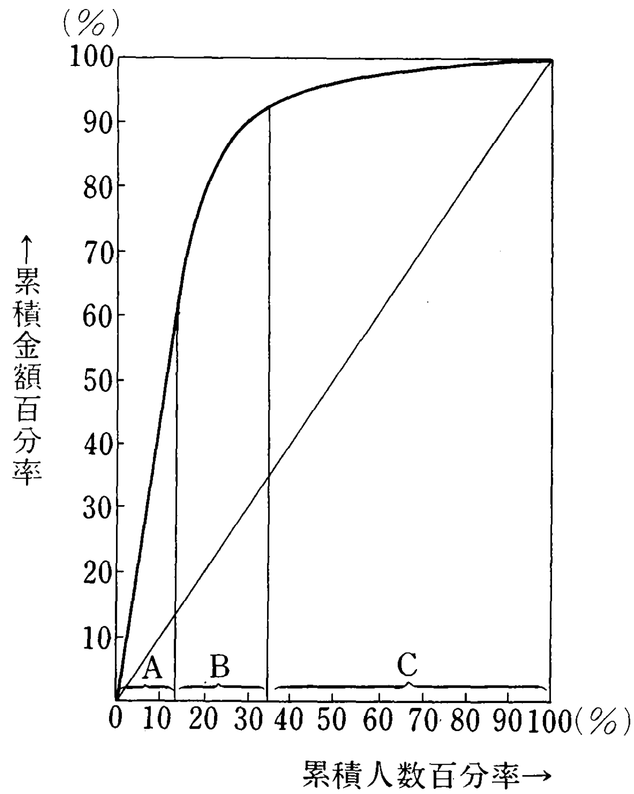


図1　パレート図(セールスマンの売上高)

■パレート図のかき方と活用の仕方

ある営業所に100人のセールスマンが所属していて，1期間での売上高が表1のようになったとする。これからパレート図をつくるには，次のようにする。

手順1　階級値と人数をかけ合わせたものを集計して，総額とする。すなわち，

$$総額=5500\times5+\cdots+50\times30$$
$$=102500\ (万円)$$

手順2　表1の第(3)の列の数値を，次の要領で書き込む。たとえば，階級1の行の第(3)の列は，階級1に属する人が全体の何％になるか，階級2の行には，階級2以上に属する人が全体の何％になるか，…を計算して記入する。

手順3　表の第(4)の列の数値を，次の要領で書き込む。たとえば，階級1の行の第(2)の列に48.8％と書いてあるが，これは階級2以上に属する人の売上高が，総額に対して48.8％にあたる，という具合に計算したものである。

これで，図1のようなパレート図が書ける。

上位10％(階級2以上)のセールスマンによる売上高は，総額の48.8％にも達していて，このクラスは，営業所にとってかけがえのない働き手である。人事管理の面に，このような考えを持ち込み，セールスマンをAクラス，Bクラス，Cクラスに分類して管理しようとするならば，それは一種のABC分析を行っていることになる。

参考文献〔52〕　(牧野)

判別分析

相異なるいくつかの群 G_1，G_2，…，G_k のいずれかに属することがわかっている個体がある。判別分析は，その個体がどの群に属するかを判別する最適な方法をみつけることである。個体について，p 項目の特性が計測されているとする。どの特性項目がそれぞれの群を識別するのに重要な役割を果たしているかという，要因分析に用いられることもある。

★解説

非倒産の企業を群 G_1，倒産の企業を群 G_2 としたとき，倒産1年前の有価証券報告書の中の財務諸表から，企業の倒産，非倒産を予知しえないかどうかを考えると，予測には2つの誤りが伴うのに気づく。一方は，群 G_1 に属する非倒産企業を倒産と判定してしまう誤りで，他方は，群 G_2 に属する倒産企業を非倒産と判定してしまう誤りである。求める判定の基準としては，たとえば，この2種類の誤りの確率の和を最小にするものが考えられよう。財務諸表の中の判別に用いる指標を x_1，x_2，…，x_p とし，変数は多変量正規分布に従い，群 G_1 と G_2 の分散共分散行列は等しいことがわかっているとする。このときの判別の基準は，

$$z = h_0 + h_1x_1 + h_2x_2 + \cdots + h_px_p$$

という線形式であり，$(x_1, x_2, \cdots, x_p)$ にデータを代入して計算した z の値が正か負かで，どちらの群に属するかを判定することになる。これを判別式とよぶが，判別式を決める係数 $(h_0, h_1, h_2, \cdots, h_p)$ は，群 G_1 に属することがわかっている N_1 個の標本と，群 G_2 に属することが分かっている N_2 個の標本から計算することになる。多変量正規分布に従うことと分散共分散行列がそれぞれの群で等しいことがわかっていれば，判別式は線形式になることから，線形判別分析とよぶこともある。2群で説明したが，考え方は多群でも同様である。

関連ページ
ベイズの定理を用いた判別　346
尤度比方式による判別　390
FUNCATによる判別　320
枝分かれ法による判別　34

分散共分散行列　334

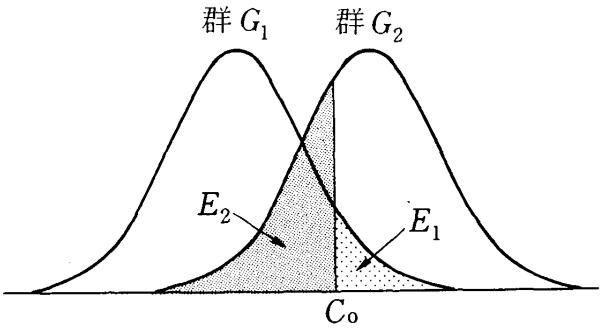


図1　最良でない判別点 C_0

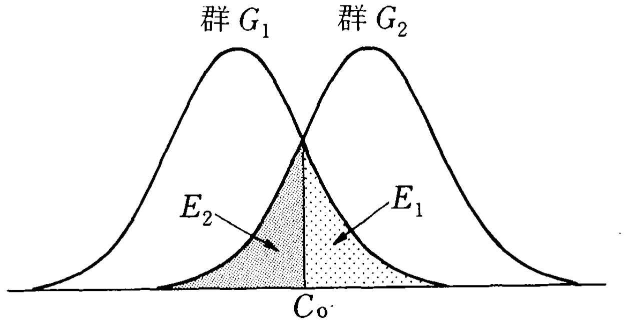


図2　最良の判別点 C_0

判別分析の考え方を把握するために1変数で説明する。非倒産企業群を群 G_1，倒産企業群を群 G_2 で表す。いま，有価証券報告書の中の経常利益率 x による判別を考える。x は正規分布に従い，群 G_1 での母分散と群 G_2 での母分散とは，等しいことがわかっているとする。

群 G_1 の平均を μ_1，群 G_2 の平均を μ_2 とすると，図1，2からわかるように誤判別の確率の和が最小であるという意味で最良の判別の基準点は，

$$C_0=\frac{\mu_1+\mu_2}{2}$$

である。このとき，群 G_1 に属する標本を群 G_2 からと判定してしまう誤りの確率 E_1 と，群 G_2 に属する標本を群 G_1 からと判定してしまう誤りの確率 E_2 の和

$$E_1+E_2$$

を最小にしていることになっている。

いま，非倒産企業群 G_1 の経常利益率の平均を4.0％，倒産企業群 G_2 の経常利益率の平均を -2.0 ％とすると，最良の判別点 C_0 は，

$$C_0=\frac{4.0+(-2.0)}{2}=1.0\%$$

となる　1％を判別点としたときが，誤判別の確率の和は最小になっている。

一般には経常利益率だけでなく，金利負担率，総資本回転率，売上債券回転期間，自己資本比率，流動比率，当座比率，固定比率…等々の多くの指標を用いて判別式を作ることになる。有価証券報告書の中の指標をすべて用いるのは賢明ではなく，現実にはどのような変数の組み合わせで判別式を作るのがよいのか，という変数選択の問題（→304ページ）が出てくる。

判別を行う人が蓄積してきた知識と経験とから，この指標は是非用いたいというものがあれば，それらの指標は統計的な変数選択の枠の中からはずして，採用しておくべきである。それらがすでに採り入れられたという条件のもとで，残りの指標の中から変数選択を行うことになる。　　　　（杉山）

判別分析の変数選択

たくさんある計測項目のすべてを用いて判別式を作ることが，必ずしも信頼性の高い判別式を得ることにはならない。判別力があまり変わらないならば，できる限り少ない変数の組で判別式を作った方が，結果として信頼性の高い判別式を得ることになる。その意味で，判別分析の変数選択は，実際上は非常に重要である。

★解説

よく用いられる変数増減法について，二群 G_1，G_2 で説明する。最初に選択する変数は，標準偏差で基準化した平均間の距離(**1 変数のマハラノビス距離**)が最も大きい変数 x_i を選択する。次に x_i にどの変数を追加したときに，2 変数による**判別効率**が最も大きくなるかを調べる。x_i に x_j を追加したときに判別効率が最大であれば，x_j を追加と判定する。x_i と x_j が選択されたという状況のもとで，残りの変数の中のどの変数を追加するのが判別効率を最大にするかを調べ，その変数 x_k を追加する。これを繰り返していくのだが，これが変数増減法の中の「増」の部分である。

変数 x_i と x_j に x_k が追加された状況の下で，3 変数で考えたとき x_i は判別に有用な変数か否か，x_j は判別に有用な変数か否かを調べる。有用でない変数があれば，それを除去する。有用でない変数が複数個あれば，有用性の最も低い変数を 1 つだけ除去する。これが「減」の部分である。

変数の増減は 1 つずつ行うのが規則である。その変数が有用か否かは，判別効率で表された ***F* 統計量**で行う。パッケージプログラムでは，変数を追加するか否かを決める基準値 F_{IN} と，変数を除去するか否かを決める基準値 F_{OUT} は，データ数が多いときはいずれも 2.0 に設定するのがよい。

関連ページ

マハラノビス距離　378

判別効率　379

F 統計量　325

百貨店とスーパーマーケットは，その成立過程において大きな相違があった。昭和40年代初めまでは，百貨店が小売業の売り上げの大部分を占めていたのに対し，50年代初めにはスーパーマーケットが百貨店につぐ状態になり，その数年後には売上高においてスーパーマーケットが百貨店を上回るような状況になった。両者は立地条件・販売理念においても異なる業種であったが，現在では大きな差がなくなりつつあるように思える。ところで両者は，経営状態においてどのような違いがあるか，その点を財務比率のデータから見ることにする。

百貨店は三越，東急百貨店，高島屋，大丸，松坂屋，……，名鉄百貨店の17店である。スーパーマーケットはダイエー，イトーヨーカ堂，ジャスコ，西友ストアー，ニチイ，ユニー，……，平和堂の20店である。取り上げた財務比率は46指標である。

変数選択の基準値 F_{IN}，F_{OUT} は，データ数が多いときは2.0が妥当であるが，ここではデータ数が少ないので，ともに3.0に基準値を設定して判別分析の変数選択を行った。選択された指標は次の6つである。

経営収支比率(17.6)
売上高債権対買入債務比率(67.7)
キャッシュフロー(8.9)の安定性
の3つの指標
売上高償却前利益率(14.1)
使用総資本利益率(13.4)の収益性
の2つの指標
公共分配率(14.1)

カッコ内の数字は最終段階での F 値であり，F 値が大きいということは，選ばれた6つの指標の中で，その指標の判別力が大きいことを意味している。

売上高債権対買入債務比率の F 値が67.7と大きく，これが百貨店とスーパーマーケットの判別に大きく効いている指標である。買い入れ債務においてはその規模に違いはあるが，両者の間に大きな違いがあるとは思えない。問題は売上債務の方であり，百貨店の売上債務はスーパーマーケットのそれよりも大きい。その理由は百貨店が特定のお得意さんを訪問し商売する外商によるのである。残りの指標のそれぞれを考察するのは省略するが，判別分析の変数選択により，財務比率から見た百貨店とスーパーマーケットの相違を知ることができる。これは判別分析の変数選択を，要因分析の手法として用いた例である。

誤判別率は以下のようである。

	百貨店と判定	スーパーと判定
百貨店のデータ	17	0
スーパーのデータ	0	20

（注） 日本経済新聞社の井出正宜氏が日経 NEEDS を用いて分析した結果を使わせていただいた。　(杉山)

汎用統計パッケージ

統計パッケージは，特定分野の研究者用に開発された専用統計パッケージと不特定ユーザーの広い要求をみたす汎用統計パッケージに分けられる。汎用統計パッケージは，幅広い統計解析機能のほか，レポートライティング，グラフィック出力，データ管理機能をもつものが多い。日本で利用できる代表的なものとして，BMDP, SPSS, SAS などがある。

★解説

■BMD，BMDP BMD (BioMeDical Computer Programs)は，1956年に開発が始まった医学に限定されない汎用統計パッケージで，カリフォルニア大学で開発された。その後，1967年にBMDP，1977年にBMDP-77が発表されている。BMDPは，BMDの統計処理手法が整理統合され少なくなっている。また，BMDにおいてパラメータが固定フォーマットであったものが，BMDPでは言語形式で指定できるようになった。

■SPSS SPSS (Statistical Package for the Social Sciences)は，1965年にスタンフォード大学で開発された。現在はSPSS Inc.が開発サポートを行っている。名前は社会科学向けであるが，実際は汎用と考えてよい。しかし，BMDPやSASに比べて処理できる範囲が狭い。

■SAS SAS (Statistical Analysis System)は，1966年にノースカロライナ州立大学でJ. H. Goodnightらを中心に開発された。現在はSAS Inc.で開発サポートされている。現存する統計パッケージとしては最も優れており，DSS (意思決定支援システム)や第4世代言語の特徴をもつ。統計解析のみならず，ORやシミュレーション，グラフィックなどを広くサポートしている。

関連ページ
SAS　136

統計パッケージのファイル管理　264

汎用プログラミング言語　308

■パッケージの分類　BMDは，図1の(1)のような個別プログラムの集まりである。この場合，ユーザーは各プログラムの仕様にあわせて入力データを準備しなければいけないし，プログラム間が独立しているので，1回限りの解析になる。BMDPは解析を連続して行えるように，解析結果をSAVEファイルに出力し，それを再度入力できるように工夫されている。さらに，BMDPプログラム・コントロール言語を用いて統一的に処理できる。

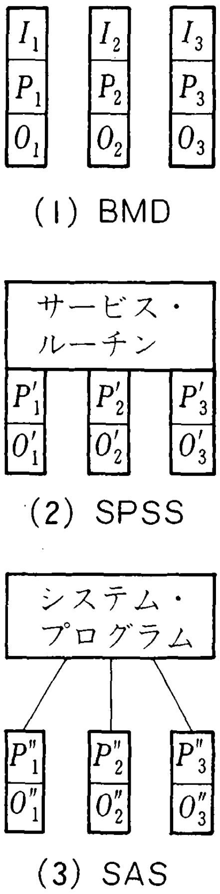


図 1

プログラムを機能面からみれば，入力部分(I)，処理部分(P)，出力部分(O)に分かれる。処理部分はアルゴリズム主体であり，サブルーチンとして汎用化しやすい。統計サブルーチンは各コンピュータ・メーカーとも提供している。サブルーチンの問題点は，ユーザーが入出力をコーディングしなければならないことと，サブルーチンごとに，パラメータ引き渡しなどの仕様が統一されていないことが多く，その都度マニュアルの引用が必要になる。

そこで，SPSSのようにSPSSコントロール・カードにより各プログラムに共通のサービスルーチンを利用できるものが表れた(2)。(1)に比べ，各プログラムに共通した部分をサービスプログラムとして共通化できたが，新規プログラムの追加が難しい。

そこで，(3)で表されるように，サービスルーチンの機能を強化し，オペレーティングシステムに似たスーパーバイザー機能をもつSASのようなパッケージが登場してきた。システムプログラム開発の困難を除けば，ユーザーにとっては各プログラムが互いに独立しており，新規プログラムの登録の容易なデータ解析システムとしてとらえることができる。

以上述べたことは，汎用統計パッケージ開発の歴史でもある。不特定のニーズをもったユーザーには，(1)，(2)，(3)の順に使いやすいといえる。しかし，この順に，マシン負荷も増大するため，特定の処理しか行わないユーザーにとっては，(1)またはサブルーチンの形式が望ましいこともある。

■ソフトウェア・パッケージの評価

評価尺度としてはいろいろあげられるが，米国のDatamation誌がユーザーに対して行っている調査結果が最も信用がある。その評価尺度としては，総合満足度，導入／初期利用，サービス，オペレーションの項目であり，統計部門ではSASが最優秀の評価を受けている。　(新村)

汎用プログラミング言語

汎用プログラミング言語は，シミュレーション言語や統計パッケージなどのコマンド言語のような特定問題向きの言語と異なり，汎用性のある不特定問題向けの言語である。コンピュータの命令は0と1の2進コードからなる機械語から作られている。これを人間に分かりやすい記号に置き換えたものをアセンブラ言語(assembler language)という。アセンブラ言語は機械語と一対一に対応しているため，人間の思考レベルからみて細かく煩雑である。そこで，機械語の数命令を人間が日常使っている表現に近い形の命令語に置き換えてプログラミングできる高水準言語(high-level language)が開発されている。

★解説

■高水準言語　アセンブラ言語では，よく用いられる定型処理を行うための一群の命令群に名前(マクロ命令)を与え登録できる。このマクロ命令をプログラム中で用いれば，対応する数命令をプログラムしたことになり，アセンブラ言語でのプログラムが容易になる。

高水準言語はこのマクロ命令に似た命令からできている。分かりやすい英単語や演算記号を用いてプログラミングできる。一命令が機械語の数命令に対応する。

プログラムを一度に機械語に置き換えることをコンパイル(編成翻訳)という。この方式をとる言語をコンパイル言語という。科学技術計算に用いられるFORTRAN, ALGOLや事務計算用のCOBOL，両方に用いられるPL／1などがある。

これに対し一命令ずつ置き換え実行していくものを解釈言語(interpretive language)という。科学技術計算用のBASICやAPLなどがある。プログラム開発には便利であるが，実行時間がコンパイル言語よりもかかる。

関連ページ
コンピュータ・ソフトウェア　124
汎用統計パッケージ　306
シミュレーション言語　168

■アセンブラ　磁気のオンとオフ，電気の正と負など，コンピュータに用いられる素子は０と１の２値に対応する。このような２値表示された単位をビット(binary digit)という。４ビットのメモリーは $2^4=16$ 通りを表す。すなわち16個の命令を表すことができる。機械語はこのような０と１の組み合わせで表される。これを英数字の符号に一対一に対応させて，人間に親しみやすくしたものがアセンブラ言語である。コンピュータにより，命令を表すために用いるビット数や命令語そのものが異なるため，異なった機種間の互換性はない。

高水準言語は，命令の種類や機能を標準化してある。このため，FORTRANで作製したプログラムは，少ない修正で異機種で実行できる。これは，機種ごとに異なったFORTRANコンパイラーが開発され，その機種独自の機械語に翻訳するからである。

ただし，各メーカーとも，ユーザーが他メーカーのコンピュータを使わないようにするため，そのメーカー独自の命令を拡張機能として提供していることが多い。もし，どの機種でも処理しようと(マシン移植性に富むという)考えている場合には，共通した命令だけでプログラミングすることが望ましい。

■FORTRAN (formula translation)

通常の数式表現に近い形でプログラムできるように1956年にIBMで開発された。今日利用されている高水準言語では最も古く，利用者の多い技術計算用の言語である。基本的な言語の習得は４～５日で行える。

■ BASIC (Beginner's All purpose Symbolic Instruction Code)　フォートランの入門用として1965年にアメリカのダートマス大学で開発された。最近ではインタープリタ型の言語としてパソコンやポケコンレベルの小さな計算機の言語として普及している。対話的にプログラムの作成が可能である。パソコンでBASICプログラムを開発し，大型機のフォートランに書き換えることも考えられる。BASIC言語はメーカーごとに文法が標準化されておらず，マシン移植性が悪い。

■ COBOL (COmmon Business Oriented Language)　フォートランと並ぶ代表的な言語であり，ファイルの取り扱いが容易なため，事務計算用に用いられる。フォートランよりは習得に時間がかかる。

■PL/1 (Programming Language 1)

フォートランとコボルの両機能を併せもつ汎用言語である。両言語に比べ機能が多いため習得が難しく，小型機での利用は難しい。また，すでに両言語で開発された莫大なプログラムの世界(財産)があり，PL/1はこれから徐々に普及すると思われる。　(新村)

非協力ゲーム

プレーヤー同士が互いに協力することなく，独立に自分のとるべき戦略を決定するようなゲーム。ゼロ和２人ゲームでは，一方のプレーヤーがその利得を大きくすることは，必然的に相手の利得を小さくすることを意味するので，非協力ゲームになっている。非ゼロ和２人ゲームや n 人ゲームでは，非協力ゲームも協力ゲームもどちらもありうる。

★解説

■双行列ゲーム　ゼロ和２人ゲームの利得は行列の形で与えられた。しかし，非ゼロ和２人ゲームでは一方の利得がそのまま他方の損失となるわけではないので，双方のプレーヤーのとった戦略の組ごとに，双方の利得を与えなければ利得を表現できない。つまり，行列の要素がそれぞれプレーヤー１，２の利得を表す２つの数の対として表される必要がある。このような形の行列，たとえば，

$$\begin{bmatrix} (0,\ 1) & (-2,\ 5) & (4,\ -1) \\ (-3,\ 3) & (3,\ 4) & (-2,\ 2) \\ (3,\ -1) & (2,\ -2) & (-1,\ 3) \end{bmatrix}$$

のようなものを**双行列**とよぶので，この形で利得が与えられるゲームを双行列ゲームとよぶ。

■双行列ゲームの均衡点　上例の双行列ゲームで，プレーヤー１が先手であるとする。彼が第１の手をとったとき，プレーヤー２は第２の手をとるのがよい。これを (1，2) で表す。同様に (2，2)，(3，3) の手がとられうるから，先手は第２の手をとって (2，2) という手で均衡し，利得３を確保するであろう。先手・後手の区別がないときは，混合戦略の範囲内で，ゼロ和ゲームの最適戦略の自然な拡張となっている**ナッシュ均衡点**が定義され，その求め方などが調べられている。その他，右ページの例のように，最適でない手で均衡することもある。

関連ページ
ゼロ和２人ゲーム　216
協力ゲーム　80

■**囚人のジレンマ**　次の双行列で利得が与えられるような先手・後手の区別のない2人ゲームで，プレーヤー2が，もし，第2の手をとったとすると，プレーヤー1にとっては第2の手をとった方が得で，それが分かればプレーヤー2も第1の手に代えない方が得ということになる。それで，両方のプレーヤーとも第2の手をとり，利得は双方とも2ということで落ち着いてしまう。もし双方とも1の手をとれば，双方にとって利得は倍になるにもかかわらず，一方が第1の手をとることを明らかにしていれば，他方は第2の手をとって自己の利得をより増やすことができるから，宣言をするわけにもいかず，さりとて黙って手1をとっても他方が第2の手をとることもありうるので，損を覚悟しなければならないであろう。こうして，このタイプの双行列ゲームでは，双方にとって最適でない手で均衡してしまう。

$$\begin{bmatrix} (4,\ 4) & (1,\ 6) \\ (5,\ 1) & (2,\ 2) \end{bmatrix}$$

パレート最適とは，一方の利得を下げることなしに他方の利得を上げることが不可能であるような均衡利得をいう。左ページの例の均衡点はパレート最適になっているが，上の囚人のジレンマゲームにおける(2, 2)はパレート最適ではない。

■**非協力均衡点**　プレーヤー1，2の混合戦略をp, qとし，qをq^*に固定してプレーヤー1の期待利得を最大にする混合戦略がp^*であり，逆にpをp^*に固定したとき，プレーヤー2の期待利得を最大にする混合戦略がq^*であるような組$(p^*,\ q^*)$を**非協力均衡点**，もしくは**ナッシュ均衡点**とよぶ。単に**均衡点**ということもある。双行列ゲームには少なくとも1つ均衡点が存在する。

ナッシュ均衡点を具体的に求めることは，一般的には簡単ではない。不動点とよばれるものを求めることに帰着されるので，不動点アルゴリズムが利用されるが，それは数理計画法の中で研究の進められている分野である。

均衡点は1つとは限らず，しかも均衡戦略は最適戦略とは限らない。次の例もそれを示している。

■**逢い引きのジレンマ**　次の双行列で利得が与えられるような，先手・後手の区別のないゲームでの均衡点は，純戦略として双方のプレーヤーがともに手1をとったとき(その場合の利得は(2, 1))，あるいはともに手2をとったとき(利得(1, 2))のほかに，混合戦略として，プレーヤー1が(3/5, 2/5)，プレーヤー2が(2/5, 3/5)をとる場合がある。この期待利得は(1/5, 1/2)で最適ではない。またこの均衡点はパレート最適ではないが，純戦略による均衡点はともにパレート最適である。参考文献〔30〕　(森村)

$$\begin{bmatrix} (\ \ 2,\ \ 1) & (-1,-1) \\ (-1,-1) & (\ \ 1,\ \ 2) \end{bmatrix}$$

非線形計画法

数理計画問題のうち，線形計画でないものを解く手法の総称であるが，整数変数をもつ場合の整数計画法はこれに含めないのがふつうである。対象とする問題の範囲は広く，したがって用いられる算法の種類も多い。一般に線形計画法に比べ，計算に必要な手間はずっと多く，実際に用いる際には適切な算法を選び，さらにそれを問題に合わせて手直ししていくことが必要である。単に**最適化法**とよぶこともある。

★解説

■問題の形　数理計画法であるから，

目的関数　$f(x_1, x_2, \cdots, x_n) \to$最小 [最大]

制約条件　$g_i(x_1, x_2, \cdots, x_n) \geq 0, \quad i=1, \cdots, m$

$h_j(x_1, x_2, \cdots, x_n) = 0, \quad j=1, \cdots, l$

という形の問題を扱う。

■非線形性　線形計画法であげた生産計画の例では，利益は生産される製品の量に比例していると仮定しているが，現実には多く作れば規模の経済性から単位利益は増加し，目的関数は直線でなく右上がりの曲線となる。また，使用される電力量も製品の生産量に比例するのではなく，多量に生産すれば効率が上がり，直線よりは凹関数となるであろう。このような非線形性を忠実にモデルにとり入れて最適解を求めようとすると，もはや線形計画法では扱えず，非線形計画法を用いなければならない。実際の数理計画応用問題ではもともと線形では表現できない場合も多く，これらすべてが非線形計画法の対象となる。

■線形計画法との選択　上のように現象を忠実に表現しようとすると，どうしても非線形モデルとなるが，最適解を得るための計算の手間は線形計画法に比べてはるかに増える。うまい構造になっていないと，3変数でも解くのは容易でない。線形計画法とのうまい使い分けが必要である。

関連ページ
数理計画法　186
線形計画法　218

■制約なし最適化問題　制約条件の g_i や h_j がない ($m=l=0$) 場合，f が微分可能ならば，(局所的)最適解の条件は1次と2次の偏微分係数で表すことができ，簡単な場合には解析的に値が求まることもある。

しかし一般には下のようなアルゴリズムを用いて数値的に反復計算をする。

■降下法　図1のように，ある点から出発して，目的関数 f の値が必ず減少するように次々と点を求めていき，最適点へ近づいていく方法を一般に降下法とよぶ。降下法では，次の点を求める方向を示す**方向ベクトル**と，その方向にどれだけ進むかを示す**ステップ幅**をどのように決めるかによっていくつかの種類がある。

方向ベクトルの決め方として最もポピュラーなのは，その点の1次偏微分係数を用いて f が減る率が最も大きい方向を選ぶもので，**最急降下法**とよばれる。また，2次までの偏微分係数を用いて最も効率的と思われる方向を選ぶものに**ニュートン法**がある。

ステップ幅は，方向ベクトルの方向で f の値が最も小さくなるように選ぶのが自然であるが，厳密に最小となる点を求めるには無限回の計算が必要となる。そこで，有限回の手続きで近似解が求められるよう，分割法，ニュートン法などを用いることが工夫されている。

■共役方向法　降下法ではその都度方向ベクトルを求めていたが，はじめに次元の数 n だけの方向ベクトルを決めておき，それを順繰りに繰り返し使っていくことが考えられる。この n 本の方向ベクトルは互いに直交していると都合がよいが，その条件を少しゆるめて互いに共役であるようにとるのがこの共役方向法である。

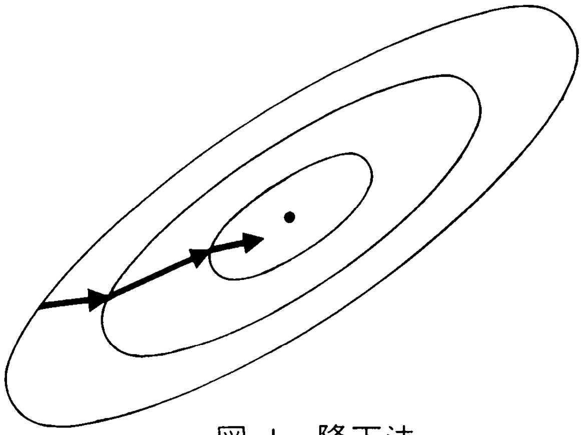


図1　降下法

■等式制約条件付き最適化問題　h_j のような等式制約条件のみで，g_i のような不等式制約条件がない ($m=0$, $l>0$) 場合には，**ラグランジュの未定係数法**を用いて制約なしの場合に帰着させることができる。

■制約条件付き最適化問題　一般に g_i のような不等式制約条件が付くと，問題はかなり難しくなる。アルゴリズムとしては，Kuhn-Tucker 条件を利用して制約なし最適化問題の形に導いて計算する変換法や，制約式を直接考慮に入れて制約式で定められる許容領域の中で最適値を求める射影法などがよく用いられている。

非線形計画法の詳しいアルゴリズムについては〔23〕参照。　(高橋)

標本分布

母集団から無作為に取り出された標本 $(X_1, X_2, \cdots, X_n)$ から作られる関数 $f(X_1, X_2, \cdots, X_n)$ を**統計量**といい，統計量の従う確率分布を**標本分布**とよぶ。その代表的なものとして，χ^2 分布，t 分布，F 分布などがある。

★解説

ここでは，統計学でよく使われる**正規母集団** $N(\mu, \sigma^2)$ からの標本($X_1, X_2, \cdots, X_n$)により作られる統計量

標本平均：$\overline{X}_n=(X_1+X_2+\cdots+X_n)/n$

標本分散：$S_n^2=[(X_1-\overline{X}_n)^2+\cdots+(X_n-\overline{X}_n)^2]/n$

などに関する標本分布を紹介する。

■ **χ^2 分布（カイ２乗分布）　自由度 f** の χ^2 分布は，ガンマ分布の特別な場合にあたり，その密度関数は，

$$f(x)=\frac{1}{2^{\frac{f}{2}}\Gamma(f/2)}x^{\frac{f}{2}-1}e^{-\frac{x}{2}} \qquad (x\geqq 0)$$

で与えられる。平均値は f，分散は $2f$ である。① $Z_1, Z_2, \cdots, Z_n$ が互いに独立で標準正規分布 $N(0, 1)$ に従うとき，$Z_1^2+Z_2^2+\cdots+Z_n^2$ の分布は自由度 n の χ^2 分布。② 上述の標本分散 nS_n^2/σ^2 の分布は自由度 $n-1$ の χ^2 分布。この性質を用いて，母分散 σ^2 の検定，区間推定を行う。

■ **t 分布　自由度** f の t 分布の密度関数は次のとおり。

$$f(x)=\frac{\Gamma(f+1/2)}{\sqrt{f\pi}\,\Gamma(f/2)}\left(1+\frac{x^2}{f}\right)^{\frac{f+1}{2}} \qquad (-\infty<x<\infty)$$

$f=1$ のときはコーシー分布ともいう。①X と Y が独立で，X が $N(0, 1)$，Y が自由度 f の χ^2 分布に従うとき，$X/\sqrt{Y/f}$ は自由度 f の t 分布に従う。②上述の $\overline{X}_n$，S_n^2 は独立で，$\overline{X}_n$ は正規分布 $N(\mu, \sigma^2/n)$，nS_n^2 は自由度 $n-1$ の χ^2 分布に従っているから，統計量 $(\overline{X}_n-\mu)/\sqrt{S_n^2/n-1}$ は自由度 $n-1$ の t 分布に従う。σ^2 未知のときの，母平均 μ の検定，区間推定に利用する。

関連ページ
正規母集団　200
ガンマ分布　66
正規分布　200
仮説検定　62
区間推定　82
コーシー分布　411

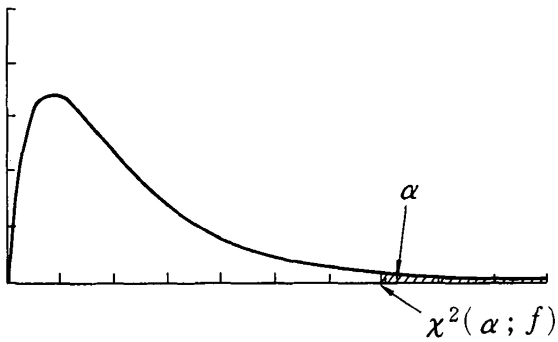


図 1　χ^2分布の密度関数と上側 $100\alpha\%$点$\chi^2(\alpha:f)$

■ ***F* 分布**　密度関数(ただし $x\geqq 0$)

$$f(x)=\frac{\Gamma\left(\frac{f_1+f_2}{2}\right)f_1^{\frac{f_1}{2}}f_2^{\frac{f_2}{2}}}{\Gamma\left(\frac{f_1}{2}\right)\Gamma\left(\frac{f_2}{2}\right)}\quad\frac{x^{\frac{f_1}{2}-1}}{(f_1x+f_2)^{\frac{f_1+f_2}{2}}}$$

をもつ分布を，**自由度** $(f_1,\ f_2)$ のF分布といい，$F(f_1,\ f_2)$ で表す。①XとYが独立で，それぞれ自由度f_1およびf_2のx^2分布に従うとき，f_2X/f_1Yの分布は $F(f_1,\ f_2)$。②X が $F(f_1,\ f_2)$ に従うとき，$1/X$ の分布は$F(f_2,\ f_1)$。③ $X_1,\ X_2,\ \cdots,\ X_m$；$Y_1,\ Y_2,\ \cdots,\ Y_n$ は互いに独立で各 X_i は正規分布 $N(\mu_x,\ \sigma_x^2)$ に従い，各 Y_j が正規分布 $N(\mu_y,\ \sigma_y^2)$ に従うものとする。$\{X_i\}$，$\{Y_j\}$ の標本分散をそれぞれ S_x^2，S_y^2 とすると，

$$\frac{mS_x^2}{(m-1)\sigma_x^2}\Big/\frac{nS_y^2}{(n-1)\sigma_y^2}$$

の分布は $F(m-1,\ n-1)$ となる。つまり，2つの標本分散の比がF分布に従い，これより標本分散の大小を比較できる。この考えは**分散分析法**(→336ページ)に利用される。

■**非心 *t* 分布**　XとYが独立で，Xは $N(0,\ 1)$，Yは自由度fのχ^2分布

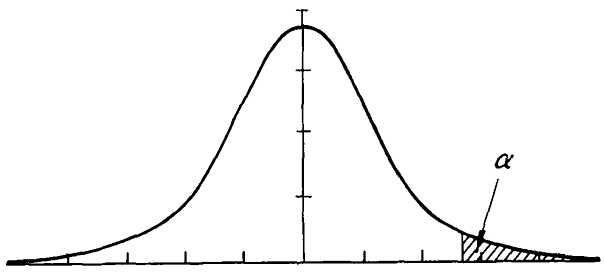


図 2　t分布の密度関数と上側 $100\alpha\%$点$t(\alpha:f)$

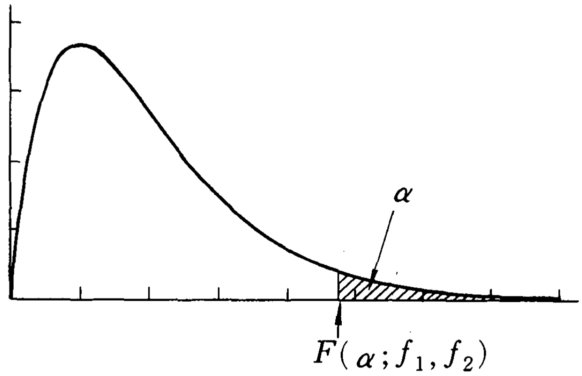


図 3　F分布の密度関数と上側 $100\alpha\%$点$t(\alpha:f)$

に従うとき，$(X+\delta)/\sqrt{Y/f}$ (δは定数)の従う分布を**自由度f，偏心率δの非心 *t* 分布**とよぶ。分散の等しい2つの正規母集団の**平均値の差の検定**(→324ページ)を行うとき，その**検定力**(→64ページ)を計算するのに用いる。同様に**非心 χ^2 分布，非心 *F* 分布**も与えられ，χ^2検定やF検定の検出力の計算に用いる。

■**順序統計量の分布**　ある母集団からのn個の標本 $X_1,\ X_2,\ \cdots,\ X_n$ を小さい方から順に並べた $(X_{[1]},\ X_{[2]},\ \cdots,\ X_{[n]})$ を**順序統計量**とよぶ。母集団の密度関数を $f(x)$ とすると，その同時密度関数(→228ページ)は，$x_1\leqq x_2\leqq\cdots\leqq x_n$に対して $n!f(x_1)\cdot f(x_2)\cdots\cdot f(x_n)$ となる。　(森)

比率

2つの集団の間の比例的関係を表す数値として比率が用いられる。これは，計算上は（分子／分母）の形で誘導される統計データであるが，この場合，分母としてとる集団を基礎集団といい，分子としてとる集団を比較集団という。

★解説

■構成比と対立比　ふつうに使われている比率は，**構成比**と**対立比**に大別される。構成比は，全体集団に対する部分集団の比率である。この場合，分子，分母の統計データについて，分子が分母の内訳区分に対応している。たとえば，人口総数のうち60歳以上の人口比率がそれである。このように構成比は，1つの統計集団の構造的特性を明らかにするために有用な比率である。構成比は，一般にはその値をパーセントで表す。パーセントは，全体集団の統計を100とした百分比であり，その計算式は(部分集団／全体集団)×100による。

これに対し，構成比以外の比率を対立比という。したがって，対立比はさまざまな比率をふくむ。分子，分母が同種集団のものである場合もあるし，異種の場合もある。分子，分母が同種集団のものであっても，そこには総数と内訳という関係がないので，その値が1をこえる可能性もある。たとえば，ある都市に住む人口数を分母にとり，その都市の事業所等に勤務する人口数を分子にとった比などがそれである。

対立比の利用の機会は，経済統計，社会統計とそれにも増して，経営統計の領域に多い。たとえば，資本や設備，労働力をそれぞれ生産高に対比させて，生産性の指標とする。また，計量経済学，ORの分野で有効に用いられる**弾力性係数**も対立比の1種である。

関連ページ

オペレーションズ・リサーチ　48

■弾力性係数　ある人の収入が x 円のとき，その家庭での交際費支出が y 円であったとする。この人の収入が Δx 円だけ増加したので，交際費支出を Δy 円だけふやした。このとき，

$$\eta=\left(\frac{\Delta y}{y}\right)\Big/\left(\frac{\Delta x}{x}\right)$$

という量を考えると，この家計に占める交際費支出の意味がはっきりしてくる。上の η を，「収入 x についての交際費 y の弾力性係数」という。弾力性係数は単に**弾力性**とか**弾性値**ともよばれ，上のことを簡単に，「教育費の収入弾力性」ともいう。

(例)　所得税の計算(改正前)

階級	所得金額(千円)	税率(%)	控除額(千円)	税額(千円)	実効税率(%)	弾性値
4	2,000 ～2,399	16	72	248 ～312	12.4 ～13.0	1.29 ～1.23
5	2,400 ～2,999	18	120	312 ～420	13.0 ～14.0	1.38 ～1.29
6	3,000 ～3,999	21	210	420 ～630	14.0 ～15.8	1.50 ～1.33
7	4,000 ～4,999	24	330	630 ～870	15.8 ～17.4	1.52 ～1.38
8	5,000 ～5,999	27	480	870 ～1,140	17.4 ～19.0	1.55 ～1.42
9	6,000 ～6,999	30	660	1,140 ～1,440	19.0 ～20.6	1.58 ～1.46
10	7,000 ～7,999	34	940	1,440 ～1,780	20.6 ～22.2	1.65 ～1.53
11	8,000 ～9,999	38	1,260	1,780 ～2,540	22.2 ～25.4	1.71 ～1.50
12	10,000 ～11,999	42	1,660	2,540 ～3,380	25.4 ～28.2	1.65 ～1.49
13	12,000 ～14,999	46	2,140	3,380 ～4,760	28.2 ～31.7	1.63 ～1.45
⋮						
19	80,000 以上	75	14,240	45,760 以上	57.2 以上	1.31 以下

右の例は，とかく問題になった所得税の計算方式である(昭和50年に改正をみたが，右表は改正前のもの)。

何が問題かというといくつかあって，その1つは，中高所得層の重税感であった。これを表から，次のように読みとることができる。課税対象となる所得金額が2000千円(階級4)の人の税額は，

$$2000\times0.16-72=248\ (\text{千円})$$

であるが，これは所得に対して12.4%に当たる。この階級の人が(階級4以内にとどまる範囲で)増収となった場合，その分については16%の所得税を納めなければならない。したがって，「所得 x についての所得税 y の弾力性係数」(表では，これを単に弾性値として示してある)

$$\eta=\left(\frac{\Delta y}{y}\right)\Big/\left(\frac{\Delta x}{x}\right)=\left(\frac{\Delta y}{\Delta x}\right)\Big/\left(\frac{y}{x}\right)$$

を考えると，年収2000(千円)の人については，

$$\frac{y}{x}=0.124,\ \frac{\Delta y}{\Delta x}=0.16$$

であるから，$\eta=\dfrac{0.16}{0.124}=1.29$ になっている。ところが，年収8000千円(階級11)の人にとっては，これが1.71にも上っている。このようなことから，年収5000～6000(千円)以上の人における重税感がきびしいと判断されたわけである。　(牧野)

品質管理

JISでは，品質管理とは「買い手の要求に合った品質の品物またはサービスを経済的に作り出すための手段の体系」としている。したがって，品質管理では製造，サービスの企業などにおいて顧客を十分に満足させることのできる品質の商品を作り出し，販売する組織的な活動の体系といえる。

★解説

品質管理(Quality Control)はQCとも略称される。とくに，近代的な品質管理は統計的手法をよく用いるので，これを**統計的品質管理**(Statistical Quality Control),または略して**SQC**ということもある。

品質管理は一部の部門だけで行っても効果がないので,全社的に行うことが多い。たとえば，市場調査，研究・開発，製品企画，設計，生産から販売，アフターサービスの各段階を通して品質の維持と向上のための管理が進められるので，このような品質管理を**全社的品質管理**(company-wide quality control, Total Quality Control)，または略して**TQC**という。現在，日本の**TQC**はその名が外国にもよく知られている。

品質管理を通じて成果をあげるためには，種々の手法が必要となる。この中でも，統計的な手法に関するものとしては，**管理図**，**パレート解析**，**抜取検査**，**実験計画法**などがある。また，統計的ではない手法としては，**特性要因図**，品質機能展開，および**FMEA**，**FTA**などがあげられる。

品質管理は，とくにわが国がよく発達していることが，外国においても高く評価され，これが日本製の製品の優秀性にも貢献しているので，日本の品質管理を学ぼうとする国々が少なくない。

関連ページ

統計的品質管理　262

管理図　68
抜取検査　108
実験計画法　160
特性要因図　263
FMEA　36
FTA　36

■品質管理の歴史　品質管理のはじまりは，1930年前後にまでさかのぼる。当時，ベルシステムにいたシュハート(W. A. Shewhart)は，ベルシステムが電話機材の購入に当たって，きびしい検査を行っていたのに対して，電機メーカーの工程管理に重点をおくべきことを指摘した。彼は統計的手法を用いて工程を安定な状態に維持し，このことによって工程においてよい品質を作り込む品質管理を啓蒙したが，これは統計的品質管理またはSQCとよばれた。

その後，品質管理は米国の各工場の中にASME(米国機械学会)などの尽力によって浸透したが，第二次世界大戦中に，米国で品質管理を中心とした戦時規格Z 1.1, 1.2, 1.3が制定されるに至って，急速に定着した。

日本に，米国のSQCが初めて組織的に導入されたのは1950年前後である。このSQCは日本独自の管理技術として発達したので，1950年から1960年は日本のSQCの発展期といえる。

1960年を過ぎると，日本のSQCは企業の全体の活動に結びつけられるようになり，TQCという名でよばれるようになった。この間に，日本の製品の品質が著しく向上したので，この品質管理は海外の専門家にも注目されるようになり，これが"company-wide quality control"とよばれたことにより，全社的品質管理がTQCの別名のように用いられることも少なくなくなった。

1970年を過ぎると，品質管理と品質保証の関係も重要となり，製品の安全性のみではなく，消費者ニーズが多様化したので，商品の顧客の要求に対する適合性も重視され出した。

■品質の意味　品質管理でいう品質は，単なる性能や機能で表される狭義の品質を含んで，次の4つの考え方からなっている。

(1)狭義の品質……機能，デザイン
(2)量，納期の品質
(3)コストに関する品質
(4)アフターサービスの品質

たとえば，いかに狭い意味の品質が優れていても，納期が遅れたり，コストが高ければ意味がないので，品質の考え方は，上の4つの要素を総合してとらえねばならない。

■管理の意味　管理とは，簡単にいえば，"plan-do-check-action"の管理のサイクルによって説明される。企業が商品の企画と販売の計画を立てることはplanであり，商品を作って販売することはdoとなり，この結果と計画とのズレを検討し，その原因を追究することがcheckである。また，この原因を改善し，仕事の質の向上対策を講ずることをactionと考える。このサイクルを何度も回しながら，質の向上をはかることが管理である。(真壁)

FUNCAT による判別

FUNCAT (FUNctions of CATegorical responses as a linear model) は，汎用統計解析システム SAS にも収録されている質的データの判別手法である。数量化Ⅱ類と同様に，外的基準の分類を質的データの説明変数にカテゴリースコアを与えて判別を行う。

★解説

■アルゴリズム　外的基準が r 個に分類(群)され，説明変数が p 個のアイテムを入力データとする。次に，説明変数のすべてのカテゴリーの組み合わせ(サンプルとよぶ)が s 個とする。各 2 カテゴリーをもつ 3 アイテムでは 8 サンプルになる。入力データをこのサンプルと外的基準で 2 重クロス集計して，表 1 のデータを得る。たとえば，説明変数がサンプル 1 の値をとる組み合わせの件数が，G_1 群では n_{11} 件，G_r 群では n_{1r} 件あり，その小計が n_1 件になる。n_{11} から n_{1r} を n_1 で除したものは，サンプル 1 が与えられたときの r 個の群の表れる出現確率 P_{1i} とみなせる。FUNCAT では，この確率の間に適当な関数関係を決めて，これを外的基準とする。この外的基準の値を，サンプルを説明変数として一種の重回帰分析を行う。たとえば，表中の $(r-1)$ 個の組などが，外的基準として考えられる。

サンプル	$G_1\, G_2 \cdots\cdots G_r$	計	$G_1\, G_2 \cdots\cdots G_r$	外的基準
1	$n_{11}\, n_{12} \cdots n_{1r}$	n_1	$P_{11} P_{12} \cdots P_{1r}$	$\log(P_{11}/P_{1r}) \cdots \log(P_{1r-1}/P_{1r})$
2	$n_{21}\, n_{22} \cdots n_{2r}$	n_2	$P_{21} P_{22} \cdots P_{2r}$	$\log(P_{21}/P_{2r}) \cdots \log(P_{2r-1}/P_{2r})$
$\vdots$	$\vdots \quad \vdots \quad \vdots$	$\vdots$	$\vdots \quad \vdots \quad \vdots$	$\vdots \quad \vdots$
s	$n_{s1}\, n_{s2} \cdots n_{sr}$	n_s	$P_{s1} P_{s2} \cdots P_{sr}$	$\log(P_{s1}/P_{sr}) \cdots \log(P_{sr-1}/P_{sr})$

表 1　FUNCAT のデータ構造

関連ページ

説明変数		デザイン行列			頻	度		確	率	尤度比	外的基準	サンプルスコア
X_1	X_2	β_{11}	β_{12}	β_{21}	G_1	G_2	計	G_1	G_2			
1	1	1	0	1	8	2	10	0.8	0.2	8/2	log(4)	$\beta_{11}+\beta_{21}$
1	2	1	0	−1	2	3	5	0.4	0.6	2/3	log(2/3)	$\beta_{11}-\beta_{21}$
2	1	0	1	1	14	6	20	0.7	0.3	7/3	log(7/3)	$\beta_{12}+\beta_{21}$
2	2	0	1	−1	1	3	4	0.25	0.75	1/3	log(1/3)	$\beta_{12}-\beta_{21}$
3	1	−1	−1	1	3	2	5	0.6	0.4	3/2	log(3/2)	$-\beta_{11}-\beta_{12}+\beta_{21}$
3	2	−1	−1	−1	0	0	0	−	−			$-\beta_{11}-\beta_{12}-\beta_{21}$
				合計	28	16	44					

表1　FUNCATのデザイン行列と外的基準

■FUNCATの扱うデータ　外的基準が2群 G_1 (28件)と G_2 (16件)で与えられている2群判別の問題を考える。説明変数は X_1 (3 カテゴリー)と X_2 (2カテゴリー)の2アイテムとする。

FUNCATでは，数量化Ⅱ類と同じく，44件のデータが入力される。そして表1に示すように，説明変数のすべてのカテゴリーの組み合わせと外的基準のクロス集計を行う。$X_1=1$，そして $X_2=1$ のデータをサンプル(1, 1)とよぶことにする。サンプル(1, 1)は G_1 群が8件，そして G_2 群が2件の計10件である。サンプル(3, 2)は0件である。この頻度表の各行の値を，その行の和で除したものが次の確率表である。確率の比は尤度比とよばれる。

■アルゴリズム　アイテム X_1 はデザイン行列の最初の2列に対応する。すなわち，カテゴリー1, 2, 3がそれぞれ(1, 0), (0, 1), (−1, −1)になる。アイテム X_2 は3列目に対応する。FUNCATでは，このデザイン行列の3列を新しく説明変数とみなし，尤度比の対数をとったものを外的基準とする重回帰分析の問題として扱う。すなわち，カテゴリースコア β_{11}, β_{12}, β_{21} が求まれば，サンプルスコアが計算できる。この値が外的基準のよい予測値になるように β_{11}, β_{12}, β_{21} が決められる。サンプル(3, 2)はデータが0件のため計算に用いられない。しかし，サンプルスコアの値が $-(\beta_{11}+\beta_{12}+\beta_{21})$ であり，この値が正の場合には G_1 群の確率が G_2 群の確率より大きいので，このサンプルは G_1 群に判別すればよい。同様にして，$(\beta_{12}-\beta_{21})$ の値が負であれば G_2 群の確率の方が G_1 の確率より大きいので，サンプル(2, 2)は G_2 群に判別できる。

ここで考えた外的基準は，G_i 群の確率を P_i とすれば $\log(P_1/P_2)=\log(P_1/(1-P_1))$ になり，ロジット変換とよばれるものになる。　(新村)

フェース法

多次元の尺度で測る状況を総合的に判断するため，人間の顔をこれら尺度の値で構成し，その表情を見て直観的な判定を下そうという方法。1971年にチャーノフ(Chernoff, H.)が原型を発表したが，その後いろいろに変形されながら各所で用いられている。人間の持つパターン認識の特技を積極的に利用するという点ですぐれているが，適用対象ごとに表情の持つ意味が異なるから，判定には若干の修練も必要になる。

★解説

■チャーノフの顔　チャーノフによると，顔の6つの部分に，計18個の尺度を割り当てている。その概形は，図1及び下表に示すとおりである。

変数	顔の特性
1	中心からSまでの長さ
2	Sの方向
3	顔の長さ（=UL）
4	顔の上半分のふくらみ
5	顔の下半分のふくらみ
6	鼻の長さ
7	口の位置
8	口の曲がり方
9	口のはば
10	目の位置
11	目の中心の離れぐあい
12	目の傾き
13	目の形
14	目の大きさ
15	ひとみの位置
16	眉の位置
17	眉の傾き
18	眉の長さ

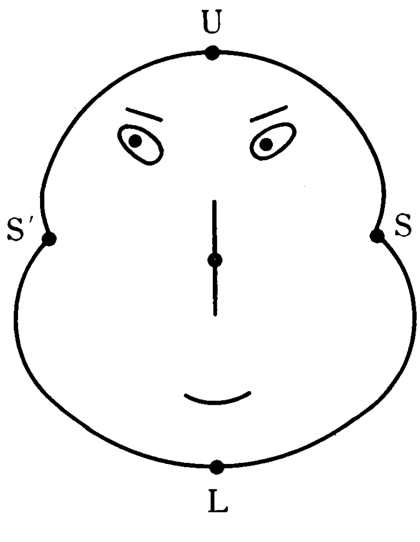


図1

関連ページ
発想的OR　294

左ページに示した変数の使い方を，もう少し詳しく説明する。

(1)**顔の形**　中心(Oとする)からSまでの距離を変数 x_1 を用い，$a(1+x_1)/2$ と定める。a は任意で，大きさを定める定数である。OS の方向は水平方向とのなす角で定める。この角度は，変数 x_2 を用いて，$(2x_2-1)\times 45°$ とする。顔の形はすべて左右対称とする。顔の縦方向の長さは，x_3 を用いて $a(1+x_3)$ と定め，これを b とおく。

図1で顔は上，下2つの部分に分かれ，別々のふくらみを持っていることが見られるが，これは，変数 x_4，x_5 を用いてそれぞれ定められている。すなわち，どちらもSと U(L) を通る楕円であるが，その離心率が x_4 (x_5) となっている。

(2)**鼻**　原点を中心とし，垂直方向に長さ bx_6 の線分として表す。

(3)**口**　口の中心はOの下 $h(x_7+(1-x_7)x_6)$ の位置に置く。そこを通り，半径 $b/|x_8|$ の円を書き，$x_8>0$ のときは中心を上方に置いて下に凸の形とし，$x_8<0$ のときは逆向きにする。この半径が b よりも小さければ，半径の x_9 倍を口のはばとして定め，顔からはみ出すときは，口の中心と顔の外周までの水平距離の x_9 倍として，外に出ないよう考慮する。

(4)**目**　目は楕円で表すが，その中心は，縦座標が $b(x_{10}+(1-x_{10})x_6)$，横座標が顔の外周までの長さの $(1+2x_{11})/4$ 倍に等しい位置である。目の傾き θ は $(2x_{12}-1)\times 36°$，目のはばは中心の横座標の x_{14} 倍を原則とする。ただし，目の中心が半分より外側に寄っているときは,外周までの距離の x_{14} 倍に修正する。このはばが目の楕円の長軸の長さに等しい。楕円の離心率は x_{13} である。

■**ひとみ**　目の楕円の中心から $(2x_{15}-1)\cdot C$ の位置に置く。ここで，C はややめんどうで，目のはばに $(d=\cos^2\theta+\sin^2\theta/x_{13}^2)^{-\frac{1}{2}}$ を乗じたものである。

(5)**眉**　目の中心から，そのはばの $x_{13}(x_{16}+0.3)$ の位置に，傾き $\theta+2(x_{17}-1)\times 36°$，はば $d(2x_{18}+1)/2$ を満たす線分として与えられる。

■**標準化**　上で用いられた変数はすべて標準化してある。これは z_i を原データ,その最大値を M_i,最小値を m_i とするとき,次表に与えられる α_i，β_i を用いて，

$$x_i=d_i-(\beta_i-\alpha_i)\,\frac{z_i-m_i}{M_i-m_i}$$

として計算される。

i	1	2	3	4	5	6	7	8	9
α_i	1	1	1	2	2	1	1	5	1
β_i	0	0	0	0.5	0.5	0	0	−5	0

i	10	11	12	13	14	15	16	17	18
α_i	1	1	1	0.8	1	1	1	1	1
β_i	0	0	0	0.4	0	0	0	0	0

参考論文：雑誌『数理科学』1973年12月号の萱原・八柳両氏による論文に応用例もある。　　（森村）

2つの母集団の母数の差の検定

2つの母集団の母数，たとえば母平均 μ_1，μ_2 に関して，それぞれの母集団からとられた任意標本 $(X_1,\ X_2,\ \cdots,\ X_{n_1})$，$(Y_1,\ Y_2,\ \cdots,\ Y_{n_2})$ をもとにして，$\mu_1=\mu_2$ という仮説を検定したいことがある。このような検定を，2つの母集団の平均の差の検定という。2つの母集団の母数の検定として，2つの母集団の比率の差の検定や分散についての検定なども行われる。

★解説

2つの**母集団**からとられた2組の**標本**に基づいて，**母平均** μ_1，μ_2 の間に有意差があるかどうかを**検定**したいことがある。このような検定のことを，**母平均の差の検定**(平均値の差の検定)，くわしくは母平均の差についての**有意差検定**という。検定の結果，“有意差が認められる”ということは，標本の間にみられる差が，**標本抽出**にともなう誤差の範囲をこえていると判定される，ということである。検定の結果，“有意差あり”と判定されたとしても，差の具体的な解釈は別問題である。有意差ありと出る場合は，問題領域の知見からすれば，それが自明であることが多い。

母平均の差の検定を行う手順は，母数の検定と同じで，

帰無仮説　H_0；$\mu_1=\mu_2$

対立仮説　H_1；$\mu_1 \neq \mu_2$

とおけば，**両側検定**になる。しかし，「μ_1 が μ_2 より小さいはずがない」ということが事前にわかっているものとすると，対立仮説を $\mu_1>\mu_2$ とおいて，**片側検定**することになる。片側検定の方が，両側検定よりも H_0 を棄却しやすい状況にあることはいうまでもない。一般に，仮説検定においては，帰無仮説を棄却することに意味があるので，両側検定よりも片側検定の方がよい判断ができる。それは，μ_1 が μ_2 より小さくないという事前情報のもたらす効果であるともいえる。

関連ページ
母集団　260
標本　260
検定　65
標本抽出　258
帰無仮説　62
対立仮説　62
両側検定　64
片側検定　64

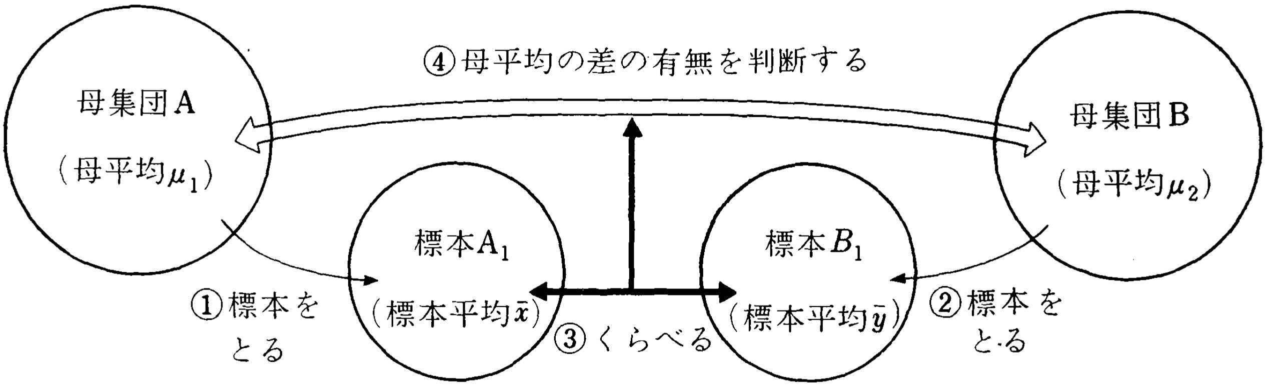


図1　母平均の差の検定の考え

2つの正規母集団 $N(\mu_1,\ \sigma_1^2)$, $N(\mu_2,\ \sigma_2^2)$ からの，それぞれ大きさ n_1, n_2 の任意標本を，

$$(X_1,\ X_2,\ \cdots,\ X_{n_1})$$
$$(Y_1,\ Y_2,\ \cdots,\ Y_{n_2})$$

とする。これを用いて，母平均の差の検定，等分散性の検定を行うには，次のようにする。

■母平均の差の検定　上で，分散は同じ σ^2 であると考えてよいことがわかっているとき，母平均について，

帰無仮説　H_0 ; $\mu_1=\mu_2$

対立仮説　H_1 ; $\mu_1 \neq \mu_2$

を検定するには，統計量

$$T=\frac{\overline{X}-\overline{Y}}{\sqrt{n_1 s_1^2+n_2 s_2^2}}\cdot\sqrt{\frac{n_1 n_2(n_1+n_2-2)}{n_1+n_2}}$$

を計算する。ただし，s_1^2, s_2^2 は，それぞれの標本における標本分散である。上の T が自由度 (n_1+n_2-2) の t 分布（→314ページ）にしたがうので，t 分布表を用いて，

$$|T|>t_{n_1+n_2-2}(0.05)$$

ならば有意水準5％で，また，

$$|T|>t_{n_1+n_2-2}(0.01)$$

ならば有意水準1％で，帰無仮説 $\mu_1=\mu_2$ を棄却して $\mu_1 \neq \mu_2$ と判定する。ただし，$t_{n_1+n_2-2}(\alpha)$ は，自由度 (n_1+n_2-2) の t 分布での両側100パーセント点である。

■等分散の検定　左で等分散を仮定したが，これを検定するには，標本から不偏分散

$$\frac{n_1 s_1^2}{n_1-1},\ \frac{n_2 s_2^2}{n_2-1}$$

を計算し，その大きいほう（かりに添字1をつけたほう）を分子にして，両者の比

$$F=\frac{n_1 s_1^2/(n_1-1)}{n_2 s_2^2/(n_2-1)}$$

をつくる。この比のことを，**不偏分散比**とよぶ。上の統計量 F は，自由度 $(n_1-1,\ n_2-1)$ の F 分布（→315ページ）にしたがうので，F 分布表を用いて検定する。　（牧野）

不平等度の計測

データの均等度，逆にいうと不平等度を表すのに，ローレンツ曲線やジニ係数が用いられる。たとえば，所得金額の不平等度を測るのに，金額の小さい方から累積して，累積人数百分率，累積金額百分率を計算し，それらを横座標，縦座標としてグラフをかく。このようなグラフをローレンツ曲線という。また，データの平均偏差を用いて，不平等度を測ることもできる。ジニ係数は，そのような量である。

★解説

所得格差に代表される，格差とか不平等とよばれる現象を計測するのに，しばしば**ローレンツ曲線**が用いられる。右ページの図に示すように，横軸 100 %，縦軸 100 %の点と原点を結ぶ直線のことを均等線とよび，ローレンツ曲線と均等線とで囲まれた部分の図形を，**不平等度を表す弓形**という。弓形の面積 λ が大きいほど，不平等の度合いが強い。一方，データの 2 つずつの対をとって，**平均偏差**を計算し，それを**平均値**の 2 倍で割ったものをジニ係数とよぶが，これはまた，弓形の面積を，均等線から下の三角形の面積で割った値になっている。ジニ係数がゼロというのは，データが均等線上に並ぶ場合に限り，逆に格差が大きいときには，ジニ係数は 1 に近づく。

OR，とくに在庫管理の分野で，**パレート図**とよんでいる曲線は，本質的にはローレンツ曲線と同じものである。ただ，ローレンツ曲線をかくには，データを小さい方から順に累積するのに対し，パレート図は，大きい順に累積してかかれているという違いがあるだけである。したがって，これら 2 つの曲線は，右ページの図に示す正方形の 2 つの対角線の交点に関して，互いに対称な曲線になる。

集中という概念は，格差の裏返しであるから，ローレンツ曲線は集中の度合いを測る量であるということもできる。

関連ページ

オペレーションズ・リサーチ　48
パレート図　300

■ローレンツ曲線とジニ係数 右に簡単な例を示してあるが，実際には，多くのデータをもとに，累積百分率の表をつくり，ローレンツ曲線をかくので，ローレンツ曲線は，右図のような折れ線でなく，もっと滑らかな曲線になるのがふつうである。ただし図1では，100％を1と目盛ってある。

ジニ係数は次のようにして求められる。いまデータがn個あって，

x_1, x_2, …, x_n

であるとするとき，すべてのi，jについて$|x_i-x_j|$を計算し，それらの和を求め，n^2で割る。これが平均偏差である。これを，平均値の2倍で割るとジニ係数が求まる。

右の例では，$n=3$であるから，

$$\text{平均偏差}=\frac{1}{3^2}\{|12-15|+|12-27|+|15-12|+|15-27|+|27-12|+|27-15|\}=\frac{60}{9}$$

で，平均値＝18であるから，ジニ係数M_Gは，

$$M_G=\left(\frac{60}{9}\right)/(2\times18)=\frac{5}{27}=0.185$$

になる。

これが，図1の弓形の面積λを，三角形の面積で割った値になっていることを，容易に確かめることができる。

ところで，3人に期末手当として，本俸の2ヵ月分，つまり，

(例) ある会社の3人の本俸が，それぞれ，12，15，27(万円)のとき；

本俸 (万円)	累積人数		累積金額	
	人数	百分率(%)	金額(万円)	百分率(%)
12	1	33.3	12	22.2
15	2	66.7	27	50.0
17	3	100.0	24	100.0

表1 累積百分率の表

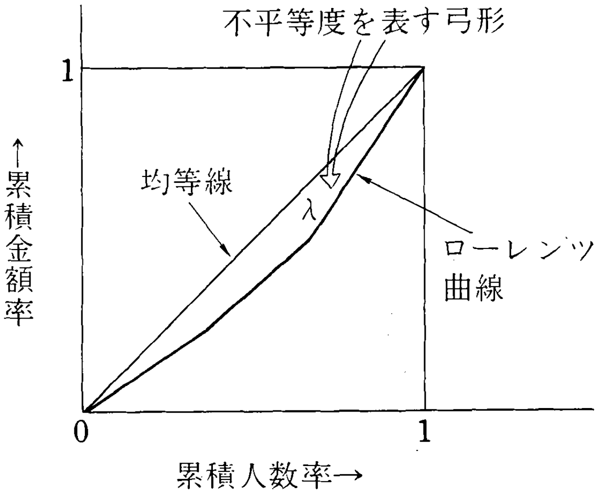


図1 ローレンツ曲線

24, 30, 54 (万円)

が支給されたとするとき，不平等度を表す弓形の面積はいくらになるかを考えてみる。この場合，累積人数百分率と累積金額百分率を計算してみると，上の表と全く同じ値になる。したがって，ローレンツ曲線は変わらない。もちろん弓形の面積も同じである。実は，この例に限らず，データを定数倍しても，ローレンツ曲線は変わらないということを，容易に証明することができる。しかし，期末手当として，本俸の2ヵ月分プラス一律5万円というように，一律加算分が加わると，不平等度を表す弓形の面積は，もとの面積よりも小さくなる。参考文献〔52〕 (牧野)

分割表による検定

n 個のデータを，2つの属性 A，B に注目して分類したとき，(A_i, B_j) に属するものの度数を n_{ij} とし，表1の形にまとめられたとする。このような表のことを，$l \times m$ 分割表という。分割表を用いて A，B の独立性，すなわち $\mathrm{P}(A_i \cap B_j)=\mathrm{P}(A_i)\cdot\mathrm{P}(B_j)$ を検定する問題を，分割表内の**独立性の検定**という。これは適合度検定の一種であって，χ^2 分布表を用いて検定することになるが，とくに応用されることの多いものである。

★解説

表1を用いて A と B との独立性の検定の仕方を説明する。

帰無仮説　H_0；A と B は**独立**である

対立仮説　H_1；A と B は独立でない

を検定するわけであるが，そのために各マス目にはいっている**実現値**に対応する**理論値**(**期待値**)を計算する。i 行 j 列の実現値 n_{ij} に対して，理論値は次のように計算される。もし，A_i と B_j とが独立であるとすると，あるデータがそのマス目におちる確率は，

$$\frac{n_{i\cdot}}{n} \times \frac{n_{\cdot j}}{n}$$

であるから，理論値は，

$$n_{i}.n._{j/n}$$

である。そこで，これと実現値とのくい違いを評価するために，ふつうの**適合度検定**のときと同じように，

$$\chi^2=\left(n_{ij}-\frac{n_{i}.n._{j}}{n}\right)^2 \Big/ \left(\frac{n_{i}.n._{j}}{n}\right)$$

を計算し，すべての i，j について加えあわせる。このような統計量 χ^2 が，近似的に自由度 $(l-1)(m-1)$ の **χ^2 分布**にしたがうので，χ^2 分布表を用いて仮説を検定することができる。

関連ページ

帰無仮説　62
対立仮説　62

理論値　248
期待値　58

適合度検定　248

χ^2 分布　314

帰無仮説H_0：AとBは独立である
対立仮説H_1：AとBは独立でない

分割表

A \ B	B_1	B_2	…………	B_i	…………	B_m	計
A_1	n_{11}	n_{12}		n_{1j}		n_{1m}	$n_{1\cdot}$
A_2	n_{21}	n_{22}		n_{2j}		n_{2m}	$n_{2\cdot}$
⋮							⋮
A_i	n_{i1}	n_{i2}		n_{ij}		n_{im}	$n_{i\cdot}$
⋮							⋮
A_l	n_{l1}	n_{l2}		n_{lj}		n_{lm}	$n_{l\cdot}$
計	$n_{\cdot 1}$	$n_{\cdot 2}$	…………	$n_{\cdot j}$	…………	$n_{\cdot m}$	n

[χ^2値の計算]$=\left[\left\{\left(n_{ij}-\frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2\Big/\left(\frac{n_{i\cdot}n_{\cdot j}}{n}\right)\right\}\right.$をすべての$i, j$について計算して加えあわせる]

自由度$(l-1)(m-1)$としてχ^2分布表から$\chi^2_{(l-1)(m-1)}(0.05)$の値を読む。上の$\chi^2$値がこの値よりも大きければ，有意水準5％で帰無仮説を棄却。

表1　独立性の検定の手順

■税務署別の申告所得金額　京橋税務署（東京都），松戸税務署（千葉県）管内において，昭和57年の年間所得（申告）金額が1000万円以上であった人の資料を分類して，表2，表3がつくられている。表2は，所得金額1000万円以上，2000万円未満についてのものであり，表3は2000万円以上，3000万円未満についてである。このような分割表に基づいて，各税務署と所得金額の独立性について検定してみる。そのために，表の各マス目にはいる理論値（期待値）を計算する。マス目の（　）内の数値がそれである。これは，たとえば表2について，次のように計算される。1行1列の300に対応する理論値は，

$$1897\times 867/4644=354.2$$

税務署 \ 金額（万円）	1,000以上~1,200未満	1,200~1,400	1,400~1,600	1,600~1,800	1,800~2,000	計
京橋	300 (354.2)	237 (209.8)	163 (134.6)	85 (95.2)	82 (73.2)	867
松戸	1,597 (1,542.8)	887 (914.2)	558 (586.4)	425 (414.8)	310 (318.8)	3,777
計	1,897	1,124	721	510	392	4,644

χ^2値$=25.38$，$\chi^2_4(0.01)=13.277$

表2　税務署別の申告所得金額（1,000万円以上~2,000万円未満）

税務署 \ 金額（万円）	2,000以上~2,200未満	2,200~2,400	2,400~2,600	2,600~2,800	2,800~3,000	計
京橋	77 (65.9)	49 (50.3)	36 (50.3)	45 (41.1)	37 (36.3)	244
松戸	257 (268.1)	206 (204.7)	219 (204.7)	163 (166.9)	147 (147.7)	992
計	334	255	255	208	184	1,236

χ^2値$=7.91$，$\chi^2_4(0.05)=9.488$

表3　税務署別の申告所得金額（2,000万円以上~3,000万円未満）

2行5列については，

$$392\times 3777/4644=318.8$$

などである。これらを用いて，χ^2値を求める。すなわち，

$$\chi^2=(300-354.2)^2/354.2=25.38$$

自由度$(2-1)\times(5-1)=4$のχ^2分布での1％点をχ^2分布表から読むと，

$$\chi^2_4(0.01)=13.277$$

であるから，上で求めたχ^2値は$\chi^2_4(0.01)$の値より大きい。よって，税務署と所得金額との独立性は，有意水準1％で棄却される。同様の計算を，表3について行ってみると，独立性が棄却されないことがわかる。　（牧野）

分割法

1つの因子Aの水準を割り付けた実験単位を分割して，他の因子Bの水準を割り付ける実験。分割法では無作為化を何回にも分けて行うため，水準変更が完全無作為化実験に比べて少なくなり，実験が容易。

★解説

ある化学反応工程で収率向上のため，2**因子**A(反応温度，80°，100°，120°の3水準)，B(反応時間，5分，10分，15分の3水準)をとりあげ，交互作用 $A\times B$ が存在する可能性もあるので，繰り返し2回,計18回の実験を計画した。

18回の実験をすべて**無作為化**すると，最悪の場合には，17回の反応温度の水準変更をする可能性がある。もし,温度の水準変更が難しいあるいは時間を要する場合には，まず反応温度Aの3水準の実験順序を無作為化してAの水準の順番を決めて，つぎに反応時間Bの3水準の実験順序を無作為化する。そして，このような実験を2回反復したほうのが，温度の水準変更は多くても5回となるから，より実験が容易になる。

この実験方法を分割法とよび，この方法による実験を分割実験という。また，Aの各水準を実験する場を1次単位，1次単位の実験の場を分割して得られる実験の場を2次単位とよぶ。そして，1次単位の因子を1次因子，2次単位の因子を2次因子という。さらに,この分割実験では,分割が1段階で行われているので1段分割実験とよぶ。2次単位をさらに分割して3次単位に因子Cを割り付けて，分割を2段階に分けて実験を行うこともできる。この分割実験は2段分割実験とよぶ。一般に,多段分割実験を考えることができる。

分割法は，完全無作為化実験よりも水準変更の回数が少なくてすむので，実用的には重要なものである。

関連ページ
因子　20
無作為化　384

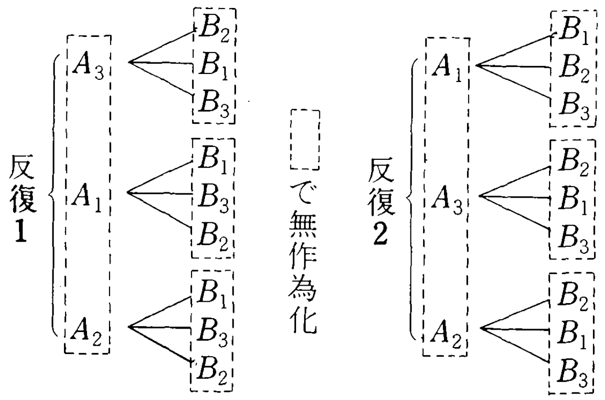


図 1　1段分割法

1段分割法の実験データ　左ページの分割法による実験順序は，図1のようになる。分割実験をした場合には，同一1次単位のデータ，たとえば第1反復での A_3B_2，A_3B_1，A_3B_3 は互いに独立ではなく，共通の誤差を含むことになる。これは無作為化を2回に分けて行っていることによる。

このことを考慮すると，第 k 反復での A_iB_j におけるデータ x_{ijk} の構造模型は，

$$x_{ijk}=\mu+\gamma_k+\alpha_i+\varepsilon_{ik}^{(1)}+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk}^{(2)}$$

ここで　μ：一般平均

γ_k：反復 k の効果

α_i：A_i 水準の効果

$\varepsilon_{ik}^{(1)}$：1次誤差

β_j：B_j 水準の効果

$(\alpha\beta)_{ij}$：A_iB_j の交互作用

$\varepsilon_{ijk}^{(2)}$：2次誤差

となり，誤差には1次誤差と2次誤差の2種類がある。

1段分割法の分散分析　データの構造模型より総平方和 S_T は，

$$S_T=S_R+S_A+S_{E_1}+S_B+S_{A\times B}+S_{E_2}$$

とすることができる。A の水準数を l，B の水準数を m，反復の回数を n とすれば，これらの平方和は，

$$CT=(x_{111}+x_{112}+\cdots+x_{lmn})^2/lmn$$
$$S_T=(x_{111}{}^2+x_{112}{}^2+\cdots+x_{lmn}{}^2)-CT$$
$$S_A=(T_{1\cdot\cdot}{}^2+\cdots+T_{l\cdot\cdot}{}^2)/mn-CT$$
$$S_R=(T_{\cdot\cdot1}{}^2+\cdots+T_{\cdot\cdot n}{}^2)/lm-CT$$
$$S_{AR}=(T_{1\cdot1}{}^2+\cdots+T_{l\cdot n}{}^2)/m-CT$$
$$S_{E_1}=S_{AR}-(S_R+S_A)$$
$$S_B=(T_{\cdot1\cdot}{}^2+\cdots+T_{\cdot m\cdot}{}^2)/ln-CT$$
$$S_{AB}=(T_{11\cdot}{}^2+\cdots+T_{lm\cdot}{}^2)/n-CT$$
$$S_{A\times B}=S_{AB}-(S_A+S_B)$$
$$S_{E_2}=S_T-(S_A+S_R+S_{E_1}+S_B+S_{A\times B})$$

と求めることができる。ここで $T_{1\cdot\cdot}$ は A_1 水準におけるデータの和を示す。他の記号も同様。これを分散分析表にまとめると表1のようになる。

要因	平方和	自由度	不偏分散	不偏分散比
反復(R)	S_R	$\phi_R=n-1$	$V_R=S_R/\phi_R$	V_R/V_{E_1}
A	S_A	$\phi_A=l-1$	$V_A=S_A/\phi_A$	V_A/V_{E_1}
1次誤差(E_1)	S_{E_1}	$\phi_{E_1}=(l-1)(n-1)$	$V_{E_1}=S_{E_1}/\phi_{E_1}$	V_{E_1}/V_{E_2}
B	S_B	$\phi_B=m-1$	$V_B=S_B/\phi_B$	V_B/V_{E_2}
$A\times B$	$S_{A\times B}$	$\phi_{A\times B}=(l-1)(n-1)$	$V_{A\times B}=S_{A\times B}/\phi_{A\times B}$	$V_{A\times B}/V_{E_2}$
2次誤差(E_2)	S_{E_2}	$\phi_{E_2}=l(m-1)(n-1)$	$V_{E_2}=S_{E_2}/\phi_{E_2}$	
計	S_T	$\phi_T=lmn-1$		

表 1　分散分析表

この1段分割実験では，1次単位の要因効果は1次誤差に対して，2次単位の要因効果は2次誤差に対して比較しなければならない。また，1次誤差は2次誤差と比較することにより，その変動が大きいか否かを判定することができる。（宮村）

分散

データの散らばりの度合いを表す特性値として，いちばんよく用いられる量は分散およびその平方根である標準偏差である。分散は，データのそれぞれの値から平均をひき，それらを2乗して加え合わせたものを，データの数で割ることによって求められる。

★解説

■分散の求め方　n 個のデータ x_1, x_2, …, x_n があるとき，分散 s^2 は次式によって求められる。

$$s^2=\frac{1}{n}\{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots+(x_n-\overline{x})^2\} \tag{1}$$

ただし，$\overline{x}$ はデータの**平均値**である。

分散を求めるには，**仮の平均** a を用いて，次のように計算してもよい。

$$s^2=\frac{1}{n}\{(x_1-a)^2+(x_2-a)^2+\cdots+(x_n-a)^2\} -(\overline{x}-a)^2 \tag{2}$$

x_1 のものが f_1 個，x_2 のものが f_2 個，…，x_k のものが f_k 個というように，データが整理されているときには，

$$s^2=\frac{1}{n}\{(x_1-a)^2f_1+(x_2-a)^2f_2+\cdots+(x_k-a)^2f_k\} -(\overline{x}-a)^2 \tag{3}$$

を用いるとよい。ただし，$f_1+f_2+\cdots+f_k=n$ である。

また，ふつうよく用いられる分散の計算法は，

$$s^2=(\text{データを2乗したものの平均}) -(\text{データの平均})^2 \tag{4}$$

である。

関連ページ

平均値　344
仮の平均　345

(例)　ある営業所でのA，Bグループ，それぞれ20人ずつによる1週間の売上高が次表のようであった。

売上金額

金額（百万円）	Aグループ	Bグループ
3	6 人	9 人
4	18	12
5	1	4
計	25 人	25 人

1人あたり平均金額は，いずれも380万円であるが，Bグループによる売上高のバラツキは，Aグループのそれよりも大きい。右図は，その様子を示している。

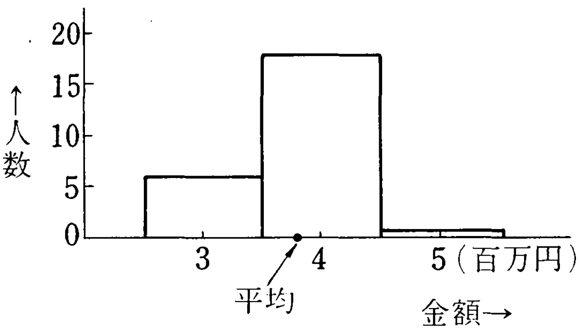


図1　Aグループの売上高

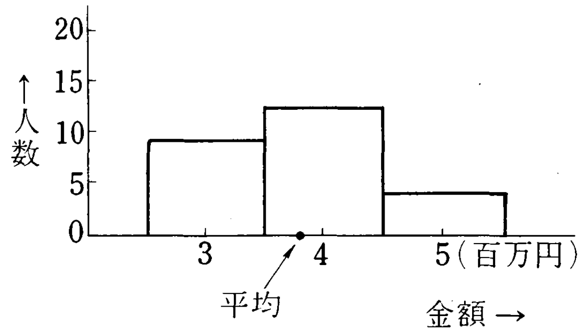


図2　Bグループの売上高

上の例で，Aグループの売上金額の分散を計算してみる。

$x_1=3,\ x_2=4,\ x_3=5$；

$f_1=6,\ f_2=18,\ f_3=1$；

仮の平均　$a=4$

として，左ページの式(3)を用いる。$n=25,\ \overline{x}=3.8$ であるから，

$$s_1^2=\frac{1}{25}\times\{(3-4)^2\times 6+(4-4)^2\times 18$$
$$+(5-4)^2\times 1\}-(3.8-4)^2$$
$$=\frac{7}{25}-0.04=0.24$$

これを，式(4)を用いて算出するには，次のようにする。

（データの2乗の平均）

$$=3^2\times\frac{6}{25}\times 4^2\times\frac{18}{25}+5^2\times\frac{1}{25}$$
$$=14.68$$

（データの平均の2乗）

$$=3.8^2=14.44$$

よって，

$$s_1^2=14.68-14.44=0.24$$

標準偏差は，

$$s_1=\sqrt{0.24}=0.49$$

になる。

同様に，Bグループの分散を計算すると，

$$s_1^2=0.48$$

標準偏差は，

$$s_2=\sqrt{0.48}=0.69$$

になる。

これらのことから，Bグループの分散のほうが，Aグループのそれよりも大きいことがわかる。　（牧野）

分散共分散行列と相関行列

変数 $x_1, x_2, \cdots, x_p$ において，x_i の分散を σ_{ij}，x_i と x_j の共分散を σ_{ij}，相関係数を ρ_{ij} で表す。分散共分散行列 Σ と相関行列 P は，それぞれ p 行 p 列の行列で次のようなものである。

$$\Sigma=\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix} \qquad P=\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{pmatrix}$$

★解説

変数 x_i と x_j の**共分散** σ_{ij} と，x_j と x_i の共分散 σ_{ij} は等しいから，分散共分散行列 Σ は対角要素に関して対称な対称行列である。相関係数に関しても $\rho_{ij}=\rho_{ji}$ であり，相関行列も対称行列である。

標本から計算した**分散を s_{ij}**，x_i と x_j の**共分散を s_{ij}** で表すと，**標本相関係数 r_{ij}** は，

$$r_{ij}=\frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}}$$

である。標本分散共分散 S と標本相関行列 R は，それぞれ次のようにかける。

$$S=\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix} \qquad R=\begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \cdot & \cdot & \cdots & \cdot \\ r_{p1} & r_{p2} & \cdots & 1 \end{pmatrix}$$

母分散共分散行列 Σ と標本分散共分散行列 S とが，どちらをさしているのか，文章の前後関係から明らかなときは，単に分散共分散行列と書くのが普通である。相関行列に関しても同じである。

関連ページ
共分散　229
分散　332

中学3年生137人の国語と社会の成績は次のようであった。

生徒番号	1	2	…	…	137
国語x_1	58	23	…	…	60
社会x_2	92	60	…	…	64

それぞれの教科の平均は，

$$\bar{x}_1=\frac{58+23+\cdots+60}{137}=45.9$$

$$\bar{x}_2=\frac{92+60+\cdots+64}{137}=66.0$$

である。また，国語 x_1 の分散 s_{11}，社会 x_2 の分散 s_{22} は，

$$s_{11}=\frac{1}{137-1}\{(58-45.9)^2+(23-45.9)^2+\cdots+(60-45.9)^2\}=343.9$$

$$s_{22}=\frac{1}{137-1}\{(92-66.0)^2+(60-66.0)^2+\cdots+(64-66.0)^2\}=534.4$$

となる。これは母分散 σ_{11}，σ_{22} の推定値である。母共分散 σ_{12} の推定値 s_{12} は，

$$s_{12}=\frac{1}{137-1}\{(58-45.9)(92-66.0)+(23-45.9)(60-66.0)+\cdots+(60-45.9)(64-66.0)\}=359.9$$

になる。国語 x_1 と社会 x_2 で説明したが，一般に任意の2教科 x_i と x_j についても平均 $\bar{x}_i$，$\bar{x}_j$ 分散 s_{ii}，s_{jj} 共分散 s_{ij} の計算は同様である。9教科であれば9個の分散と，$9(9-1)/2=36$ 個の共分散が得られる。

137人の生徒の成績に基づく9教科の分散行列は次のようになる。

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
国語 x_1	344								
社会 x_2	360	534							
数学 x_3	410	530	858						
理科 x_4	399	547	624	680					
音楽 x_5	358	447	563	517	612				
美術 x_6	279	367	406	388	344	510			
体育 x_7	226	284	347	335	293	215	462		
技家 x_8	373	452	520	518	489	376	327	619	
英語 x_9	443	541	678	631	539	389	346	523	772

分散共分散行列から，

$$s_{11}=344 \qquad s_{22}=534 \qquad s_{12}=360$$

より，x_1 と x_2 の相関係数は，

$$r_{12}=\frac{360}{\sqrt{344}\sqrt{534}}=0.84$$

を得る。同様の手順で36個の相関係数を計算すると，相関行列

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
国語 x_1	1						
社会 x_2	.840	1					
数学 x_3	.756	.783	1				
理科 x_4	.826	.908	.817	1			
音楽 x_5	.780	.783	.777	.802	1		
美術 x_6	.667	.703	.614	.659	.616	1	
体育 x_7	.566	.571	.551	.598	.552	.444	1
技家 x_8	.808	.786	.714	.798	.794	.669	.611
英語 x_9	.859	.843	.833	.871	.784	.621	.579

	x_8	x_9
技家 x_8	1	
英語 x_9	.757	1

を得る。これは母相関行列 P の推定値である。

多変量解析では，分散共分散行列と相関行列が，分析において非常に重要な役割を演じる。 (杉山)

分散分析

実験により得られたデータ全体の平方和を，いくつかの要因効果に対応する平方和と，その残りの誤差平方和に分解して検定や推定を行うこと。

★解説

一般にデータ x_1，x_2，…，x_n が与えられているとき，これらデータの変動は，個々のデータの全体の平均値 $(\overline{x}.)$ からの偏差の 2 乗和

$$S=(x_1-\overline{x}.)^2+(x_2-\overline{x}.)^2+\cdots\cdots+(x_n-\overline{x}.)^2$$

ここで，

$$\overline{x}.=(x_1+x_2+\cdots+x_n)/n$$

として求めることができる。これを(偏差)平方和とよぶ。

平方和 S は，見かけ上

$$(x_1-\overline{x}.),\ (x_2-\overline{x}.),\ \cdots\cdots,\ (x_n-\overline{x}.)$$

の n 個の成分の 2 乗和であるが，これら n 個の成分はすべて加えると 0 になるという関係

$$(x_1-\overline{x}.)+(x_2-\overline{x}.)+\cdots\cdots+(x_n-\overline{x}.)$$
$$=x_1+x_2+\cdots+x_n-n\overline{x}.=0$$

がある。したがって，独立な成分の数は $(n-1)$ である。このことから，平方和 S は $(n-1)$ 個の独立な成分の 2 乗和であると考えられる。このように，平方和が独立な成分の何個の 2 乗和になっているかを考えたとき，独立な成分の個数を平方和の自由度とよぶ。また，平方和 S をその自由度 $(n-1)$ で除したもの $V=S/(n-1)$ を不偏分散または平均平方とよぶ。

平方和 S をさらに意味のある要因別に分解することを平方和の分解といい，これには，データがどのような実験計画の下でとられているかを明確にする必要がある。この意味で，後の解析の目的を考慮した計画的なデータの収集が極めて重要である。

関連ページ

要因効果　394
実験計画法　160
構造模型　110

水準	繰り返し				和	平均
A_1	x_{11}	x_{12}	……	x_{1r}	$T_{1\cdot}$	$\bar{x}_{1\cdot}$
A_2	x_{21}	x_{22}	……	x_{2r}	$T_{2\cdot}$	$\bar{x}_{2\cdot}$
⋮	⋮	⋮		⋮	⋮	⋮
A_l	x_{l1}	x_{l2}	……	x_{lr}	$T_{l\cdot}$	$\bar{x}_{l\cdot}$

表 1　1 因子要因実験による実験データ

1 元配置実験の分散分析　1 因子 A のみを取り上げて，その最適条件を求めるため l 水準 A_1, A_2, …, A_l を選び，各水準で繰り返し r の計 lr 回の実験を無作為に行うと，実験データは表 1 のように得られる。このとき，データ x_{ij} の構造模型は，

$$x_{ij}=\mu+\alpha_i+\varepsilon_{ij}$$

ここで，α_i：A_i 水準の効果，ε_{ij}：平均 0，分散 σ^2 の正規分布に従う誤差となる。

全データの変動は，総平方和 S_T

$$S_T=(x_{11}-\bar{x}_{\cdot\cdot})^2+\cdots+(x_{lr}-\bar{x}_{\cdot\cdot})^2$$

ここで，

$$\bar{x}_{\cdot\cdot}=(x_{11}+\cdots+x_{lr})/lr$$

によってはかられる。この平方和はさらに，A の水準の違いによるばらつきと，実験誤差によるばらつきによるものに分解することができる。すなわち，

$$S_T=\underbrace{r\{(\bar{x}_{1\cdot}-\bar{x}_{\cdot\cdot})^2+\cdots+(\bar{x}_{l\cdot}-\bar{x}_{\cdot\cdot})^2\}}_{A\text{間平方和}\ (S_A)}$$
$$+\underbrace{\{(x_{11}-\bar{x}_{1\cdot})^2+\cdots+(x_{lr}-\bar{x}_{l\cdot})^2\}}_{\text{誤差平方和}\ (S_E)}$$

となる。このとき，平方和 S_T, S_A, S_E の自由度は，それぞれ $(lr-1)$, $(l-1)$, $l(r-1)$ となり，

$$lr-1=(l-1)+l(r-1)$$

の関係が成り立つ。

これらの平方和を用いて，因子 A の効果があるかどうか，すなわち，

$$\alpha_1=\alpha_2=\cdots=\alpha_l=0$$

であるか否かの検定はつぎのようにできる。まず各平方和をその自由度で除して，不偏分散を求める。

$$V_A=S_A/(l-1),\quad V_E=S_E/l(r-1)$$

さらに A の不偏分散 V_A を誤差分散 V_E と比較するため，不偏分散比

$$F_0=V_A/V_E$$

を求める。もし A の効果が大きければ，S_A は大きくなり，したがって V_A も大きくなるから F_0 の値は大きくなる。したがって F_0 がある値以上になれば，A の効果があると判断できる。この判断基準の値は，有意水準（A の効果がないときに誤って効果があると判断する確率）を α（通常 $\alpha=5$ %がよく用いられる）とすれば，自由度 $(l-1)$, $l(r-1)$ の F 分布の上側 α 点 $F^{l-1}_{l(r-1)}(\alpha)$ により与えられる。

以上のことを表にまとめると，表2の

要因	平方和 (S)	自由度 (ϕ)	不偏分散 (V)	不偏分散比 (F_0)	F 値
A	S_A	$\phi_A=l-1$	$V_A=S_A/\phi_A$	V_A/V_E	$F^{l-1}_{l(r-1)}(\alpha)$
E（誤差）	S_E	$\phi_E=l(r-1)$	$V_E=S_E/\phi_E$		
計	S_T	$\phi_T=lr-1$			

表 2　1 因子要因実験の分散分析表

分散分析表が得られる。

なお平方和の計算は，定義式を用いるよりもつぎのように計算した方がよい。

$$CT=(x_{11}+x_{12}+\cdots+x_{lr})^2/lr$$
$$=(\text{全データの和})^2/lr$$
$$S_T=(x_{11}{}^2+x_{12}{}^2+\cdots+x_{lr}{}^2)-CT$$
$$=(\text{個々のデータの 2 乗和})-CT$$
$$S_A=(T_1.{}^2+T_2.{}^2+\cdots+T_l.{}^2)/r-CT$$
$$=\frac{(A_1\text{ 水準のデータの和})^2}{A_1\text{ 水準のデータの個数}}$$
$$+\frac{(A_2\text{ 水準のデータの和})^2}{A_2\text{ 水準のデータの個数}}+\cdots$$
$$+\frac{(A_l\text{ 水準のデータの和})^2}{A_l\text{ 水準のデータの個数}}$$
$$-CT$$
$$S_E=S_T-S_A$$

ここで，CT は平方和の計算にはつねにでてくる項で修正項とよばれる。

■2元配置実験の分散分析　ここでは，2つの因子 A (l 水準) B (m 水準) をとりあげ，各水準組み合わせの繰り返し r の計 lmr 回の実験を無作為化して行った実験データの分散分析について考える。この場合の実験データは，表3のように x_{ijk} で表すことができる。ただし，添字 i は因子 A の水準を，j は因子 B の水準を，k は繰り返しの番号を示す。また，記号 $T_{i..}$ は A_i 水準のデータの和，$\bar{x}_{i..}$ は A_i 水準のデータの平均値を表す。他の記号も同様に定義する。

データ x_{ijk} の構造模型は，

$A \backslash B$	B_1	B_2	……	B_m	和	平均
A_1	x_{111} ⋮ x_{11r}	x_{121} ⋮ x_{12r}		x_{1m1} ⋮ x_{1mr}	$T_{1..}$	$\bar{x}_{1..}$
A_2	x_{211} ⋮ x_{21r}	x_{221} ⋮ x_{22r}		x_{2m1} ⋮ x_{2mr}	$T_{2..}$	$\bar{x}_{2..}$
⋮						
A_l	x_{l11} ⋮ x_{l1r}	x_{l21} ⋮ x_{l2r}		x_{lm1} ⋮ x_{lmr}	$T_{l..}$	$\bar{x}_{l..}$
和	$T_{.1.}$	$T_{.2.}$		$T_{.m.}$	$T_{...}$	
平均	$\bar{x}_{.1.}$	$\bar{x}_{.2.}$		$\bar{x}_{.m.}$		$\bar{x}_{...}$

表3　2因子要因実験のデータ

$$x_{ijk}=\mu_{ij}+\varepsilon_{ijk}$$
$$=(A_iB_j\text{ 条件下での真値})+(\text{誤差})$$

と表されるが，μ_{ij} はさらに

$$\mu_{ij}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}$$
$$=(\text{一般平均})+(A_i\text{ 水準の効果})$$
$$+(B_j\text{ 水準の効果})$$
$$+(A_iB_j\text{ の交互作用})$$

と書けるから，結局

$$x_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk}$$

となる。ただし，

$$\alpha_1+\beta_2+\cdots+\alpha_l=0$$
$$\beta_1+\beta_2+\cdots+\beta_m=0$$
$$\left.\begin{array}{c}(\alpha\beta)_{11}+\cdots+(\alpha\beta)_{1m}=0\\ \vdots\\ (\alpha\beta)_{1m}+\cdots+(\alpha\beta)_{lm}=0\end{array}\right\}(l+m)\text{個}$$

である。交互作用は，A，B どちらか一方の添字について和をとれば0にな

$A \backslash B$	B_1	B_2	$\cdots$	B_m
A_1	$T_{11\cdot}=x_{111}+\cdots+x_{11r}$	$T_{12\cdot}=x_{121}+\cdots+x_{12r}$		$T_{1m\cdot}=x_{1m1}+\cdots+x_{1mr}$
A_2	$T_{21\cdot}=x_{211}+\cdots+x_{21r}$	$T_{22\cdot}=x_{221}+\cdots+x_{22r}$		$T_{2m\cdot}=x_{2m1}+\cdots+x_{2mr}$
$\vdots$				
A_l	$T_{l1\cdot}=x_{l11}+\cdots+x_{l1r}$	$T_{l2\cdot}=x_{l21}+\cdots+x_{l2r}$		$T_{lm\cdot}=x_{lm1}+\cdots+x_{lmr}$

表 4　AB2元表

要因	平方和	自由度	不偏分散	不偏分散比
A	S_A	$\phi_A=l-1$	$V_A=S_A/\phi_A$	V_A/V_E
B	S_B	$\phi_B=m-1$	$V_B=S_B/\phi_B$	V_B/V_E
$A\times B$	$S_{A\times B}$	$\phi_{A\times B}=(l-1)\times(m-1)$	$V_{A\times B}=S_{A\times B}/\phi_{A\times B}$	$V_{A\times B}/V_E$
E	S_E	$\phi_E=lm(r-1)$	$V_E=S_E/\phi_E$	
計	S_T	$\phi_T=lmr-1$		

表 5　2因子要因実験の分散分析表

る。また誤差 ε_{ijk} は平均0，分散 σ^2 の正規分布に従うものとする。この場合の仮説検定は，

(1)　Aの主効果は存在するかどうか。

(2)　Bの主効果は存在するかどうか。

(3)　AとBの交互作用 $A\times B$ は存在するかどうか。

の3つになる。

Aの主効果，Bの主効果，$A\times B$の交互作用を検定するために，1因子要因実験の場合と同様に総平方和 S_T を

$$S_T=(x_{111}-\bar{x}_{\dots})^2+(x_{112}-\bar{x}_{\dots})^2+\cdots+(x_{lmr}-\bar{x}_{\dots})^2$$
$$=S_{AB}(AB\text{間平方和})+S_E(\text{誤差平方和})$$

に分解することが必要になる。S_{AB}を計算するには，表4のAB 2元表を作成すると便利である。表4より，S_{AB}は，

$$S_{AB}=(T_{11\cdot}{}^2+\cdots+T_{lm\cdot}{}^2)/r-CT$$

と求められる。ここで，CT は修正項で，

$$CT=T_{\dots}{}^2/lmr$$
$$=(\text{全データの和})^2/lmr$$

である。S_{AB} はさらに，

$$S_{AB}=S_A+S_B+S_{A\times B}$$

に分解することができる。ここで，S_AはAの水準間の効果の違いを表す平方和，S_BはBの水準間の違いを表す平方和，$S_{A\times B}$は交互作用 $A\times B$ の大きさを表す平方和である。S_A, S_B は，

$$S_A=T_{1\cdot\cdot}{}^2/mr+T_{2\cdot\cdot}{}^2/mr+\cdots+T_{l\cdot\cdot}{}^2/mr-CT$$
$$S_B=T_{\cdot1\cdot}{}^2/lr+T_{\cdot2\cdot}{}^2/lr+\cdots+T_{\cdot m\cdot}{}^2/lr-CT$$

と，そして $S_{A\times B}$ は，

$$S_{A\times B}=S_{AB}-(S_A+S_B)$$

と求められる。なお，誤差平方和 S_E は，

$$S_E=S_T-S_{AB}$$

から求められる。

以上のことを分散分析表の形にまとめると，表5のようになる。これより，有意水準 α で

$V_A/V_E>F^{\phi_A}_{\phi_E}(\alpha)\Rightarrow A$は有意

$V_B/V_E>F^{\phi_B}_{\phi_E}(\alpha)\Rightarrow B$は有意

$V_{A\times B}/V_E>F^{\phi_{A\times B}}_{\phi_E}(\alpha)\Rightarrow$ $A\times B$は有意

と判断する。最適条件を求めるには推定の考え方が必要。（宮村）

分析に対する行動科学的知見

ORの問題を数理的技術の問題としてではなく，組織・人間の問題として措定し，多くの行動科学者，組織科学者がこの議論に参加した。分析に対する行動科学的知見は①組織内の二大文化圏，及び②OR活動推進のための行動科学的技法という形で論ずることができる。

★解説

行動科学的研究によって得られた分析活動に対する知見にはいろいろあるが，これを整理すると以下のようになる。

意思決定者ないしは管理者の世界と分析者(ないしはOR研究者)の世界は，さまざまな観察・調査・実験などから大変違っているといわざるをえない。一言でいえば，組織内に2つの異なった文化圏が存在し，これら文化圏の相異に起因し，両者が共同作業する際の足並みがみだれたり，誤解が生じたりすることになる。管理者と分析者の違いについては，①知覚の仕方，②パーソナリティ，③認知の仕方，④問題解決スタイル，⑤管理者と分析者の仕事の性質，中身，ペース(仕事環境)，⑥政治や権力に対する考え方，⑦動機付けや目標，⑧バックグラウンドなど，さまざまな側面から指摘されている。概してこのような違いを互いが気づかず，管理者も分析者も同じ人間であるということで，2つの文化圏の間に軋轢を引き起こしているとされる。

またこうした行動科学的研究の中から，技法とはいえないが，一応の行動科学的ノウハウとしては，①コミュニケーション技法，②参加的(集団的)問題解決の奨め，③管理者と分析者の仲介(2つの文化圏のコーディネーター)を主業務とする第三者(integrator)の提案などがあげられる。

関連ページ
ORの実施理論
44

■組織内の二大文化圏

(1) **知覚の差異** 管理者(以下Mとする)が，分析者(以下Rとする)やOR手法の利用に対してどのような知覚を有しているかに関して，MはORのような経営科学的技術によって自分の役割がとって代わられるのではないかとか，自己の地位を維持するために分析結果に抵抗しているとチャーチマンは述べ，またアージリス(C. Argyris)は，Mは経営科学技術の浸透によって自己の自由裁量の幅やessentiality(この仕事は自分でなければこなせないといった感覚)が減少するのではないかと恐れていると報告している。また同一のデータに対しMとRはまったく違った知覚をもっているという報告もある。ヘイズ(R. H. Hayes)とノラン(R. L. Nolan)によると，MはRの行う分析手法に対する約束事(promises)が過大過ぎると知覚していると述べており，またMは自分の感じているイメージにぴったりしたモデルをよいモデルと考えていて，現実そのものを正確にとらえることではないと報告している。多くの点でMとRは知覚のズレが生じている。

(2) **パーソナリティの問題** チャーチマンは，分析結果の提案に対するMの反応の仕方は人によって大変異なっており，かりに提案を拒否する場合でも全面的に否定するケースは少なく，言い方，接し方でその抵抗はゆるめられもするとし，RはMのパーソナリティをよく理解すべきだとしている。Mは現時点の，特定の，その場その場(ad hoc)の，口頭形式で語られた問題を好む傾向がある(ミンツバーグ，H. Mintzberg)。またアージリスによるとMは他者によって自己の目標水準や成功基準に口をはさまれたとき，心理的敗北感(psychological failure)を受け，フラストレーションを感じると報告している。

(3) **認知スタイルの相異** マッケニー(J. L. Mckenney)とキーン(Keen)は，Mは直観的・教訓的(関係を気にする)思考をし，Rはシステマティックな思考で，詳細を気にするタイプが多いとしており，ハモンド(J. S. Hammond)は，Mはヒューリスティックで，Rは大変分析的であるとしている。

(4) **問題解決スタイルの相異** 認知スタイルの差異は問題解決様式にも大きな差異をもたらす。ハモンドによると，MはRよりルーズで，明示的ではなく，しばしば広いスコープで問題をとらえており，マッケニーらによると，Mは直観的問題解決を行う傾向が高く，確率を0か1でしか考えず，考える前に行動する人が多いとしている。この思考様式の違いについては，最近の大脳生理学の発達から左脳が直線的・論理的思考を司り，言語や数学は左脳の発達と関係し，右脳が全体的・同時的・空間的思考を司るとされてい

る(これについては発達中のこともあり，問題はそれほど単純ではないとされてる)。左脳タイプの人は，良構造問題(例：解答のある数学の問題)を好んで選択し，verbal-analyticテストを高得点し，このタイプはRに多く，右脳タイプの人は，open-endedな問題(あいまいな自由解答問題，条件不足問題など)を好んで選択し，intuitive-spatialテストを高得点し，Mはこのタイプが多いという実験報告もある。

(5) **仕事内容と仕事環境の相異**　Mはinterpersonal roles(看板として，外部との接触役割，リーダーとして)，情報処理役割，意思決定機能など多様な役割をこなさねばならず，仕事のペースが残酷(不連続的，多様さ，じっくりと思考する時間がない。絶えず外乱が発生し，人と話をしている)であり，Rと仕事環境が大変違っている。また人により内容も異なる。Rは持続して仕事に集中できる。

(6) **権力と政治への無関心**　多くのRは，組織における権力と政治の役割を理解していないとよくいわれている。スター(M. Starr)は，Rはもっと政治に対する敏感性をもつべきだとしている(政治的誤謬)。またRはもっと現存する意思決定システムのメカニズムの分析(本当に決めているのは誰かとか，どこで情報が発生したり消滅したりしているかなど)をすべきであるとされる(エイコフ，R. L. Ackoff)。一般に，Rは手法(数理，コンピュータ・プログラムなど)にのみ関心を持ち過ぎ，政治に疎く，無関心であることが多い(フェンスケ，R. W. Fenske)。

(7) **目標(ないしは動機)の相異**　Mは実用性，即時性を重んじ，**特定目的，特定問題**に関心(problem in content)があり，Rは問題を**一般的な形式**(problem in form)でとらえ，モデルの質や**格好良さ，エレガントさ**など専門性の追求を動機とする。目標達成度の理解についても，Mは"acceptable"な解(70％ならOKとする)を，Rは"right" answer(100％を狙う)を目指すという報告もある。Mが組織全体の目標よりは当面の自部門(ないしは自己)の目標に動機付けられるのに対し，Rはいわばコスモポリタン的で，組織全体目標を目指すとされている。

(8) **受けた訓練(バック・グラウンド)の違い**　Mは経験をベースにしているのに対し，Rはフォーマルなトレーニングを積んで，定量的手法をある程度マスターした人が多い。

(9) **年齢の相異**　Mは一般に年輩者でRは比較的若者が多い。

以上，行動科学的な側面から調査・研究された諸知見を述べてきたが，RとMの仕事環境や仕事内容，問題解決スタイル，知覚や動機付けの仕方などが大変違っているということを，十分にRは気づいた上で，意思決定メカニズムの分析やモデル構築・解析，解の提言，実施な

どをすることが必要とされている。

■ **OR活動推進のための行動科学的技法**

(1) **伝達技術1(聞き上手)** M(管理者)からR(分析者)への伝達であり，OR活動の**初期段階**に特に大切で，初期段階は**ゆっくり時間を割き**，問題の定義の際Mの観点を十分に取り入れることである。M自身が何をやりたいかを明確にとらえていないことも多く，それを引き出すには**非指示的インタビュー技術**が必要で，分析の進行中もMとの間に**継続的対話**を確保することも必要である。大事な情報(これはあまり手放したがらないとも言われている)がRに教えられないことが起こりうる。

(2) **伝達技術2(話し上手)** R側からM側への伝達技術で，いわば発表・報告技術である。部厚い，難解な，数式を盛り込んだ，あるいは単なる文書のみの報告等はいましめられ，平易な，Mの言葉で，図表を用い，口答で何回でも報告し，しかも職務階層よりも**最も効果的な人に報告**することが大事であるとされる。

(3) **人間関係の技術** 個人間・集団間に発生する個人間葛藤(interpersonal conflict；感情的，情動的歪みなど)の処理技術もRに必要だとされる。

(4) **参加的マネジメント(participative management)の必要性** ORグループ単独のプロジェクト推進はいたるところでいましめられ，**集団的問題解決**(ORグループの人，M，トップ，影響を受ける関連部門)の必要性が強調される。集団的問題解決のもつ意味は，Mはモデル構築への参画によって**総合的体験**が得られる点にあり，また，モデルを用いることよりも，**作ること**の方に価値があるとされる。さらによいモデルとは，それが意思決定に役立つことではなく，**教育・開発の手段**を与えることにあるともいわれている。OR活動の最初からMを取り込め(management involvement)ば，お互いが恥をかかずにすむ。OR研究費を**関係部門に負担**させ，責任者もそこのMにすると，プロジェクトに熱心になるともいわれている。またMの定期移動などで，プロジェクトが挫折することも多く，次席支配力をもつ人を含ませることも大事である。

(5) **コンティンジェンシー・アプローチ** それぞれの問題状況は異なるため，**唯一最善の方法(ユニバーサルな方法)**はないので，Rは**その**会社，**その**問題の条件をよく考えて臨機応変に対応する必要がある。

(6) **ジョブ・ローテーション** Rの人間とラインの人間の間にローテーションを行い，RにMの体験をさせるなども指摘されている。

(7) **第三者(integrator)の導入** MとRの間に立って，Rの専門性を向上させつつ，Mのかかえる実問題に答えていくために，2つの文化圏のバッファー役としての仲介を専門とする第三者の導入も提言されている。 (中野)

平均値

もっとも基本的な特性値である代表値として，いちばんよく用いられるものは平均値である。平均値を単に平均ともいう。平均は，データの合計の値をデータの個数で割ることによって求められる。

★解説

関連ページ
確率変数の平均値（期待値） 58

n 個のデータ x_1, x_2, …, x_n があるとき，平均値 $\overline{x}$ (エックス・バーと読む)を求めるには，

$$\overline{x}=\frac{1}{n}(x_1+x_2+\cdots+x_n)$$

を計算すればよい。平均値 $\overline{x}$ は，次のような性質をもっている。いま，ある数 h とデータとの差の平方和

$$y=(x_1-h)^2+(x_2-h)^2+\cdots+(x_n-h)^2$$

を考え，これを最小にする h を求めてみると，$h=\overline{x}$ となることがわかる。

平均値を求めるには，いろいろの仕方がある。そのいくつかを右ページに示してあるが，データが，x_1 という値をもつもの f_1 個，x_2 が f_2 個，…，x_k が f_k 個で，合計 $f_1+f_2+\cdots+f_k=n$ (個)から成り立っているとき，平均値 $\overline{x}$ は，

$$\overline{x}=x_1\cdot\frac{f_1}{n}+x_2\cdot\frac{f_2}{n}+\cdots+x_k\cdot\frac{f_k}{n}$$

で求めることもできる。上の式で，f_1/n, f_2/n, …, f_k/n は**相対度数**とよばれる。これらをそれぞれ p_1, p_2, …, p_k と書くことにすると，平均値 $\overline{x}$ は，

$$\overline{x}=x_1\cdot p_1+x_2\cdot p_2+\cdots+x_k\cdot p_k$$

によって求めることができる。つまり平均値 $\overline{x}$ は，データの値にその相対度数をかけて加え合わせることによって求められる。

例1　仮の平均を用いた計算

データ;

15.3, 15.3, 15.0, 15.4,
15.2, 15.7, 15.1, 15.3,
15.3, 15.2 (単位：cm)

のとき，仮の平均を15.0(cm)として，

$$\text{平均値}\,\bar{x}=15+\tfrac{1}{10}\times\{(15.3-15)+(15.3-15)+\cdots+(15.0-15)\}$$
$$=15+0.28$$
$$=15.28\,(\text{cm})$$

例2　相対度数の和として平均が求められる例

得点 x_i	度数 f_i	相対度数 p_i
0点	4人	0.08
1	10	0.20
2	18	0.36
3	12	0.24
4	4	0.08
5	2	0.04
計	50	1.00

q_0　q_1　q_2　q_3　q_4

0.92+0.72+0.36+0.12+0.04
=2.16

■平均値の求め方　例1は同種の製品10個の寸法を測って得た値の平均を求めるのに，仮の平均を用いて計算した例である。

一般に n 個のデータ

$$x_1,\ x_2,\ \cdots,\ x_n$$

があるとき，大まかに平均値に近い数 a を定めておいて，

$$\bar{x}=a+\frac{1}{n}\{(x_1-a)+(x_2-a)+\cdots+(x_n-a)\}$$

のように，平均値 $\bar{x}$ を求めることができる。このような a のことを**仮の平均**という。

上の例では，仮の平均を15.0として計算してあるが，仮の平均はいくらにしておいてもよい。たとえば，仮の平均を16.0とすると，

$$\text{平均値}\ \bar{x}=16+\frac{1}{10}\{(15.3-16)+(15.3-16)+\cdots+(15.0-16)\}$$
$$=16-0.72=15.28\ (\text{cm})$$

となる。

例2は，50人の生徒に対するテスト(5点満点)の結果が，度数表にまとめられているとき，相対度数を用いて，平均値を求める方法を示している。

$$x_0=0,\ x_1=1,\ \cdots,\ x_5=5$$

に対する相対度数は，

$$p_0=0.08,\ p_1=0.20,\ \cdots,\ p_5=0.04$$

である。いま，

$$q_i=p_{i+1}+p_{i+2}+\cdots+p_5\ (i\geqq 0)$$

とおくと，$\bar{x}=q_0+q_1+\cdots+q_4$ で平均 $\bar{x}$ が求められる。　(牧野)

ベイズの定理を用いた判別(ベイズ診断)

ベイズの定理に基づき，事後確率を用いて判別分析を行う手法。質的データの判別手法が未発達な時代において，利用できた代表的な判別手法である。人間の意思決定と相性がよく，欧米の医療関係者の間では今でも利用されている。一般にベイズ診断とよばれている。

ベイズ診断は，データがない場合でも，専門家の主観や経験に基づく確率を用いることができ便利である。しかし，信頼性の高いデータがある場合には，判別関数などの多変量解析手法の方が一般に判別成績がよい。

★解説

■ベイズ診断　n 個の群 G_k $(k=1, \cdots, n)$ を，質的な説明変数の値(カテゴリー) S を用いて判別する。説明変数が複数の場合，カテゴリーの組み合せのパターンをあらためて S として考える。連続変数の場合には，値をいくつかの点で分割し，カテゴリー化して用いる。

S が観測されたもとでの G_i の事後確率 $P(G_i|S)$ は，事前確率 $P(G_k)$ と尤度 $P(S|G_k)$ が与えられたとき，ベイズの定理を用いて次のように表される。

$$P(G_i|S)=\frac{P(G_i)P(S|G_i)}{\sum_{k=1}^{n}P(G_k)P(S|G_k)} \tag{1}$$

ここで n 個の事後確率のうち l 番目の $P(G_l|S)$ が最大値をとり，他の $(n-1)$ 個の値に比べ十分に大きければ，観測値 S をとるデータを群 G_l に判別すればよい。

データがある場合には，説明変数と外的基準の 2 重クロス表から，$P(G_k)$ と $P(S|G_k)$ の値を推定できる。前者は G_k の周辺確率になり，後者は S と G_k の同時確率になる。

関連ページ

クロス集計　90

FUNCAT による判別　320

数量化Ⅱ類　190

決定行列　104

ROC 曲線　14

尤度比方式による判別　390

AID　32

■例題　がん患者10人と対症群(コントロール群)10人に対して，新しく開発された検査を実施し，表1の結果を得た。

	がん	正常	
+	8人	3人	11人
−	2人	7人	9人
	10人	10人	

	がん	正常	
+	8/20	3/20	11/20
−	2/20	7/20	9/20
	10/20	10/20	

表1　がんの検査データ

上の左の表は，がん症例中の8人が検査＋，2人が検査−になったことを示す。この20人の検査結果から経験確率を計算したものが上の右の表である。表中の確率を同時確率という。たとえば，がんで検査＋の同時確率は P(がん$\cap$+)=0.4 である。表側と表下の確率は，検査または疾病単独の周辺確率である。周辺確率は各行各列の合計になっている。

がんと正常の周辺確率 P(がん)=P(正常)=1/2 を特に事前確率という。一方，がん症例で検査＋または−，または正常症例で検査＋または−の条件付確率を尤度とよび，次の値をとる。

P(+|がん)=0.8

P(−|がん)=0.2

P(+|正常)=0.3

P(−|正常)=0.7

疾病に関する周辺確率と条件付確率を，事前確率と尤度というように言い換えるのは，疾病が原因となる事象で，検査がそれに伴って起こる事象を表すからである。

次にがんの疑いをもって来院した患者にこの検査を実施したところ，＋の結果を得た。果たして，この患者はがんであろうか。この疑問に答えるため，次の条件付確率をベイズの定理(1)を用いて計算する。

$$P(\text{がん}\mid +)$$

$$=\frac{(10/20)(8/10)}{(10/20)(8/10)+(10/20)(3/10)}$$

$$=\frac{80}{80+30}=\frac{8}{11}$$

$$P(\text{正常}\mid +)=\frac{3}{11}$$

以上から，検査を実施する前は，この患者ががんであるか正常であるかは，事前確率が1/2であるので判断できなかった。検査実施後は＋という情報が得られたので，がんである確率が8/11と大きくなり，この症例をがんと判別することになる。この条件付確率を，事前確率に対応させて，事後確率とよぶ。すなわち，条件付確率は，尤度と事後確率というように使い分けられる。

以上の過程は次のようにまとめられる。事前確率は，がんも正常も0.5である。検査＋という情報は，P(がん|+)+P(正常|+)=11/20 の確率を1に変化させる。これにより，検査＋が確定後の事後確率は，がんの方が正常より大きくなる。　（新村）

変動係数

2つの種類のデータの散らばり具合を比べるとき，単に分散や標準偏差で比べないで，それぞれの(標準偏差)/(平均)で比べたほうがよい。このような量を，変動係数(または変異係数)という。

★解説

10人ずつ，2つのグループA班，B班に分かれて，異なった区間の測量を行ったとしよう。A班の人たちによる測定結果のバラツキを計算したところ，**標準偏差** $s_1=0.1\mathrm{m}$ であった。一方，B班の人たちによる標準偏差は $s_2=0.06\mathrm{m}$ であった。このことだけで，B班のほうがバラツキが少なかったと判断するのは適当でない。A班による測定区間の**平均**の長さは $\bar{x}=1000\mathrm{m}$ であったのに対し，B班の平均距離は $\bar{x}_2=200\mathrm{m}$ でしかなかったとすると，むしろA班のほうが，よい測定結果を得ているといったほうがよいと思われる。いいかえれば，標準偏差は平均値のまわりの散らばりの大きさを示す量であるから，2つの異なる種類のデータを比較するのには適していない。そのような場合，単に標準偏差の大小で比べるよりも，

(標準偏差)/(平均値)

で比べたほうがよい。このような量を**変動係数**(coefficient of variation)という。変動係数をc.v.と略記することがある。

上の例では，

$$\text{A班のc.v.} \quad =\frac{s_1}{\bar{x}_1}=\frac{0.1}{1000}=0.0001$$

であるのに対し，

$$\text{B班のc.v.} \quad =\frac{s_2}{\bar{x}_2}=\frac{0.06}{200}=0.0003$$

であるから，平均値に対する相対的な散らばりの大きさを表す変動係数で比べると，A班のほうが精度のよい結果を得ているということができる。なお，変動係数はふつうデータが正の値のものから成り立っているときに用いる。

関連ページ

標準偏差　59, 333

平均　59, 344

順位	社　　名	所得金額 (千万円)
1	日本銀行	3547
2	トヨタ自動車	3336
3	日産自動車	1880
4	アラビア石油	1658
5	日立製作所	1618

表1　全企業内での上位5社

順　位	社　　名	所得金額 (千万円)
1.(2264)	三木プーリ	104
2.(2736)	日本開閉器工業	85
3.(3607)	矢野新商事	64
4.(4199)	利康商事	55
5.(4420)	六桜商事	52

表2　中小企業内での上位5社
(カッコ内は全企業内での順位)

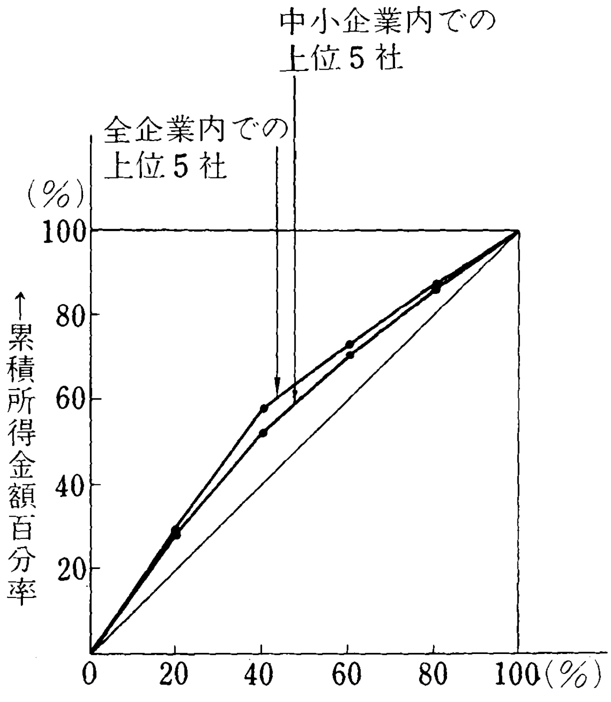


図1　パレート図

■法人の所得金額ランキング　中小企業家同友会という組織があって，そこで発行している機関紙によると，「法人所得50,000社ランキング83年版」(日経マグロウヒル社)に記載されている会員企業上位5社は，表2のようであるという。中小企業ということについての定義はきちんとしておかなければならないが，ここでは上に記した組織の会員ということにしておく。

一方，全企業内での法人所得上位5社は，表1のようであった。両者におけるバラツキ，いいかえれば集中の度合いを調べるために変動係数を計算してみよう。

まず，全企業内での上位5社については，

平均 $\overline{x}_1=2407.8$ (千万円)

標準偏差 $s_1=851.3$ (千万円)

となるから，変動係数は，

$$\text{c.v.} =s_1/\overline{x}_1=0.354$$

である。これに対し，中小企業内での上位5社については，

平均 $\overline{x}_2=72.0$ (千万円)

標準偏差 $s_2=19.7$ (千万円)

であって，変動係数は

$$\text{c.v.}\ s_2/\overline{x}_2=0.274$$

となる。つまり，全企業内での上位5社のほうが，変動係数が大きい。これはいいかえれば，全企業のほうが，集中の度合いが強いということであり，パレート図(→300ページ)が，このことをよく表している。　(牧野)

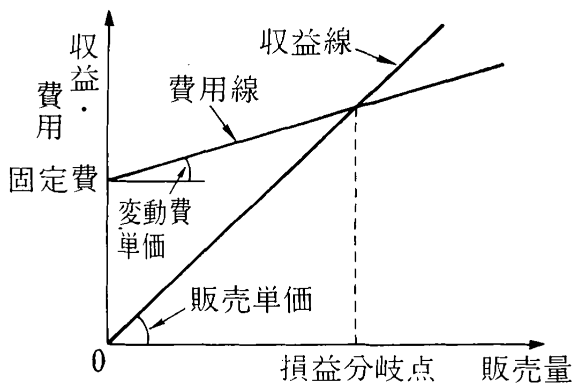
変動費と固定費による損益分岐点分析

操業度(生産量，販売量)に応じて変化する費用を変動費(variable cost)といい，操業度が変動しても総額として変わらない費用を固定費(fixed cost)という。製品の製造販売を1単位増加させると，固定費は変わらないが変動費が増し，収益も販売単価ぶん増加し，(販売単価－変動費単価)だけ利益が増大する。この利益額のことを限界利益(marginal profit)という。

固定費を限界利益で割った値を損益分岐点(break-even-point)という。損益分岐点は，利益がゼロになる操業度を意味する。

★解説

費用の内容を変動費と固定費でとらえ，操業度に応じて利益が変化する様子をグラフで示したのがいわゆる利益図表(profic chart)である。



p：販売単価

v：変動費単価

F：固定費

D：販売量（操業度）

利益 $\pi=(p-v)\times D-F$

損益分岐点 $D_0=\dfrac{F}{p-v}$

図1　利益図表

利益図表は利益計画を立案するのに便利な図表として使われているが，経済性の分析でも活用できる。特に，販売量(需要量)によって代替的な生産方策の有利さがどう変わるかを，利益図表に表してみるとよい。変動費，固定費，販売量が，方策の有利さを左右する様子がわかりやすくなる。

関連ページ
経済性の比較の原則　98
埋没費用　364

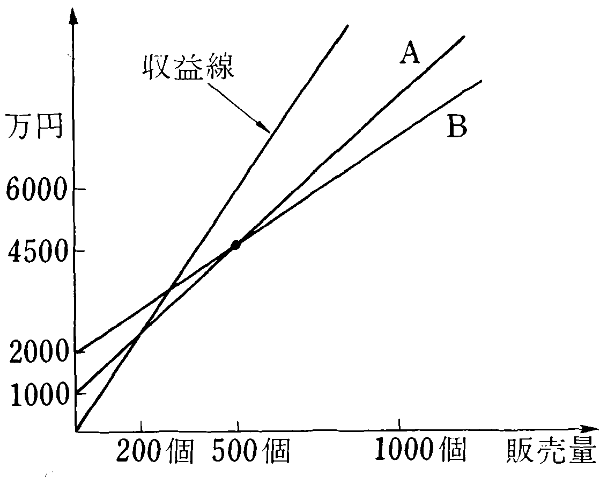
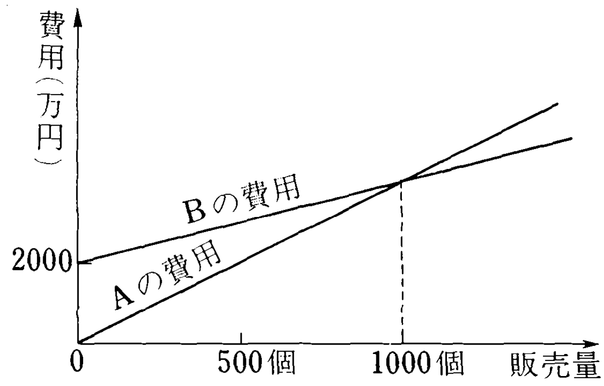


図 2　代替生産方策



$$優劣分岐点=\frac{2000\text{万円}}{(7-5)\text{万円/個}}=1000個$$

図 3　生産方策の優劣分岐点

■利益図表での代替生産方策の比較

販売単価が12万円の製品を生産するのに，次のような代替案A, Bがある。

案	固定費	変動費単価
A	1000万円	7万円/個
B	2000万円	5万円/個

すなわち　両者を比べると，A案は変動費にたより，B案は固定費にたよる案といえる。この2つの案の費用線を利益図表に描くと図2のようになる。

図2からもわかるように，販売量が500個を境にして，A案とB案の有利さが変わってくる。この500個の販売量を優劣分岐点という。また，A案の損益分岐点は200個である。このように，販売量の変化に応じて案の有利さの変わる様子が読みとれる。

■方策の優劣分岐点　前記の例で，A案とB案の優劣分岐点は500個になった。この値は次式から求められる。

$$優劣分岐点=\frac{固定費の差}{変動費単価の差}$$

$$=\frac{(2000-1000)\text{万円}}{(7-5)\text{万円/個}}=500個$$

各案の固定費が設備費だとしよう。そして設備Aはすでに取得して工場にあり，設備Bはこれから購入するものだとする。この場合は，設備Aの固定費1000万円は，設備Aを使ってもBを使ってもいずれにしても共通にかかる費用となる。したがって，AとBの比較には考慮しないでよいことになる。その結果，AとBの優劣分岐点は図3で表されることになる。

別の場合として，設備A, Bともに以前に取得したもので工場にある場合は，変動費単価の安い設備Bが販売量のいかんにかかわらず有利になる。固定費も含めて費用を比較してはいけない。

このように設備費の取り扱い方は，設備をこれから入手するのか，すでに入手済みかで変わってくることに注意しなければならない。　　　　(中村)

ポアソン分布

たとえば，放射性物質の崩壊にともなって α 粒子が発生するときのように，事象が時刻に関係なくランダムに発生する場合，一定時間内に発生するその個数の分布はポアソン分布(Poisson distribution)となる。

★解説

「ランダムに発生する」とは次のような意味あいである。

(1)**独立性**　ある時間区間に発生する事象の数は，それ以前の事象の発生の仕方とは無関係(**独立**)である。

(2)**順序付可能**　同時に2個以上発生する確率は無視できる。

(3)**定常性**　同じ長さの区間であれば，その間に発生する事象の個数の分布は，観測開始時点によらない。

ある時間区間に，このような仕方で発生する事象の個数を X とすると，その確率分布と特性値は，

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda} \qquad (k=0,\ 1,\ 2,\ \cdots)$$

平均　$E(X)=\lambda$　　　分散　$\mathrm{Var}(X)=\lambda$

となる。この分布をパラメータ λ のポアソン分布とよぶ。

また，これは次のようにも導ける。ある区間を n 等分する。各小区間の間に事象の発生する確率を p，発生しない確率を $1-p$ とし，各区間に発生するか否かは独立とする。全区間で k 個発生する確率は，n 回コインを投げて k 回表のでる確率と同じと考えてよく，**二項分布**となる。全区間での発生個数の平均は np であるが，これを一定値 λ として，n を大きくすると(つまり，$p=\lambda/n$ は小さくなる)，ポアソン分布に近づく(図1)。フランスのポアソン(S. D. Poisson, 1781－1840)が導いたのでこの名がついた。このような事象の発生の仕方を一般化して述べたのが，上の(1)～(3)である。

関連ページ

平均　58
分散　58
二項分布　276

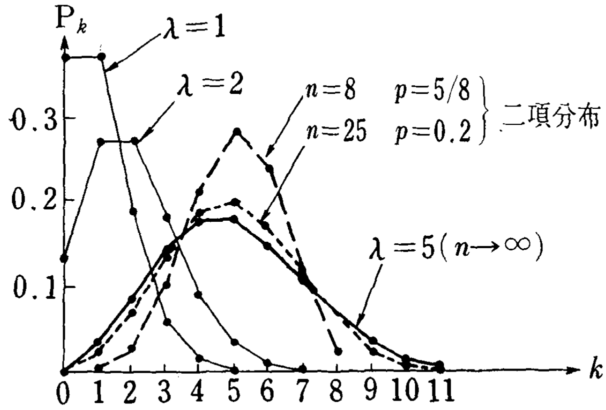


図 1　ポアソン分布(実線)と二項分布(点線)

台　数	観 測 値	理 論 値
0	180	173
1	112	121
2	41	42
3	12	10
4	1	2
5	2	0
6 以上	0	0
計	348	348

表 1　ゲートへの車の到着数

■有料道路ゲートへの車の到着　表1は森村らが観測した15秒ごとの観測値と理論値である(〔56〕参照)。理論値は観測値から平均を求め，それを平均とするポアソン分布から計算したものである。つまり，

$$\text{平均}=\frac{1\times112+2\times41+3\times12+4\times1+5\times2}{348}$$

$$=0.70$$

であるから，$\lambda=0.7$ とおく。たとえば，15秒間に3台到着する確率は，

$$P\{X=3\}=\frac{0.7^3}{1\times2\times3}e^{-0.7}=0.028$$

となる。これに総回数348をかけると，3台の到着に対する理論値 $9.74\fallingdotseq10$ が求まる。$\lambda=0.7$ のポアソン分布によるあてはめが妥当か否かは，適合度検定(→248ページ)を行えばよいが，見たかぎりではかなりよい。

このほか，ポアソン分布は，ある電話局の交換機に接続を要求してくる呼の数の分布などの社会現象にもよくあてはまることが多い。その理由は，1人1人としては定まった時間区間内に電話をする可能性はきわめて小さいけれど，そのような潜在客が数多くいてそれぞれ勝手に行動するので，その呼の発生の仕方は(1)～(3)の性質を大体満たすためであろう。

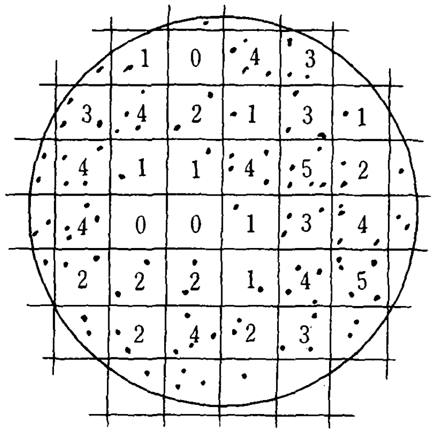


図 2　バクテリアのコロニーの発生

■バクテリアのコロニーの発生(〔44〕参照)　図2はペトリ皿上に発生したバクテリアのコロニーを顕微鏡で眺めたものである。これまで説明の都合上，時間軸上でランダムに発生する事象についてのみ説明してきたが，空間上に発生する点についても(1)～(3)と同様な性質があれば，各メッシュ内の点の個数の分布はポアソン分布となる。都市を適当な大きさのメッシュに切ったとき，各メッシュ内に含まれる病院の数などもポアソン分布となることが知られている。　　(森)

貿易統計

貿易による商品の流れや，それに伴うカネの決済，輸出入による海上運輸，海外旅行，海外投資，海外諸国のわが国に対する借款など国際経済取引の実態を明らかにした統計を貿易統計という。貿易統計は主として関税収入を得るための必要性から生まれたものである。一応の整理がなされた貿易統計は，フランスでは17世紀後半に，ついで18世紀末までに，イギリス，プロセイン，アメリカの諸国ではじめられ，わが国では，関税法が施行された1899年に本格的な貿易統計が作成された。

★解説

貿易業務をその契約からカネの決済までの流れを輸出についてみると，①輸出契約の成立，②信用状の接受，③為替銀行の認証，④貨物の輸出，⑤為替銀行によるカネの決済，の各段階を経て完了する。その各段階ごとに統計がとられている。このうち，税関の手続きが終わった貨物の金額，数量を集計した通関統計と，貿易のみならず，あらゆる国際経済取引における金銭の受け払いの関係を総体的にとらえた国際収支統計が最も基本的な統計である。

通関統計(日本外国貿易統計)は，輸出入の金額と数量を，相手国別，商品分類別に集計しており，この統計結果から，貿易数量指数，貿易価格指数が作成されている。貿易のどの段階で通関統計に計上するかについては，輸出がFOB価格(Free On Board)，輸入がCIF価格(Cost Insurance and Freight)で統一されている。

国際収支統計は，「ある一定期間内において，主にわが国での居住者と外国の居住者との間で行われた経済取引を体系的に記録したもの」で，その基本的形態には，①対価を伴う商品とサービスの取引，②物資・サービスの交換(バーター取引)，③資本の取引(直接投資，証券投資など)，④物資・サービスの一方的供与，⑤金銭などの一方的供与，の5つがある。

関連ページ

指数　154
主な経済指数　425

(1) 輸出　(単位　兆円)

商　品	昭和55年		58年	
	金額	構成比(%)	金額	構成比(%)
総　　額	29.4	100.0	34.9	100.0
金属及び同製品	4.8	16.5	4.4	12.5
(うち鉄鋼)	(3.5)	(11.9)	(3.1)	(8.7)
機械機器	18.4	62.7	23.7	67.8
(うち自動車)	(5.3)	(17.9)	(6.2)	(17.8)
(うち船舶)	(1.1)	(3.6)	(1.4)	(4.1)
(うちテープレコーダー)	(0.7)	(2.5)	(1.6)	(4.5)
そ　の　他	6.1	20.8	6.9	19.8

(2) 輸入

商　品	昭和55年		58年	
	金額	構成比(%)	金額	構成比(%)
総　　額	32.0	100.0	30.0	100.0
食　料　品	3.3	10.4	3.5	11.8
鉱物性燃料	15.9	49.8	14.0	46.6
(うち原油・粗油)	(12.0)	(37.5)	(9.5)	(31.7)
(うち石油製品)	(1.2)	(3.6)	(1.3)	(4.5)
化学製品	1.4	4.4	1.7	5.7
機械機器	2.2	7.0	2.5	8.2
そ　の　他	9.1	28.4	8.3	27.7

表1　商品別輸出入の状況

通関統計は，商品や国別に細かく分類されているので，貿易構造の変化や地域的な依存度をつかむことができる。

昭和58年の輸出入の構造をみると，輸出では自動車が6.2兆円(全輸出額の17.8％)，鉄鋼3.1兆円(8.7％)，テープレコーダー1.6兆円(4.5％)，船舶1.4兆円(4.1％)となっており，これら4品目で35％を占めている。ここ3年間では鉄鋼以外の3品目は急増している。輸入では，原油や石油製品などの鉱物性燃料が14兆円で全輸入金額のほぼ2分の1となっており，ついで食料品の3.5兆円(11.8％)が多い。つまり日本の貿易は，原料を輸入し，これを加工して輸出するという貿易構造になっている。一般に，景気が上昇すると利益が大きくなった国内販売に力を入れるので，輸出が停滞するのに対し，原料輸入の必要性が高まるので輸入が増える。景気が後退局面に入ると，国内販売が鈍化するため輸出ドライブがかかるとともに輸入が減退する。輸出入金額はこのほか，輸出入価格の変化をうけるから，それらの影響を明別できるように輸出入価格指数が作られている。

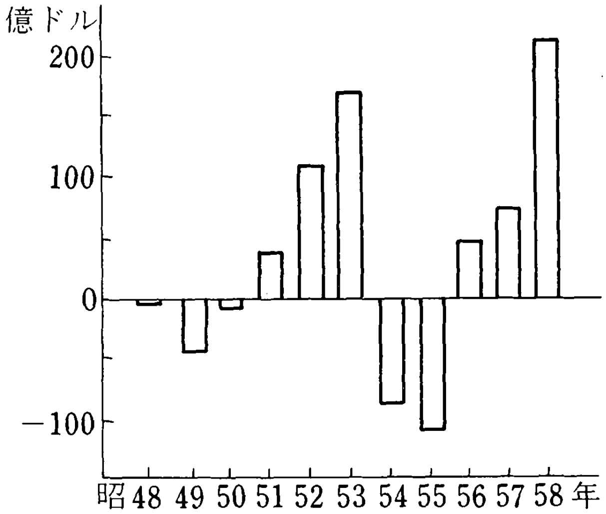


図1　わが国の経常収支

貿易収支の結果は，貿易外収支(商品の海上輸送や海外旅行費など)をも含めて経常収支によってとらえられる。

昭和58年の経常収支は，210億ドルの黒字であり，過去最高だった53年の165億ドルをも大幅に上回った。このことが主因となり，外貨準備高は12億ドル増加して，245億(58年末)ドルとなった。　(市野)

方眼紙

統計資料を図示したり，それに基づいて，何らかの傾向を引き出そうとしたりするとき，普通目盛りの方眼紙がよく用いられる。しかし，片対数方眼紙を用いたり，両対数方眼紙に打点した方が，適切な判断を下すのに，一層便利であることも多い。

★解説

ヒストグラムや**累積図表**は普通目盛りの方眼紙に書くことが多いが，**時系列**データなどは，**片対数方眼紙**に打点した方がよい場合も少なくない。片対数方眼紙の横軸は普通目盛りであるが，縦軸が対数目盛りになっている。どのようなときに片対数方眼紙を用いるとよいかについての原則的なきまりはないが，概して

・何らかの意味で，**比率**を問題にしたいとき，

・戦後の**物価指数**などのように，大小の極端に異なるものを表示したいとき，

には，片対数方眼紙がよい。

一般的には，変量X，Yの関係が，

$$Y=k\cdot\alpha^{X}$$

で表されるとき，これを図示するのに片対数方眼紙を用いるとよい。それはデータが，この方眼紙の上で，直線上にならぶからである。

一方，X，Y の関係が，

$$X^{\alpha}Y^{\beta}=C\ (一定)$$

で表されるようなときには，**両対数方眼紙**を用いるとよい。この場合，データは両対数方眼紙の上で直線上にならぶからである。片対数方眼紙を半対数方眼紙，両対数方眼紙を全対数方眼紙とよぶこともある。

関連ページ
ヒストグラム　275
累積図表　275
比率　316
物価指数　154

(例) 年間自動車生産台数

年次	車種	
	乗用車	トラック
昭和30年	2	4
32年	5	13
34年	8	18
36年	25	55

表1　年間生産台数
（単位：1万台）

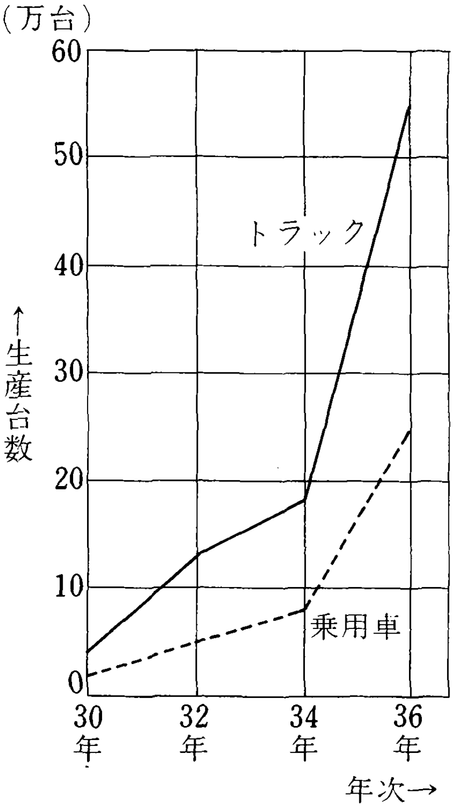


図1　普通目盛りの方眼紙への打点

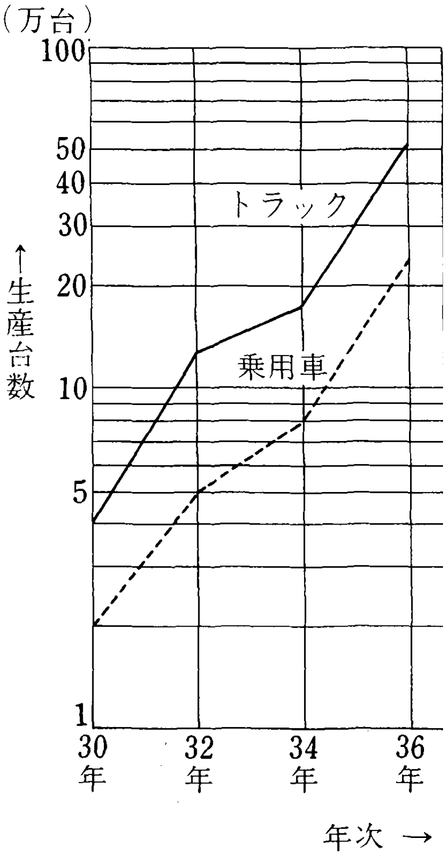


図2 片対数方眼紙への打点

表1は，わが国が経済的に一大飛躍をとげようとしていた昭和30年から36年までの自動車生産台数を示している。これを普通目盛りの方眼紙に書くと，図1のようになるが，片対数方眼紙に書くと図2のようになる。昭和34年から36年にかけては，乗用車およびトラックの相対的増加が，両者ともほぼ等しいはずなのに，図1は図2よりも，トラックの方が，急激な上昇を示しているような，誤った印象を与えがちである。これは，相対的増加は等しいが，絶対的増加でくらべると，トラックは乗用車のそれよりも上回っているからである。対数目盛りは，等しい相対的変化を，平行線で示す特性をもっているので，片対数方眼紙は，上のような変化を，視覚に訴えるのに便利である。

なお，n 組の統計資料

$$(x_1,\ y_1),\ (x_2,\ y_2),\ \cdots,\ (x_n,\ y_n)$$

を半対数方眼紙上に打点したとき，だいたい直線上にならぶことから，データに曲線

$$y=k\cdot\alpha^x \qquad (1)$$

をあてはめたいことがある。このようなとき，最小2乗法を用いて，k，α を推定する。しかし，直接式(1)に最小2乗法を適用しないで，式(1)の両辺の対数をとり，

$$Y=X+K \qquad (2)$$

（ただし，$Y=\log_\alpha y$, $X=x$, $K=\log_\alpha k$）

としておいて，式(2)に最小2乗法を適用するのがふつうである。　（牧野）

方策のキャッシュ・フロー

経済性の分析において，比較の原則(→98ページ)にしたがってまずはじめに明らかにしなければならないものが方策(案)のキャッシュ・フロー(**cash flows**)である。方策のキャッシュ・フローとは，その方策を実施することにより(実施しない場合と比べて)，収入と支出が，いつ，どれだけ増加あるいは減少するかを表したものである。

方策のキャッシュ・フローは，図で表示するとその内容が視覚的にとらえられてわかりやすくなる。

★解説

方策を実施することによって生じる収入，支出の変化の仕方は，一般には次の4通りになる。

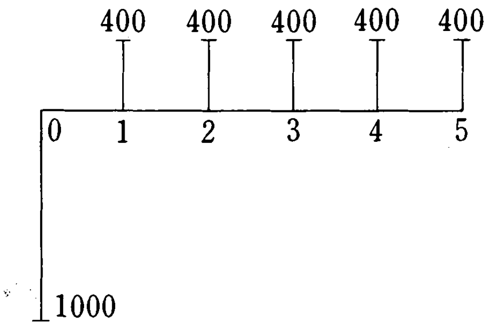
(1)　収入の増加
(2)　支出の減少
} 得を意味する。

(3)　収入の減少
(4)　支出の増加
} 損を意味する。

収入の増加と支出の減少は自己資金の増加をもたらし，得を意味する。キャッシュ・フローはプラスの値になる。収入の減少と支出の増加は自己資金を減少させ，損を意味する。キャッシュ・フローはマイナスの値になる。

たとえば，初期投資として1000万円かける(支出の増加)と，人件費がこの先5年にわたり年間400万円削減される(支出の減少)投資案のキャッシュ・フローは次のように示される。

現　在：－1000万円
1年後：　400万円
2年後：　400万円
3年後：　400万円
4年後：　400万円
5年後：　400万円



関連ページ
経済性の比較の原則　98
資金の時間的価値　144

年　度	設備費	材料費	人件費	経　費
0	5000	0	0	0
1	0	1000	5000	1500
2	0	1000	5000	1500
3	2000	1000	5000	1500
4	0	1000	5000	1500
5	0	1000	5000	1500

表 1　設備投資案のコスト（単位：万円）

年度	費用総額	収益	キャッシュ・フロー
0	5000	0	−5000
1	7500	10000	2500
2	7500	10000	2500
3	9500	10000	500
4	7500	10000	2500
5	7500	10000	2500

表 2　設備投資案のキャッシュ・フロー

方策(案)のキャッシュ・フローは，経済性の比較の原則にのっとって正しくつかまねばならない。

■キャッシュ・フローは比較の対象で違ってくる　たとえば，年当たり1億円の収益が予測される製品への投資案が，表1のように計画されたとしよう。

表1の数値は，この投資案を実施することによる(実施しない場合と比べての)支出の増加量になっている。設備費のところに減価償却費の値を用いたり，人件費や経費を表すのに配賦額を用いてはいけない。この投資案のキャッシュ・フローは，結局表2のようになる。

この投資案からえられる利益の正味

年度	設備費	省力額	キャッシュ・フロー
0	4000	0	−4000
1	0	1000	1000
2	0	1000	1000
3	2000	1000	−1000
4	0	1000	1000
5	0	1000	1000

表 3　人手案と比べたキャッシュ・フロー（単位：万円）

現価(→268ページ)は次式から求められる。ただし，$i=10\%$ とする。

$$P=-5000+\frac{2500}{(1+0.1)}+\frac{2500}{(1+0.1)^2}+\frac{500}{(1+0.1)^3}+\frac{2500}{(1+0.1)^4}+\frac{2500}{(1+0.5)^5}$$
$$=2974\ (\text{万円})$$

もし，この設備を使わず別の人手中心の製造方法が考えられるとしよう。それによると，設備費は初期投資の2000万円で済むかわりに人件費は年当たり6000万円かかることが見込まれている。

この人手中心の案を前者の設備投資案の比較の対象とすると，設備投資案のキャッシュ・フローは表3のように全く異なったものになる。

このキャッシュ・フロー(人手案と機械化案の差)の現在価値は −712万円になり，機械化の投資案は人手中心の案と比べ不利になる。　(中村)

ポートフォリオ(Portfolio)

ポートフォリオには，証券類を入れる“紙ばさみ”といった本来の意味があるが，それから転じて有価証券などの種々の金融資産の組み合わせのことをいう。各種金融資産に対して，その組み合わせ選択を行うことをポートフォリオ・セレクション(portfolio selection)という。なお，現在では，資本予算や戦略投資における投資対象の組み合わせとしての意味もあり，事業ポートフォリオや製品ポートフォリオといった使い方もされている。

★解説

株のようにその収益性(通常，投資収益率ではかる)が，かなり変動する危険性のある(リスクの高い)ものや，国債のようにかなり確実なものなど，証券の収益性はさまざまであり，同じ株や債券の中でも収益性の大きさやリスクの大きさはまたさまざまである。そこで，収益性を**確率変数**として**確率分布**で予測するならば，その分布の期待値で収益性の大きさを，また**分散**や標準偏差でリスクの大きさを表すことができる。通常の投資家は，収益性が同じ大きさならばリスクの小さい証券を好むものであり，このような行動をとる投資家をリスク回避者とよぶ。

投資家は，限られた投資可能額を複数の投資対象にうまく分散投資することにより，全体として収益性を確保しながらリスクを引き下げたいと考えている。そこで問題は，限られた予算の中で，全体としてもっとも自分にとって好ましい収益性の大きさとリスクの大きさをもたらす証券の組み合わせと投資配分率(額)を決めることである。こうした問題に対する理論的研究は，マーコビッツ(H. M. Markowitz)により創始され，**数理計画法**の2次計画法の問題としても研究されている。また，投資家の実際の投資行動の研究や，証券以外の投資対象への概念の拡張も行われている。

関連ページ

確率変数と確率分布　56

分散　332

数理計画法　186

■証券ポートフォリオ　たとえば株といった証券を考えた場合，ある経済情勢のもとでは高い収益性を示すが，別の経済情勢では低い収益性を示すものもあれば，その逆の動きを示す証券もある。さまざまな経済情勢に対し，さまざまに反応するだろう多数の証券があるとする。将来の経済情勢がどうなるか不確かな場合，もし投資家がリスクを回避したければ，どんな経済情勢が生じてもよいように，投資先をうまく分散させ，常にほどほどのもうけが得られるようにすればよいわけである。リスクの回避と収益性を高めることは，通常，同時に追求することができない。そこで，リスクを回避するには収益性を犠牲にすることになる。犠牲にしてもよい度合いは，投資家の投資態度によって違ってくる。このリスクに対する態度に応じた最適な証券の組み合わせ(ポートフォリオ)とそれぞれへの投資配分率を決めることが問題となる。

リスクの大きさを，収益性の分散で表すとすると，各証券の収益性に関する確率的推定(分布，あるいは期待値と分散)および，証券相互の間の変動性(共分散，あるいは相関係数)を推定できれば，ポートフォリオの収益性の期待値は各証券の期待値の配分率を重みとする重み付き平均値として示される。またポートフォリオのリスクは，各証券の分散・共分散を係数とする配分率の2次式で示される。そこで，投資家がリスクに対する態度を，許容できるリスクの大きさとして示したり，リスクと収益性との代替関係をリスク回避係数といった形で明示するならば，問題は2次計画法の問題として定式化され，投資家にとって最適なポートフォリオとそれぞれへの配分率を求めることができる。

他方，現実の投資専門家の投資行動を分析し，どのような情報をどのように利用してポートフォリオと配分率を決めているのかをモデル化しようとする実証的な研究が，クラークソン(P. E. Clarkson)によって始められ，計算機による投資問題支援システム化が考えられている。

■その他のポートフォリオ　経営戦略の策定にあたっている経営者の立場にたつと，資本予算を含む限られた経営資源をどの事業分野や製品分野にどう配分すべきかという問題になり，似た構造をもった問題に直面していることがわかる。そこで，今日では対象が事業や製品といった場合にも，ポートフォリオという言葉が使われるようになっている。ただし，証券の場合は収益性とリスクの2側面からポートフォリオの位置づけを考えているが，事業や製品の場合は，その市場の魅力度(成長性など)と市場での競争力とによって位置づけを行い，事業や製品の組み合わせと経営資源の配分を考えようとするものである。　(福川)

母数に関する仮説検定

母集団分布の平均値 μ が μ_0 に等しいという仮説をたて，これを標本に基づいて調べることを，母平均に関する仮説検定という。同様に母比率に関する検定，母分散についての検定などがよく用いられる。これらはいずれも母数に関する検定である。

★解説

「**仮説検定**」の項に例示してある**比率**の検定も，母数に関する検定の1つであり，その一般的な手順は，つぎの通りである。すなわち，標本数が大きい場合，母比率 p について**両側検定**を行うのに，仮説を

帰無仮説　H_0 ; $p=p_0$

対立仮説　H_1 ; $p \neq p_0$

のように立てておいて，検定すればよい。そのために標本平均 $\overline{p}$ を求めると，帰無仮説のもとでは，**統計量** $\overline{p}$ が**正規分布** $N(p_0,\ p_0(1-p_0)/n)$ にしたがう。ただし，n は標本の大きさである。このとき，統計量

$$T=(\overline{p}-p_0)\Big/\sqrt{\frac{p_0(1-p_0)}{n}}$$

を考えると，これは**標準正規分布**にしたがうので，

$|T| \geqq 1.96$

ならば，**有意水準**5％で帰無仮説を棄却してよい。また，

$|T| \geqq 2.58$

ならば，有意水準1％で帰無仮説を棄却してよい。なお，対立仮説を $p \neq p_0$ でなく，$p>p_0$（または $p<p_0$）のようにたてると片側検定になるが，この場合には，

$T \geqq 1.65$（または $T \leqq -1.65$）

のとき有意水準5％で，また，

$T \geqq 2.58$（または $T \leqq -2.58$）

のとき有意水準1％で，帰無仮説 $p=p_0$ を棄却する。

関連ページ
仮説検定　62
両側検定　64
帰無仮説　62
対立仮説　62
統計量　314
正規分布　200
標準正規分布　201
有意水準　62
片側検定　64

		正規母集団で分散既知のとき	正規母集団で分散未知のとき	標本数が大きいとき
計算すべき統計量(実現値)		$T=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$T=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n-1}}$	$T=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}$
両側検定	仮　　説	$\mu=\mu_0$	$\mu=\mu_0$	$\mu=\mu_0$
	有意水準5%での検定	$\lvert T \rvert \geqq 1.96$で$\mu=\mu_0$を棄却	$\lvert T \rvert \geqq t_{n-1}(0.05)$で$\mu=\mu_0$を棄却	$\lvert T \rvert \geqq 1.96$で$\mu=\mu_0$を棄却
	有意水準1%での検定	$\lvert T \rvert \geqq 2.58$で$\mu=\mu_0$を棄却	$\lvert T \rvert \geqq t_{n-1}(0.01)$で$\mu=\mu_0$を棄却	$\lvert T \rvert \geqq 2.58$で$\mu=\mu_0$を棄却
片側検定	仮　　説	$\mu=\mu_0$	$\mu=\mu_0$	$\mu=\mu_0$
	有意水準5%での検定	$T\geqq 1.65$で$\mu>\mu_0$と判定	$T\geqq t_{n-1}(0.10)$で$\mu>\mu_0$と判定	$T\geqq 1.65$で$\mu>\mu_0$と判定
	有意水準1%での検定	$T\geqq 2.33$で$\mu>\mu_0$と判定	$T\geqq t_{n-1}(0.02)$で$\mu>\mu_0$と判定	$T\geqq 2.33$で$\mu>\mu_0$と判定

ただしμ＝母平均，$\bar{x}$＝標本平均，α^2＝母分散，s^2＝標本分散，n＝標本の大きさで$t_{n-1}(0.05)$や$t_{n-1}(0.01)$の添字（$n-1$）はt分布表を読むときの自由度を示す。

表1　平均値の検定

■母平均の検定　母数の検定で，ふつうよく用いられるのは，平均値に関する検定である。表1は，いろいろな場合での平均値の検定の仕方を示している。検定に用いる統計量は，推定に用いるものと同じである。表には，仮説$\mu=\mu_0$と書いてあるが，これは帰無仮説のことである。対立仮説は，両側検定においては$\mu\neq\mu_0$であり，片側検定では$\mu>\mu_0$（または$\mu<\mu_0$）である。

■母分散の検定　これに用いる統計量も，母分散の推定のときのそれと同じである。すなわち，正規母集団$N(\mu, \sigma^2)$からの任意標本$X_1, X_2, \cdots, X_n$に基づいて，母分散についての検定を行うには，（両側検定の場合）仮説を

帰無仮説 H_0 ; $\sigma^2=\sigma_0^2$

対立仮説 H_1 ; $\sigma^2\neq\sigma_0^2$

のようにたてる。ここで，母平均μが既知の場合には，帰無仮説のもとで，統計量

$$\chi^2=\{(X_1-\overline{\mu})^2+\cdots+(X_n-\mu)^2\}/\sigma_0^2$$

が，自由度nのχ^2分布にしたがうことを用いて，仮説を検定することができる。しかし，母平均μが未知のときには，

$$\chi^2=\{(X_1-\overline{X})^2+\cdots+(X_n-\overline{X})^2\}/\sigma_0^2$$

が自由度（$n-1$）のχ^2分布にしたがうことを用いる。（牧野）

埋没費用

過去に現金の支出がされてしまったり，すでに支出が決定されていてとりかえしのきかない費用のことを埋没費用(sunk cost)という。経済性の分析では考慮しなくてよい費用である。

★解説

経済性の分析では，将来にわたる選択の可能性を問題にしており，済んでしまった過去の出来事を直接にとりあげることはしない。したがって，とりかえしのつかない費用，すなわち**埋没費用**については全く考慮しなくてよいのである。**経済性の比較の原則**を適用することにより，とりかえしのつかない費用はおのずと比較の対象間の違いとしてあらわれないことになる。たとえば，40万円の品物を買うのに10万円の手付金を支払ってあるとき，同じものが25万円で安く買えることがわかった。このとき，安い方を買うと手付金が損してしまうと考えてはいけない。

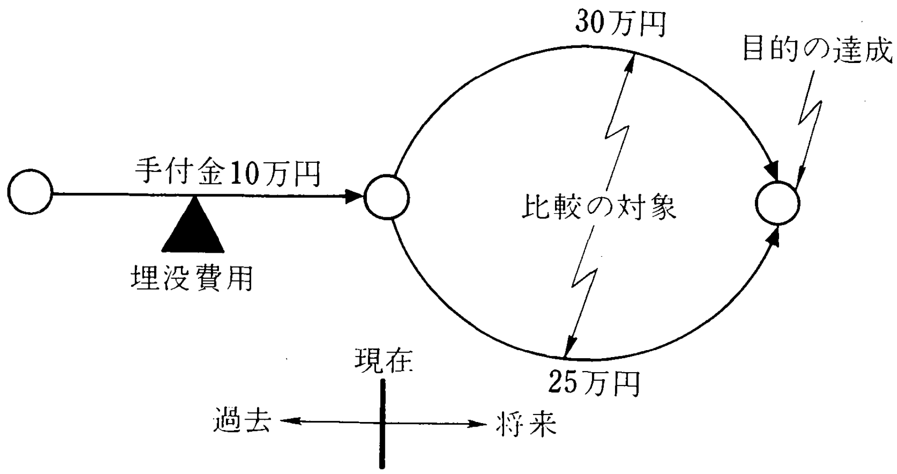


図1　埋没費用は過去のもの

埋没費用という用語の存在意義は，とりかえしのつかない費用を計算に考慮してしまう誤りをおかしがちなために，その注意を喚起する役割にあるといえよう。たとえば，設備の減価償却費は埋没費用なので，経済性の検討の際に直接とりあげてはいけない，ということになる。

関連ページ

経済性の比較の原則　98

■残ったビールの損　埋没費用の意味するところを理解するために，ごく単純な例を考えてみよう。

宴会も終わりになったころ，ビールが半分ほど余ったビール瓶をみて，「損をするから飲んじゃおう」ということがしばしばある。飲みたくてしょうがない人が我慢しているのなら，たしかに飲まなければ損するので飲むのがよい。

このビールは1本500円の値段になっているとしよう。ある人が，「半分残っているので250円損するから飲むのが得だ」といった。この人の計算は明らかに誤りである。

半分のビールを飲んでも飲まなくても，1本ぶんの500円の支出はすでにとりかえしのつかないものである。すなわち**埋没費用**である。したがって，比較の対象が〝飲む〟か〝飲まない〟かであれば，経済的な面ではどちらをとっても同じである。飲みたければ飲むのがよいし，飲みたくないのに無理して飲むことはないのである。

そうはいっても，半分残ったビールはなにか「もったいない」と思う経済観念の持ち主がいるかもしれない。もったいないと思う気持ちは大切にすべきである。余った資源を活用するアイディアを追求すべきである。

となりの部屋では宴もたけなわ。もし余ったビールを持っていくと100円で買ってくれるかもしれない。あるいは，2本もっていくと，となりで追加注文する500円の支出をしなくて済むかもしれない。こんなときは，半分のビールを飲んでしまうと，かえって経済的に損をすることになる。将来に目を向け，よいアイディア，方策を探すことが最も大切になる。

■設備費の取り扱い方　設備の減価償却費は，購入時に支出した購入価額を，使用期間に配賦(割り勘)したものである。したがって，購入済の設備を使う方策を検討する際は，償却費は埋没費用となり，計算にのせてはいけない。

工場に旧設備と新設備があって，特定の仕事をどちらで行うかを検討するとき，旧設備は減価償却が済んでいるので安くつく，などと考えるのは誤りである。実際に稼動することによってかかる費用だけで比較しなければならない。

短期的な問題と長期的な問題では，設備費の扱いは変わってくる。単純な例でみてみよう。Aさんはレンタカーを契約料1万円，km当たり200円の変動費で1日借りた。当日この車を100km走るBさんに又貸しする値段は，2万円以上ならペイする。契約料の1万円は**埋没費用**となる。Bさんが，明日も貸してほしいという場合は，契約料(設備費)の1万円を加え3万円以上で貸さないとペイしない。このときの1万円は埋没費用ではないのである。

(中村)

待ち行列

時折生ずる待ちや遅れを評価するために作られた数学的模型。もしくはそれを主な対象とする理論。確率論的な扱い方をする，ORの中心的手法の1つである。生産現場での工程管理，通信や計算機システムの性能評価，交通の遅れ解析など応用分野も極めて広い。トラヒック理論，確率的サービス・システム，混雑理論などとよばれることもある。

★解説

たとえば，スーパー・マーケットのレジでは，買物客は買った品物の料金を支払わねばならないが，もし先客がいれば自分の順番になるまで行列に並んで待つ。レジの代わりに工作機械，客の代わりに部品と考えれば，行列は仕掛かり在庫に当たる。レジの代わりに滑走路，客の代わりに航空機を考えることもできる。このように，順番待ちの行列(今後これを**待ち行列**とよぶ)のモデルによって定式化される現実の問題はすこぶる多い。

■待ち行列論の目標　待ちや遅れは，客の到着とサービス時間が確率的に変動するため，一時的に客をさばききれなくなって生ずる。その待ち時間などを定量的に評価し，客の到着に応じた適正なサービス能力を算定するための基礎資料を提供することが待ち行列論の目標といえよう。

■待ち行列モデルとその応用　待ち行列の基本モデルや基本的用語などについては，右ページで解説する。

待ち行列の応用分野の広さを反映して，待ち行列のモデルにはさまざまな型がある。それらを幾分厳密に取り扱おうとすると，一般にかなり難解なものとなるので，特殊な場合を除き，**図表**や**公式**を活用するのが実用的であろう。

また，現実には，ある待ち行列でのサービスを受け終えてから別の待ち行列に並ぶ例も多い。近年，**網状**に配置された待ち行列のモデルに対して関心が高まっている。

関連ページ

待ち行列とラッシュアワーの問題　368

待ち行列の図表　374

待ち行列の公式　370

待ち行列網　376

■**基本用語**　一般の待ち行列モデルは図1のように，**サービス**を提供する**窓口**(または**扱い者**)と，そのサービスを受けるために**到着**する客，及びそれに付随して時折生ずる**待ち行列**とから成る。待ち行列と窓口とを合わせて**待ち行列系**，系内の客の数を**系内数**という。

客が待ち行列系に到着したとき，窓口が空いていれば，直ちにサービスを受けるが，先客がサービス中のときには待ち行列を作って順番を待つ。もし，待つ場所(**待合室**ともいう)に制限がある場合には，到着客はサービスを受けることなく系外に立ち去る。このような客を**呼損**という。これは，客のことを呼とよぶこともあるので，呼が失われたという意味である。

■**複数窓口**　窓口が図のように複数個あるモデルの特徴は，待ち行列は共通で1本しかないということである。サービス中の客のサービスが終了して系から退去したとき，それがどの窓口からの退去であっても，次のサービス順の客が待ち行列からその窓口にはいってサービスを受け始める。$s=1$ のときと $s\geqq 1$ のときとでは，モデルの解析上の困難さが違うので，別々に扱われるのが普通である。

■**到着の確率法則**　ある時間区間内の客の到着数は確率的に変動するから，それがどんな確率分布(→56ページ)にしたがうかを仮定して待ち行列モデルは作られる。そのうち，最も大切なものは，ポアソン分布(→352ページ)にしたがう場合で，最も不規則的な到着のありさまを表すと考えられている(〔54〕の7ページ参照)。対照的に最も規則的な到着は，一定時間ごとに到着する場合である。前者を**ポアソン到着**，後者を**レギュラー到着**とよぶ。

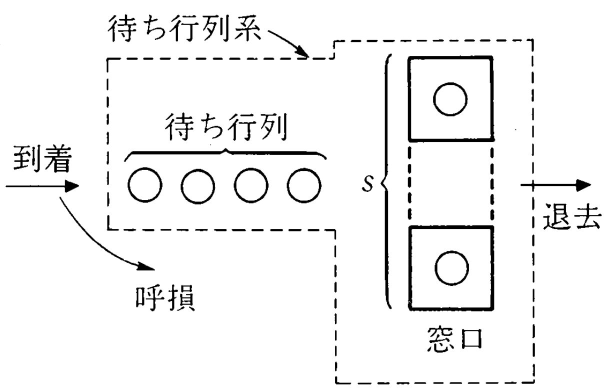


図1　待ち行列系の名称

■**サービス時間**　客が窓口を占有する時間で，これも確率的に変動する。最も不規則なのは指数分布にしたがう場合で，ポアソン到着に相当する。レギュラー到着に相当するのは，サービス時間がすべての客に対して一定の場合である。到着とサービスの時間的な偶然変動が待ちの原因であるから，ポアソン到着，指数分布のとき(数学的には簡単で**単純待ち行列**という)，待ちは大きく，現実の値は普通これを下回る。

■**ケンドールの記号**　待ち行列の型を表す記号。$A/B/c$ の形で，A と B にはそれぞれ到着とサービスの確率分布を記号で書き込む。c は窓口数，分布を示す記号は，M：指数分布，D：一定時間，G：一般分布などである。単純待ち行列は $M/M/s$ で表される。　(森村)

待ち行列とラッシュ・アワーの問題

待ち行列モデルは一種の混雑現象を記述するためのモデルではあるが，長い目で見れば処理能力は十分あって，時折，確率的に生ずる混雑を扱うのに適している。ラッシュ・アワーというのは，ある時間帯においては，とてもさばききれない量の客が殺到することに特徴があり，その数学的モデルは，ふつうの待ち行列モデルとは全く異なった「流体モデル」の方が適している。

★解説

■慢性的混雑と一時的混雑　混雑現象を解析するとき，第一に見きわめなければならないことは，それが慢性的混雑なのか，一時的混雑なのか，という点である。前者は，時折窓口の暇な状況が生ずるという特徴でとらえてもよい。たとえば，電話回線や券売機などの適正数を定める問題では，最も混む時間帯でも，爆発的に待ち行列や呼損が増加することのないよう設計しなければならない。その場合には，待ち行列論でいうトラヒック密度 ρ (→374ページ)が1より小さいという条件が必要になる。そして，そのことは，窓口が暇になる確率が正であることを保証している。

もし，$\rho>1$ となると，それは，サービスの能力を上回る数の客が殺到することを意味するが，このような事態がいつまでも続くようであると，待ちや呼損は増え続けて収拾がつかなくなる。ラッシュ・アワーとは，一時的に $\rho>1$ となる状況である。ただ，それがいつまでも続かず，しばらくすると $\rho<1$ となって平静になる。

■流体モデル　$\rho>1$ のときには，個々の客の待ちを考えるよりは，客を連続な流体とみなし，じょうごを使って瓶の中に液体を入れると溢れてしまうという現象と同じと考えると都合がよい。この考え方には確率変動がはいっていないので過小評価となる。

関連ページ
待ち行列　366

流体近似と拡散近似　408

■ラッシュのときの待ち行列　左のページでも説明したように，窓口への客の到着がサービスの能力を上回るときがラッシュ・アワーである。それは一時的にしか起こらない。図3には，単位時間あたりの到着客数(これを**到着率**とよぶ)の時間的変化の例を示している。

一方，サービス能力，つまり単位時間当たりにサービスを終えられる客数(これは平均サービス時間の逆数に等しい)は一定値 μ に等しいとしよう。

時点 t_0 までは $\rho<1$ の状況なので，水を瓶の中に注ぐ話になぞらえると，図2のような状況で水はじょうごの中に全く溜まっていない。同様に待ちは全く生じない。

しかし，t_0 を過ぎると $\rho>1$ の状況になり，図1のように水がじょうごの中に溜まり出す。つまり図4に示すように待ち行列ができ始める。図3の影の部分の面積は溜まった水(または客)の量すなわち待ち行列に当たるので，図4の曲線 $Q(t)$ の縦座標に等しい。

到着率が最大になる時点 t_1 の後も待ち行列の増え方は少なくなるものの，時点 t_2 までは待ち行列は増える。t_2 を過ぎると急激に減り始め，時点 t_3 で待ちは0になる。

このように待ち行列の挙動は確率的変動が主要原因でないラッシュの状況では，流体モデルがよく実情を表している。しかし，$\rho<1$ でも待ちはできるので，この場合にまで流体モデルを適用するのは適切ではない〔42〕。(森村)

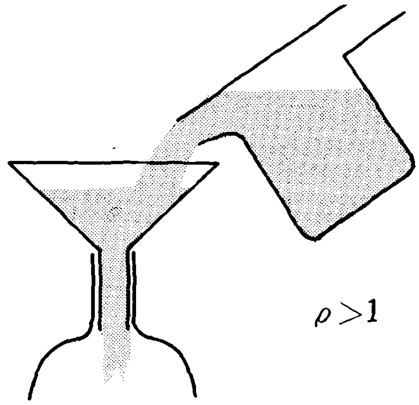


図1　流入量>流出量のときの一時的な溢れ

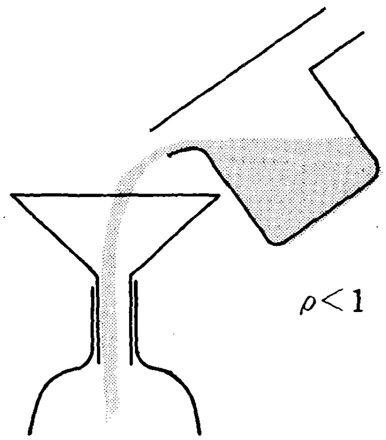


図2　流入量<流出量のときは溢れない

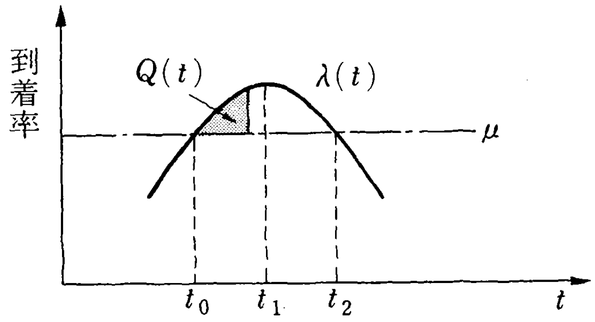


図3　到着率の時間的変化

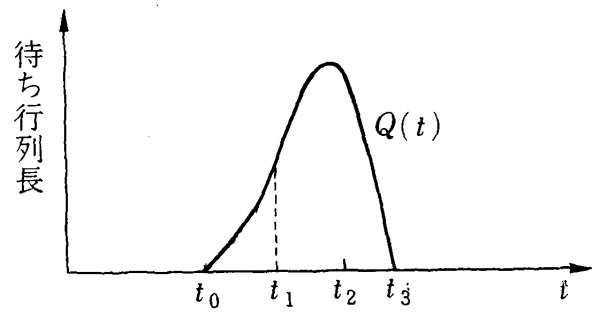


図4　待ち行列の増減

待ち行列の公式

待ち行列において，待ちの程度などを表す諸量を具体的で簡潔な形の数式で与えることは一般には困難である。わずかに $M/M/s$ の場合の諸量や $M/G/1$ の場合の平均待ち時間などを与える公式が直接の手計算に利用しうる程度である。その他，一般に成立するいくつかの関係も知られている。近年，平均待ち時間などに対するよい近似式が得られている。

★解説

■平均値の法則　系に到着した客がいつかはサービスを受ける限り，どのような待ち行列系においても，

$$L=\lambda W \qquad L_q=\lambda W_q \qquad \text{（リトルの公式）}$$

が成り立つ。ここで，λ は**平均到着率**，λ^{-1} は平均到着間隔。

■呼損系　$M/G/s$ で待合室の大きさが 0，すなわち，窓口が全部ふさがっているときに到着した客は呼損になる系の場合。系内数（すなわちサービス中の客の数）が n である確率を P_n とすると，$a=\lambda b$ とおいて，

$$P_n=\frac{a^n}{n!}\Big/\sum_{n=0}^{s}\frac{a^n}{n!} \qquad \text{（アーラン式）}$$

ここで b は平均サービス時間。

■単純待ち行列　$M/M/s$

$$P_0=\{\textstyle\sum_{i=1}^{s-1}a^i/i!+a^s[(s-1)!(s-a)]^{-1}\}^{-1}$$

$$P_n=a^nP_0/n! \quad (0<n<s), \quad =a^nP_0/s^{n-s}s! \quad (n\geqq s)$$

$$\Pi=a^sP_0/\{s!(1-a/s)\}$$

$$L_q=a^{s+1}P_0/\{(s-1)!(s-a)^2\}$$

■平均待ち時間　$M/G/1$ で

$$W_q=\frac{\lambda b_2}{2(1-\rho)},$$

b_2：サービス時間の 2 次積率

（ポラチェック・ヒンチンの公式）

関連ページ
待ち行列　366
待ち行列の図表　374

■リトルの公式

$$W=W_q+b, \quad L=L_q+\rho$$

の両式を使えば，W_q を441ページの図や，ポラチェック・ヒンチンの公式，あるいは近似式などから求めることで，L，L_q，W の諸量が求められる。ここで，b は平均サービス時間を示す。待ち行列網を扱うときなど，特に有効に利用される。具体的な計算例については375ページの例参照。

■アーラン式　1917年，アーラン(**A. K.Erlang**)によって求められた公式で，待ち行列論の創始と目される業績である。電話交換において，途中回線の話し中のため呼損となる確率，すなわち**呼損率**は，電話サービスの基準を定める量としてよく用いられているが，$n=s$ のときの P_n，つまり P_s が呼損率である。アーランは $M/M/s$ の場合に導いたが，40年後，$M/G/s$ でも同じ公式が成り立つことが知られた。

呼損率は電話事業で用いられるので〔53〕などに詳しい数表がある。しかし，$\rho=a/s$ の値が極端に1に近くなければ，P_n はポアソン分布 $(a^n/n!)e^{-a}$ で近似しても実用上困らないことも多い。例として，$a=3$，$s=6$ のとき両者の比較を次表に示す。

n	0	2	4	6
アーラン式	.052	.232	.174	.052
ポアソン近似	.050	.224	.168	.050

■ポラチェック・ヒンチンの公式　サービス時間の分布が任意のときに平均待ち時間 W_q を与えるので便利である。

1930年代初頭にポラチェックとヒンチンにより独立に求められた。$s=1$ のときに成り立つが，$s>1$ に対しては同様の公式はない。この式を変形すると，

$$W_q=(M/M/1 \text{ の } W_q)\cdot\frac{1+C^2}{2} \qquad (1)$$

となる。ここで C はサービス時間の変動係数，すなわち，標準偏差／平均値を示す。したがって，$M/D/1$ の W_q は $M/M/1$ の W_q の1/2に等しい。

■近似式　$GI/G/1$ における平均待ち時間 W_q の近似式にはいくつかのものが提案されているが，比較的簡単でかつ精度の点でも優れているものに，次の式がある。

$$W_q=\frac{\lambda b^2(C_a^2+C_s^2)}{2(1-\rho)}$$

ここで，C_a，C_s はそれぞれ到着間隔分布およびサービス時間分布の変動係数を示す。

複数窓口の場合には，$M/G/s$ に対し，式(1)を拡張した

$$W_q=(M/M/s \text{ の } W_q)\cdot\frac{1+C^2}{2}$$

も提案され，粗い近似としては有用であるが，もう少し精度の高いものとしては，

$$W_q=C^2(M/M/s \text{ の } W_q) +(1-C^2)(M/D/s \text{ の } W_q)$$

も使いやすい。　　　　（森村）

待ち行列の数値解析

各種の公式，近似式，図表などが利用できない非標準的モデルに対してモデルの特性量を求めるには，シミュレーションや数値解析が用いられる。シミュレーションは複雑なモデルも扱えるが，計算時間がかかるわりに精度が悪い。そのため，可能な場合には，数値解析を用いるのがよい。

★解説

■数値解析のいろいろ　数値解析にもいくつもの種類がある。マルコフ連鎖に帰着させ定常確率を数値的に計算して特性量を求めるもの，解析的方法で求められたラプラス変換やフーリエ変換を数値的に逆変換するもの，拡散近似を精度よく行うために数値的取り扱いをするもの，などいろいろある。とくに，最初のものは適用できるモデルのバリエーションが多いのでよく用いられる。

■状態とマルコフ連鎖　マルコフ連鎖を利用するには，マルコフ性が成り立つように待ち行列システムの状態が決められなければならない。サービス時間分布や到着間隔分布が，**指数分布**，**アーラン分布**，**超指数分布**など(このような分布を一般に相型分布という)で与えられる場合には，各分布の相や客数などの組をとることによって，時間パラメータ連続のマルコフ連鎖に帰着されることが多い(右ページの例参照)。

■定常確率とその計算法　待ち行列で用いられるマルコフ連鎖はほとんどエルゴード的なので，各状態にいる確率 α_i が推移確率 p_{ij} からつぎのように決められる。

$$\sum_i \alpha_i p_{ij} = \alpha_j, \qquad \sum_i \alpha_i = 1$$

これは α_i を変数とする連立1次方程式なので，ガウス・ザイデル反復法などの数値計算法が利用できる。

関連ページ
待ち行列　366
マルコフ連鎖　382
指数分布　156
アーラン分布　66

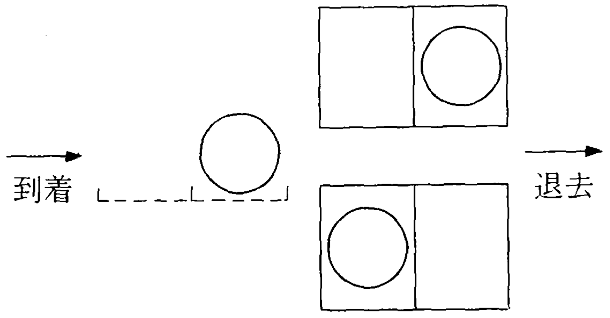


図 1 $M/E_2/2[4]$モデル

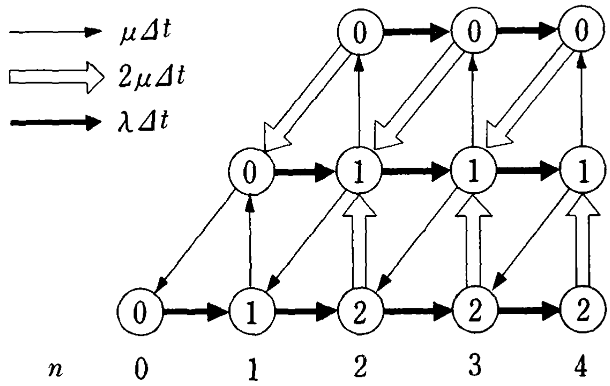


図 2 推移図式

■アーランサービス ケンドールの記号で $M/E_2/2[4]$ と書かれる場合，つまりポアソン到着(到着間隔が指数分布)，サービス分布がフェーズ 2 のアーラン分布，窓口の数は2つで，系内には 4 人までしか入れない(つまり行列で待てる人は 2 人までで，それを超えるときは呼損となる)場合を考えよう。

このときのシステムは図 1 のようなイメージとなる。客は一方の窓口に入ると，まず第 1 のフェーズで指数分布に従うサービスを受け，それが終わると，引き続き第 2 のフェーズでやはり同じ指数分布に従うサービスを受ける。

システムの状態は組 (n, i, j) で表される。ここで，n は系内の客数，i はフェーズ 1 でサービスを受けている客数，j はフェーズ 2 でサービスを受けている客数である。図 1 で示されている状態は，(3, 1, 1) というわけである。n と i が決まると j は自動的に決まるので，状態は組 (n, i) で表されるといってもよい。すると，マルコフ連鎖の推移図式は図 2 のように書ける。丸の中の数字は i，λ は到着率(平均到着間隔の逆数)，$\mu/2$ はサービス率(平均サービス時間の逆数)，Δt は小さな正数である。ただし，自分自身へ戻る矢印は省略してある。

この推移図式から推移確率を読みとり，左ページの連立 1 次方程式を解けば，各状態の定常確率が求められる。この定常確率をもとにして呼損率，平均系内人数，平均行列長などの特性量を計算するのである。たとえば，呼損率は $n=4$ の 3 つの状態に対応する定常確率を加えることによって求められる。

■マルコフ連鎖の数値解法 待ち行列の数値解析で最も計算量を必要とする部分は，マルコフ連鎖の定常確率を連立 1 次方程式から解くところである。ここには連立 1 次方程式の各種の数値解法が適用できるが，係数に特殊な構造があるため，消去法よりも反復法系統のものが適している。ブロック・ガウス・ザイデル法などがよいであろう。モデルの構造をさらに利用した数値計算法もいくつか提案されている。

(高橋)

待ち行列の図表

待ちの程度を示す量として，平均待ち時間 (W_q)，平均系内時間 (W)，平均待ち行列長 (L_q)，平均系内数 (L)，待ち率 (Π)，待ち時間分布などがよく利用される。待ち行列の基本公式を利用すると，W_q と Π の値が求まると他の諸量も求められる。$M/M/s$ の待ち行列モデルにおける Π と W_q の値を求める図表とその使用例をここに示す。

★解説

■待ち率　ある客が到着したとき，窓口がすべてふさがっていて，待たなければならない確率をいう。窓口が1個の場合には，到着やサービスの確率分布が何であっても，常に a＝平均サービス時間／平均到着間隔 に等しい。しかし，窓口が複数のときは簡単ではない。$M/M/s$ の場合には公式として与えられるが，それでもいちいち数値計算をするのはかなりめんどうである。440ページに示す図表には，a の値に対する Π の値を示してある。

■待ち時間分布　客にとって待ちの程度を端的に表すのは待ち時間が t 以上となる確率 $P(>t)$ である。この量は，

$$P(>t)=\Pi\cdot e^{-\mu st(1-\rho)} \tag{1}$$

と表される。ここで μ は平均サービス時間の逆数であり，ρ は $\rho=a/s$ として与えられ，**トラヒック密度**とよばれる量である。e は自然対数の底を意味する。

■平均待ち時間　441ページに示す図表には，ρ の値に対して μW_q の値を与えている。μ の定義から，

$$\mu W_q=\frac{\text{平均待ち時間}}{\text{平均サービス時間}}$$

つまり，平均サービス時間の何倍を平均として待つか，という量を示している。**リトルの公式**を利用すると，この値から L_q が求まり，L や W もそれらから求められる。

関連ページ
待ち行列　366

待ち行列の公式
370

■ Π を求める例題　定期便を運航している航空会社は，各空港に整備ステーションを用意しているが，そこには，さまざまな部品を予備として在庫しておかなければならない。予備品は毎日まとまって何個も使用されるものではないから，使用のたびに中央倉庫に発注して，常備品数 s はできるだけ常備しているようにする。最適な s を求めたい。いま，ある部品の予備品使用は週当たり平均 2 回，発注から納品まで平均 1 週間と見積もられる。この時間をサービス時間と見，予備品が在庫していない状態を客が窓口でサービス中と解釈すると，$M/M/s$ 待ち行列モデルで定式化できる。このとき，Π は予備品が全部使われていて，品切れを起こす確率に相当する。いま，それを 1 年に 1 度程度，すなわち週当たりでは 0.02 の確率に抑えることとしよう。440ページの図表で，横座標 $a=2$ の線上，縦座標 $\Pi \leqq 0.02$ となる最小の s は 6 であることから，この部品は常時 6 個を予備として在庫しておけばよいことになる。すると在庫切れは 1 年に 1 度程度ということになる。

■ $P(>t)$ を求める例題　オンラインの自動支払機で，利用客が確認のボタンを押してから支払いがなされるまでにあまり長時間待たせることは避けたい。したがって，中央の計算機で先客処理中のために待たせる時間はたかだか 3 秒程度に抑えたいとしよう。すなわち，3 秒以上待たせる確率 $P(>3)$ を 0.01 以下にしたいという要請があったとする。

利用客の支払い要求は，平均して毎秒 8.5 個，計算機の処理時間は平均 0.1 秒とすると，$a=\Pi=0.85$, $\mu s t=30$ となる。電卓もしくは指数関数表，あるいは参考文献〔54〕の213ページの付図を用いると，式 (1) は $0.85\times 0.012=0.01$ となって条件をちょうど満たす。

■ W_q を求める例題　銀行のオンライン自動支払機の前に長い行列ができていることを時折見かける。ある支店で，何台の支払機を使用可能とすれば客に我慢してもらえるか，を考えよう。

客の平均到着率は，$\lambda=85$ 人/時，平均取扱時間 1.4 分と推定されたとする。$a=\lambda/\mu=1.98$ であるから，窓口数 $s=2$ 以上でないと，収拾できかねる状況である。$s=2$ とすると，$\rho=a/s=0.99$ となるので，441ページの表から μW_q を求めると 50 となる。リトルの公式（→370ページ）によると，

$L_q=\lambda W_q=a\cdot\mu \quad W_q=1.98\times 50=99$

となって，平均 100 人近くの客が行列に並ぶことになる。これではひどすぎる。そこで，$s=3$ とすると，$\rho=0.66$ であるから $L_q=1.98\times 0.43=0.85$ で，待ち行列は平均 1 人以下で問題にならない。

こうして，この支店ではなんとか 3 台の支払機を使えるように考えたい，という結論になる。　（森村）

待ち行列網

複数の待ち行列がネットワーク状につながったもの。混雑しているシステムではあちらこちらに待ち行列ができることが多いので，適用範囲は広い。計算機や通信システムの設計には欠かせないモデルである。しかし解析の難しさから，実際に使われるモデルは $M/M/s$ 型の待ち行列がつながったものがほとんどである。特殊なケースとして，直列型待ち行列，巡回型待ち行列などを含む。

★解説

■直列型待ち行列　図1のように，いくつかの待ち行列が直列につながったものをいう。病院で受付，診察，会計と順に待たされる場合とか，いくつかの工程を経て加工される場合の待ち時間やスループット(システムとして単位時間当たりサービス可能な客数)の解析などに用いられる。

各段の待合室の大きさに制限がある(有限バッファの)場合があり，このときはある段でサービスが終了しても次段の待合室がいっぱいであれば先に進めず，そのサービス窓口をふさいだまま先が空くのを待つ。この現象を**ブロッキング**という。ブロッキングが生じない，つまりすべて無限バッファであれば，システムのスループットは処理能力の一番小さい段のスループットでおさえられる。しかし，ブロッキングが生じる場合には，システムのスループットは各段のスループットよりも小さくなることがある。そして，これは待ち行列の配置の順番にも依存する。

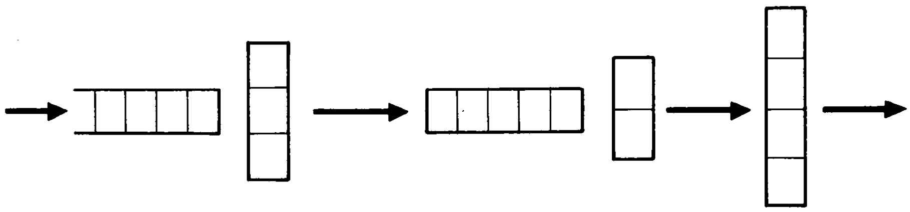


図1　直列型待ち行列

関連ページ
待ち行列　366

マルコフ連鎖
382

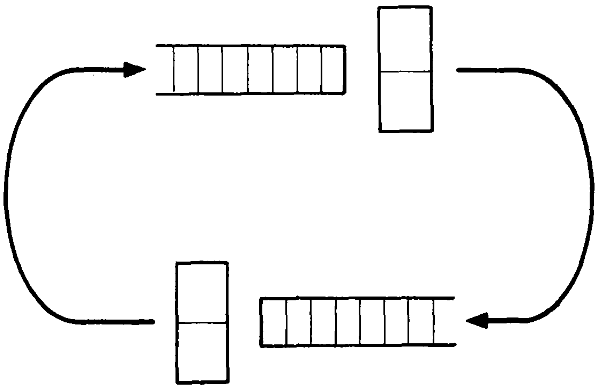


図 2　巡回型待ち行列

■巡回型待ち行列　図2のように，客がいくつかの待ち行列を順に巡るような待ち行列網を巡回型待ち行列という。ある場所から別の場所へ部品や原材料を運ぶために行ったりきたりしている船，トラック，台車，箱などの動きを解析するときなどに使われる。

■ジャクソン(Jackson)型ネットワーク　図3のように，無限待合室をもった $M/M/s$ 型の待ち行列がいくつかあり，i 番目の待ち行列でサービスを終了した客は，p_{ij} の確率で待ち行列 j に移るというモデルである。系の外から客が入ってこず，また出てもいかない場合，つまり一定の数の客が網の中を動き回る場合を**閉じた網**といい，図のように外から各待ち行列に独立にポアソン到着し，サービスを終了した客はそれぞれ決まった確率で系の外へ出ていく場合を**開いた網**という。

■積形式　ジャクソン型ネットワークでは，各待ち行列の人数(サービス中の客の数も含む，以下同じ)の組を状態にとるとマルコフ連鎖となる。その定常確率は開いた網であれば，各待ち行列で対応した人数のいる(周辺)確率の積になっている。この性質を用いると，外からの各到着率と推移確率 $\{p_{ij}\}$ を用いて各待ち行列への実質的な到着率が計算でき，$M/M/s$ の公式を使って定常確率が求められる。

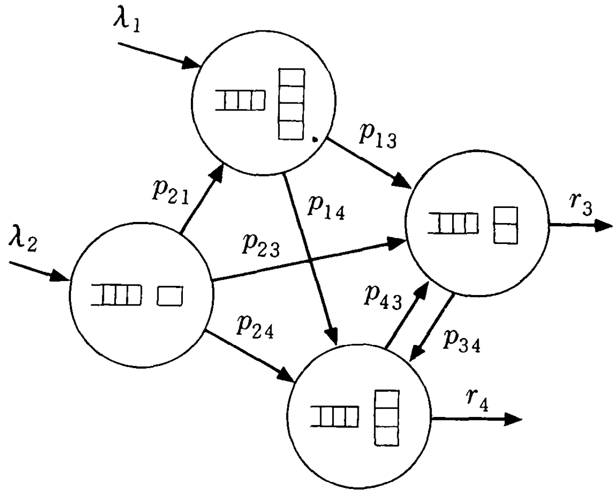


図 3　ジャクソン型ネットワーク

閉じた網でも定常確率は同様の積形式を持つが，網内人数が一定という条件があるため，ある規準化定数が掛けられる。この規準化定数は，いざ計算をしようとすると，かなりの計算量を必要とするため，うまい計算方法がいくつか研究されている。

■BCMP型ネットワーク　ジャクソン型ネットワークで，客の種類が1種類でなく，いくつかのタイプの客が存在したり，窓口におけるサービス率がその待ち行列の人数に依存するような場合でも，定常確率はやはり積形式になる。このような型の網を最初に研究した4人の学者の頭文字をとってBCMP型とよんでいる。計算機の性能評価によく用いられる。　(高橋)

マハラノビス距離

2点間の距離として，正規分布のもとで最も自然に導かれるのがマハラノビス距離である。これは正規分布の曲面にそって測っており，勾配の急な所は平面的には近くても大きな距離を，勾配の緩やかなところは平面的には遠くても小さな距離を与える。正規分布の勾配の度合いを考慮した距離であり，よく用いられるのはそれを2乗したマハラノビス平方距離である。群 G_1 か群 G_2 のいずれかに属することが確かである標本 $\underset{\sim}{x}$ がどちらに属するかは，群 G_1 の平均と $\underset{\sim}{x}$ とのマハラノビス平方距離と，群 G_2 の平均と $\underset{\sim}{x}$ とのマハラノビス平方距離との大小によって判定する。平均までの距離が近い群に，標本 $\underset{\sim}{x}$ は属すると判定する。

★解説

関連ページ

1変数で考えると，平均 μ と標本 x とのマハラノビス平方距離は，

$$D^2=\left(\frac{x-\mu}{\sigma}\right)^2$$

である。2変量のときの平均$(\mu_1,\ \mu_2)$と標本$(x_1,\ x_2)$とのマハラノビス平方距離は，

$$D^2=\frac{1}{1-\rho^2}\left\{\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right\}$$

である。ここで，σ_{11} は x_1 の分散，σ_{22} は x_2 の分散，ρ は x_1 と x_2 の相関係数である。

p 変数の場合，**主成分**を$(Y_1,\ Y_2,\ \cdots,\ Y_p)$とし，それぞれの主成分の分散を$(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_p)$で表す。このとき，主成分のはる Y 座標での平均$(\mu_1,\ \mu_2,\ \cdots,\ \mu_p)$と標本点$(y_1,\ y_2,\ \cdots,\ y_p)$とのマハラノビス平方距離は，

主成分　176

$$D^2=\frac{(y_1-\mu_1)^2}{\sqrt{\lambda_1}}+\frac{(y_2-\mu_2)^2}{\sqrt{\lambda_2}}+\cdots+\frac{(y_p-\mu_p)^2}{\sqrt{\lambda_p}}$$

である。主成分 Y は $x_1,\ x_2,\ \cdots,\ x_p$ の1次式で表されるから，この D^2 は2変量の場合のように，$x_1,\ x_2,\ \cdots,\ x_p$ によって表現することもできる。

2 変量 x_1，x_2 の場合を例にして説明する。群 G_1 に属する 27 個の標本と，群 G_2 に属する 27 個の標本が得られたとする。それぞれの群で，変数 x_1，x_2 は 2 変量正規分布に従っており，また分散共分散行列が等しいことがわかっているとする。

群 G_1 の標本を○印で，群 G_2 の標本を△印でかいたのが図 1 である。平均は，

	群 G_1 の平均	群 G_2 の平均
$\overline{x}_1$	57.9	48.5
$\overline{x}_2$	53.0	50.4

である。共通の分散は，

x_1 の分散　17.0

x_2 の分散　23.0

であり，x_1 と x_2 の相関係数は 0.57 である。

群 G_1 の平均と標本(x_1，x_2)とのマハラノビス平方距離は，上記の推定値を用いて，

$$D_1^2=\frac{1}{1-0.57^2}\left\{\left(\frac{x_1-57.9}{\sqrt{17}}\right)^2+\left(\frac{x_2-53.0}{\sqrt{23}}\right)^2 -2\times0.57\left(\frac{x_1-57.9}{\sqrt{17}}\right)\left(\frac{x_2-53.0}{\sqrt{23}}\right)\right\}$$

とかける。群 G_1 のデータ(x_1，x_2)＝(52，51)を代入すると，このデータと平均(57.9，53.0)のマハラノビス平方距離

$$D_1^2=2.37$$

を得る。

群 G_2 の平均と標本(x_1，x_2)とのマハラノビス平方距離は，

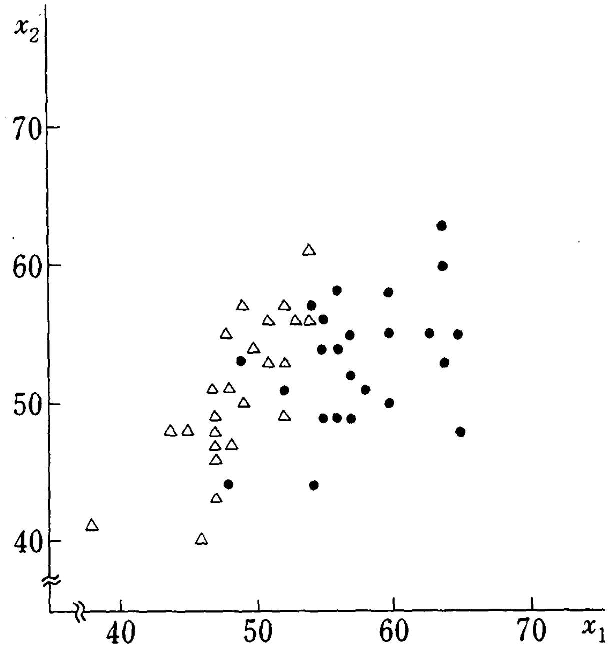


図 3　●印は群 G_1 のデータ，△印は群 G_2 のデータ

$$D_2^2=\frac{1}{1-0.57^2}\left\{\left(\frac{x_1-48.5}{\sqrt{17}}\right)^2+\left(\frac{x_2-50.4}{\sqrt{23}}\right)^2 -2\times0.57\left(\frac{x_1-48.5}{\sqrt{17}}\right)\left(\frac{x_2-50.4}{\sqrt{23}}\right)\right\}$$

であり，データ(52，51)と群 G_2 の平均(48.5，50.4)とのマハラノビス平方距離は，

$$D_2^2=0.96$$

となる。$D_1^2>D_2^2$ であり，データ(52，51)は群 G_2 に近いことになる。このデータは，判別分析では群 G_2 に属すると判定される。

判別効率とは，群 G_1 の平均と群 G_2 の平均とのマハラノビス平方距離であり，判別効率が大きければ大きいほど，誤判別の確率は小さいと期待される。

（杉山）

マルコフ決定過程

各期にとられた決定に従って，状態が確率的に(マルコル連鎖に従って)推移する多段階決定過程をいう。一般の計画期間有限の場合には，動的計画法によって最適な政策が求められる。計画期間が無限で推移確率や利得関数が定常な場合には，政策反復解法によって効率的に解かれることがハワード(R. A. Howard)によって示された(1960)。信頼性や需要が確率的に変動する生産在庫モデルの分析などに広く用いられている。

★解説

■マルコフ決定過程　第 t 期に状態が i であったとき，決定 k をとると，次の第 $t+1$ 期には $p_{ij}{}^{k}(t)$ の確率で状態 j へ推移する。このときに $r_{ij}{}^{k}(t)$ の利得が得られる。計画期間全体にわたっての，期待総利得を最大にする政策(最適政策)を求めるのが問題である。

■マルコフ決定過程のタイプ　計画期間は有限か無限か，推移確率や利得関数が定常(つまり期 t に関係しない)かどうか，利得の評価に割引率 β を導入するかどうか，などによっていくつかのタイプに分類される。

■定常政策とハワードの政策反復解法　とり得る状態や決定の数が有限で，推移確率・利得関数が定常，計画期間無限の場合には，**定常な最適政策**(つまり，期 t によらず，状態 i では必ず決定 k_i をとるという形の最適政策)が存在することが知られており，この定常最適政策を求める方法としてハワードは**政策反復解法**が優れていることを示した。この方法によれば，有限回の手続きで最適政策を求めることができる。ただし，割引率がある($\beta<1$)場合のほうが簡単で，割引率がない($\beta=1$)場合には計算だけでなく最適性の定義などにも少々面倒な考察を必要とする。

関連ページ
マルコフ連鎖　382
動的計画法　270
在庫モデル　128

■**予防保全への応用例**　ある機械は，その劣化の程度に応じて1から4までの状態のいずれかをとる。1は新品，4は使用不能で交換を要する状態である。通常の使い方に従うと，第 t 期で状態 i にあったものが，第 $t+1$ 期で状態 $i+1$ に推移する確率は0.2，状態 i にとどまる確率は0.8である。そして1期あたり，状態1では10，状態2では8，状態3では4の利得がある。状態4では機械を取り替えなければならないので，利得は -50 である。

一方，この機械は特別な手入れをしながら使うこともできる。すると状態 i から状態 $i+1$ へ推移する確率を半分の0.1にできるかわりに，得られる利得は，毎期2ずつ減って，状態1では8，状態2では6，状態3では2になる。状態4は取り替えなので -50 で変わらない。このような状況は，表1のようにまとめることができる。各欄で上段の数字は通常の使用，下段の数字は手入れをしながらの使用を示す。

割引率を $\beta=.95$ として最適政策を求めると，状態1と2では手入れをして使用したほうがよく，状態3では手入れをしないほうがよいという結果が得られる。

■**需要が確率的に変わる生産在庫モデルへの応用例**　定期発注方式の在庫モデル(→244ページ)で調達期間がある場合，需要が確率的であると，ある程度の安全余裕をみておく必要があった。

$$
P=\begin{array}{c} \\ 1 \\ \\ 2 \\ \\ 3 \\ \\ 4 \end{array}
\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\
\left[\begin{array}{cccc} .8 & .2 & & \\ .9 & .1 & & \\ & .8 & .2 & \\ & .9 & .1 & \\ & & .8 & .2 \\ & & .9 & .1 \\ 1 & & & \end{array}\right] \end{array}
\begin{array}{c} \text{利得} \\ 10 \\ 8 \\ 8 \\ 6 \\ 4 \\ 2 \\ -50 \end{array}
$$

しかし，そこで用いた安全余裕の計算方法は，扱う商品の量が大きく，変動が毎期独立で正規分布に従う場合を想定したものである。扱う商品の数が毎期数個しかなかったり，需要の変動が独立でなく相関をもっていたとすると，このマルコフ決定過程を用いて定式化しなければならない。そのときの状態変数には，在庫量や前期の需要量(需要変動に相関のある場合)などが入ってくる。

また，動的計画法の項で紹介した生産在庫管理モデル(→271ページ)でも，需要が確率的にならず，マルコフ決定過程で扱うことになる。ただし，扱わなければならない状態や決定の数が膨大となるので，うまい数値計算プログラムを作って計算しなければならないであろう。

応用例や政策反復解法については，参考文献〔43〕または〔55〕を参考のこと。

（高橋）

マルコフ連鎖

「将来における事象の起こる確率は，現在の状態だけに依存し，過去の経歴にはよらない」という性質をマルコフ性という。離散的な値(通常は自然数)をとりながら，確率的に変動する量がマルコフ性を満たすとき，マルコフ連鎖とよぶ。過去の影響を将来に及ぼすような現象の中にはマルコフ連鎖で表される例も多い。待ち行列や在庫のモデルにおいては，その理論的基礎ともなっている。

★解説

■マルコフ連鎖の例　第 n 週末の在庫量 X_n は，前週末の在庫量 X_{n-1} に，第 n 週の納入量 Y_n を加え，需要 Z_n を差し引いたものに等しい。Y_n や Z_n は確率的に変動するが，Y_n が第 $n-2$ 週以前の在庫量によって定められたものでない限り，X_n の確率分布は，X_{n-1} が与えられたとき Y_n や Z_n の確率分布を用いて表すことができる。それは X_{n-2}，X_{n-3} などには無関係である。つまり $\{X_n\}$ はマルコフ性を満たしている。その他の例については右ページ参照。

■用語　マルコフ連鎖のとりうる値を**状態**とよぶ。状態は自然数で表しておく。マルコフ連鎖が状態 i から状態 j へ変わることを i から j への**推移**とよび，その推移が生ずる確率 P_{ij} を**推移確率**という。P_{ij} をマトリックスの形に並べたもの P を**推移確率行列**という。P はマルコフ連鎖の動きを規制し，基本的に大切なものである。

■吸収マルコフ連鎖　P の形によっては，状態が 2 種類のものに分かれ，**一時的**とよばれる状態には，有限回の推移はしても，いつしか推移しなくなり，**吸収的**とよばれる状態にはいって動かなくなる。これを吸収マルコフ連鎖という。

■エルゴード的マルコフ連鎖　どの状態へも繰り返し何度でも推移するようなマルコフ連鎖をいう。待ち行列の多くはこの型である。一般の場合にはこれら 2 種の合成とみなせるものもある。

関連ページ
マルコフ決定過程　380
待ち行列　366
在庫モデル　128

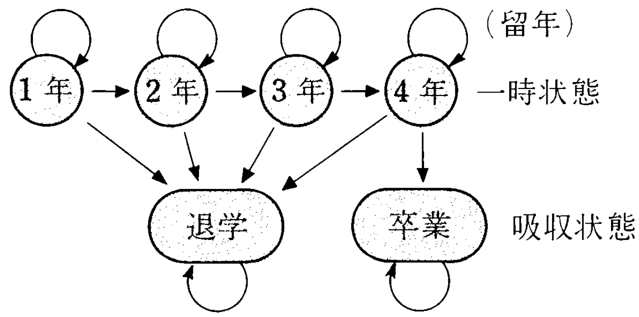


図1　進級モデルの推移図式
（吸収マルコフ連鎖の例）

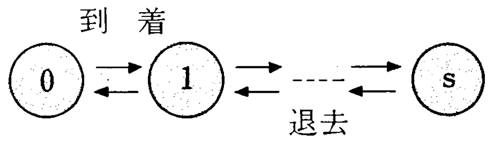


図2　呼損系の待ち行列モデル
（エルゴード的マルコフ連鎖の例）

■進級モデル　大学に入学すると，ふつうは1年生となり，以後，2年，3年，4年と進級して卒業するが，途中で退学する人もいる。図1は，これをマルコフ連鎖とみなすときの状態と，その間の推移を表したもので，**推移図式**とよばれる。1年生，…，4年生の4つの状態は，ある学生にとって，いつしかそこへは推移しなくなる状態であるから，一時状態である。そして，退学してしまったり卒業したりしたときは，その状態から動かなくなるから吸収状態である。

図1で実線の矢印は，正の確率で推移の起こるところにつけられている。留年すれば，もう一度同じ状態を繰り返すから，自分自身に戻る矢印もかかれている。吸収状態から出る矢印は必ず自身に戻る。

推移図式ではふつう矢印の傍らに推移確率の値を記入する。すると，推移図式は推移確率行列と同じ内容を表すことになる。しかし，マトリックスは演算もできるのに推移図式ではそれはできない。このため，直観的なイメージをつかむためにのみ推移図式は用いられ，通常は推移確率行列を利用して解析を進める。たとえば，平均在学期間などは，Pに基づいた行列演算で簡単に計算できる。1年生などを係員，係長，課長などとすれば人事計画に利用できるし，病気の程度とすれば，予防医学の評価にもなる。

■待ち行列モデル　呼損系の待ち行列（→370ページ）で系内数を状態にとる。客が到着すると系内数は1つ増え，サービスを終了して退去すると系内数は1つ減る。一度に2つ増えたり減ったりすることは起こらない。また，窓口が全部ふさがっているとき，すなわち系内数がsのときに到着した客は呼損となるので，系内数は増えない。それで，推移図式は図2のようになる。

ここでは，どの状態への推移も繰り返し起こり，吸収されて動かなくなるようなことはありえない。エルゴード的マルコフ連鎖では，それぞれの状態にいる確率が時間によって変化しなくなると考えてよい。参考文献〔55〕

（森村）

無作為化

実験順序と因子の効果が区別できるように，実験をどのような組み合わせから始めるかを，クジあるいは乱数表などを用いて決めること。時間的，場所的などの系統誤差を転化するための重要な方法。

★解説

実験の場で生じる誤差は，ある確率分布(多くの場合は**正規分布**)に従って生じる偶然誤差と，時間的あるいは空間的に一定の方向(増加あるいは減少)に向かって系統的に生じる系統誤差に大別することができる。

ある化学反応工程で，収率向上の目的で新しく触媒Aを添加したほうがよいかどうか調べた。触媒の量を1％から増やすにつれて収率は向上し，3％ぐらいで一番高く，それ以上増やすと収率は減少することがわかったとしよう。もし，この実験が触媒の量の少ない水準から順次高い水準へと行われたとすると，どのようなことが起こるであろうか。慣れていない実験者であれば，はじめは実験条件の設定もうまくいかないかもしれない。また，この実験が何日にもわたって行われるとすると，気温，湿度などの天候条件の影響，あるいは反応釜などの実験設備の調子の変化が収率にも影響しているかもしれない。このように，実験順序とともに環境条件が系統的に変化すると，結果である特性値の変化が，とりあげた**因子**の水準の違いによるものなのか，環境条件などそれ以外の要因による影響なのか，明確に区別することができなくなる。このように2つ以上の要因がこみになって分離できなくなるのを，交絡という。

環境条件が系統的に変化することが避けられないときには，実験順序と因子の交絡を避けるため，実験順序を乱数表などを用いて無作為化して，系統誤差が各因子の組み合わせ条件下で公平に現れるようにすることが必要になる。これを系統誤差の偶然誤差への転化という。

関連ページ
正規分布　200
因子　20

実験順序	1	2	3	4	5	6
Aの水準	A_1	A_1	A_1	A_2	A_2	A_2
系統誤差	-3	-2	-1	$+1$	$+2$	$+3$

表1　規則的に実験を行ったときの系統誤差

実験順序	1	2	3	4	5	6
Aの水準	A_1	A_2	A_2	A_1	A_1	A_2
系統誤差	-3	-2	-1	$+1$	$+2$	$+3$

表2　無作為に実験を行ったときの系統誤差

■無作為化の目的　因子Aが収率に影響するかをみるために，これを2水準とって3回繰り返し，計6回の実験を行うことにした。このとき，実験を規則的にAの第1水準を3回続けて行い，つぎにAの第2水準を3回続けて行ったとしよう。

もし環境条件が6回の実験の間に系統的に変化すれば，それが収率にも影響する。その影響が表1のようにでたものとしよう。Aの第1水準はマイナスの誤差，第2水準はプラスの誤差をもっていることになる。この系統誤差を評価することができれば，この値を補正することによって，Aの水準の違いが収率にどのように影響するか，調べることができる。しかし,系統誤差の大きさを評価することは，その原因と結果の定量的関係が把握できていない場合には難しい。したがって，この場合にもしA_2がA_1よりも見かけ上収率の高いことがデータからいえたとしても，それはAの水準の違いによるものか，環境条件の違いによるものか明確でない。これに対処する1つの方法が，無作為化である。いま，A_1，A_2のカードを3枚ずつ用意し，計6枚を無作為に抜き取って，抜き取られた順に実験を行うとしよう。その時の実験順序が表2のようになったとする。このときには，実験順序の無作為化により，A_1とA_2の系統誤差が互いに相殺されて，A_1とA_2の比較がより正確に行えることになる。無作為化により，系統誤差の影響が完全になくなることは稀であるが，その可能性をより高めることは可能であり，その影響も，プラス，マイナスがランダムになり，偶然誤差の様子をおびてくる。

■無作為化の方法　実験順序を無作為化するには，乱数表，乱数サイ，カードによるクジ引きなどの方法がある。乱数表による方法を，上の例の場合について説明しよう。いま2桁の乱数を6個無作為に選んで，

67，12，08，47，95，28

が得られたとする。このとき，奇数ならばA_1，偶数ならばA_2の実験を行うことに前もって決めておけば，表2のような実験順序が得られる。　（宮村）

メリット・デメリット分析とマトリックス分析

ある方策のよい点(メリット)と悪い点(デメリット)を書き出して質的な判断の一助とする方法を，メリット・デメリット(M/DM)分析という。その方策の影響の強さを推定するときは，インパクト分析とよばれる。

いくつかの方策をいくつかの評価項目にわたって評価するときは，マトリックスの形で評価表を作るので，マトリックス分析と名づけられる。これには数種類がある。

このほか，チェック・リストや円形グラフなどによるパターン分析なども，質的データや要因を数量化せずに分析する際に利用される方法である。

★解説

■言葉による分析　メリット・デメリット分析は，メリットとデメリットをそれぞれ言葉で表現する。それを眺めてメリットがデメリットを上回るか否かを直観的に判断し，その方策の採否の決定の助けにする。このように，質的色彩の高い評価は，無理に数量化をしないで，言葉による分析を行う方がよいことも多い。M/DM 分析は，縦の線で左右に 2 分し，左側に M，右側に DM と標記して，その下に，思いつくままに M と DM とを書き並べればよい。

■±インパクト分析　図 1 にインパクトマトリックス表の書き方の例を示すが，ある技術やプロジェクトが完成すると，どのようなインパクトを与えるかを考えつくだけあげ，それらがどのような側面で，それぞれの対象にプラスの影響を与えるか，マイナスの影響を与えるか，記号で書き込んでいる。その強さも考慮しておくと，なお役に立つであろう。図 1 の例では○印は強い影響を表し，△は＋とも－ともいえないことを示している。

関連ページ

発想的 OR　294

番号	インパクト内容	経済	心理	健康	…	土着	移住	…
1	便利さ	レ				+	⊕	
2	騒音		レ			−	−	
3	危険			レ		⊖	−	
4	美観		レ			△	△	
⋮	⋮		⋮			⋮	⋮	

図1　新線建設の評価

方策＼評価	投手	打者	走者	野手
敬遠策	5	3	2	3
臭い球	1	5		1
勝　負	3	3		1
ランナーを刺す	4	1	4	2

図2　次打者への対策評価

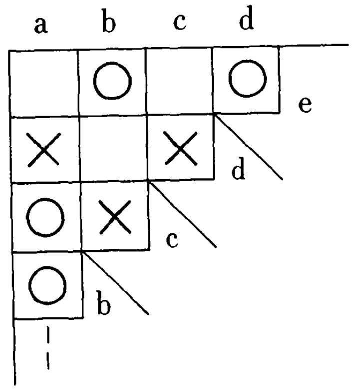


図3　3角マトリックス

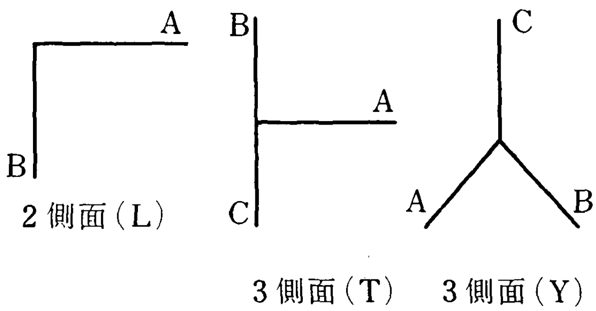


図4　マトリックス分析

■マトリックス分析　図2の例は，同点で迎えた最終回に2死から四球で走者を1塁に出したが，次打者が4番打者なのでバッテリーとしてはどういう対策を考えたらよいか，というような場合を例にとっている。方策と評価の2つの側面を縦横に配置し，考えられる方策ごとに，それぞれの面から5点法で評価をしている。

点数評価ができれば，項目ごとにウェートを定め，荷重平均を求めて，どの方策がよいか定めることが考えられる。

マトリックス分析は字義のとおりに，縦横に数字を並べたものが基本であるが，ある集団の相互関連を調べるためには，相互に関連に応じて○や×などの符号をつける方式も用いられている。ある集団が対象ならば3角形の分だけ書けばよいので，3角マトリックスといわれる(図3)。

そのほか，3つの集団の相互関連を，Aを中心にAとB，AとCと見たいときには図4の中央の配列をすればよいし，これにBとCの関連も見たいときには右側の配置にすればよい。このような配列のしかたについては，それぞれL型，T型，Y型などの名でよぶこともある。

もし総合的に3つの側面の関連を見ようとすると3次元配列になるが，そのときは適当な断面で調べるのがふつうである。参考書〔22〕　　　　(森村)

モンテカルロ法 (Monte Carlo method)

乱数(random number)を用いる手法を総称してモンテカルロ法という。フォン・ノイマン(von Neuman)とウラム(Ulam)による中性子のランダムな拡散現象のコンピュータ・シミュレーションに端を発している。このようなランダム事象のシミュレーションのほか，逆行列や固有値などの決定論的な問題を確率の問題としてとらえなおして，乱数を用いて数値解析を行うこともできる。

★解説

■乱数　0から9までの10個の数字のn個の並びが次の性質をもつとき，10進数の一様乱数という。①等確率性。各数字の出現頻度をk_i ($i=0, 1, \cdots, 9$)として，相対頻度k_i/nがnを大きくしていくと確率1/10に近づく。②無規則性。数字の並び方がでたらめであること。i番目にでた数字とj番目($i \neq j$)にでた数字に規則性がない。一様乱数からは，右ページにみるように任意の確率分布に従う乱数を発生できる。

■乱数の発生　簡単な問題を机上でシミュレーションする場合，サイコロ，乱数さい(正20面体のサイコロに0から9までの数字が2面にきざんである)や乱数表を用いる。本格的な問題には，コンピュータで算術的に発生させた凝似乱数(pseudorandom number)を用いればよい。一様乱数や正規乱数が簡単に利用できる。このほか，自然現象のランダム性を利用した物理乱数やπやeなどの超越数の各桁を用いる方法もあるが，実用的でない。

■利用分野　ランダム(確率的)な事象を含む対象のシミュレーションのすべてがモンテカルロ法の対象になる。在庫問題，待ち行列，交通システムなどの確率的な性格をもつシステムのシミュレーション，多重積分などの数値解析がある。待ち行列にはシミュレーション言語GPSSがある。

関連ページ
統計シミュレーション　256
シミュレーション　166
シミュレーション言語　168

■一様乱数の発生法　区間〔a, b〕上に一様に分布する乱数列の発生方法として，次の2通りがある。**平方採中法**は，フォン・ノイマンによって最初に提案された。たとえば，4桁の数字4567を最初の乱数とする。これを2乗して得られた8桁の数字20857489の中央の4桁8574を次の乱数とする。この方法は，すぐにあるサイクルで同じ数字が現れるので現在あまり用いられない。これに対し**合同法**とよばれる方法が提案されており，レーマー(D. H. Lehmer)により提案された乗算合同法が有名である。その方法は，乱数 x_n を λ 倍したものを m で割った余りを次の乱数 x_{n+1} とする。

$$x_{n+1} \equiv \lambda \cdot x_n \pmod{m}$$

たとえば，$\lambda=23$, $x_0=35$, $m=10^2$ とすれば $23x_0=805$ になり，余りの5が x_1 になる。10進2桁しかない計算機では3桁目の8は自動的に桁あふれするので便利である。ただし，次の混合合同法の方がより長い周期の乱数が得られる。

$$x_{n+1} = \lambda \cdot x_n + \mu \pmod{m}$$

■その他の乱数　指数分布，ポアソン分布，正規分布をする乱数列は，簡単に一様乱数から発生できる。正規乱数の発生方法の1つとして，中心極限定理を利用する方法がある。すなわち，平均値 μ，分散 σ^2 に従う分布の乱数列 $\{x_i\}$ の n 個の平均値 $\overline{x}_n=\sum_{i=1}^{n} x_i/n$ は，n が大きくなると正規分布 $N(\mu, \sigma^2/n)$ に収束する。一様乱数では，$\mu=1/2$, $\sigma^2=1/12$ であるので $n=12$ にとれば，正規分布 $N(1/2, 1/12^2)$ に収束する。ここで，$\overline{x}_n$ を次の z_n に変換すれば $N(0, 1)$ が求まる。

$$z_n=\frac{\overline{x}_n-1/2}{\sqrt{1/12^2}}=12\overline{x}_n-6$$

■任意の分布の乱数　任意の分布の累積分布関数を $y=F(x)$ とする。区間〔0, 1〕の一様乱数で0.5の値が得られた場合，これを y として x の値(1.5)を図から読みとればよい。

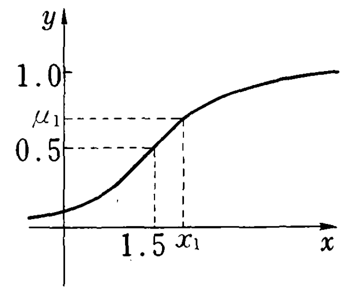


図1

連続分布であっても，x を適当に分割し，階段状の累積分布として近似すれば表として簡単に扱える。

■定積分の計算　区間〔a, b〕で，関数 $f(x)$ の定積分 $S=\int_a^b f(x)dx$ を一様乱数で計算する。

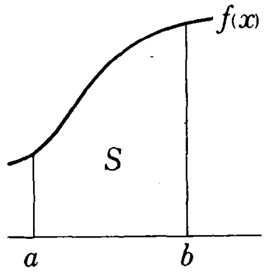


図2

〔0, 1〕の一様乱数 r は，$r'=a+(b-a)\cdot r$ で〔a, b〕の一様乱数になる。面積 S は n 個の関数値 $f(r_i')$ の平均値（S を平らにならした時の高さ）と底辺（$b-a$）の積になる。すなわち，

$$S=(\text{底辺})\times(\text{平均の高さ})$$
$$=(b-a)\left\{\sum_{i=1}^{n} f(r_i')/n\right\}$$

（新村）

尤度比方式による判別

群 G_1 の確率分布が与えられている時，G_1 のもとでの観測値 s の確率（条件付き確率）$P(s|G_1)$ を尤度という。G_1 と G_2 の 2 群の尤度の比 $P(s|G_1)/P(s|G_2)$ を尤度比という。この値が 1 より大きければ，G_1 群の確率の方が G_2 群より大きいので，s 値をとるデータは G_1 群に判別すればよい。このように，尤度比の大小で判別する方法を尤度比方式による判別という。説明変数が多次元正規分布に従う場合には，この比が 1 の値になる点は直線になり，フィッシャーの線形判別関数の判別境界点（尤度比方式による判別境界）になる。説明変数が質的データの場合には，ベイズ診断とも関係する。

★解説

■尤度と尤度比　G_1, G_2 とも s_1 から s_6 までの 6 個の離散点上の確率分布で与えられているとする。G_1 の分布で s_1 のとる確率 0.2 は，尤度 $P(s_1|G_1)$ を表す。同じく，$P(s_1|G_2)=0.1$ なので，この尤度比は 2 になる。同様にして，s_2 から s_6 までの尤度比が，1, 0.66, 1, 2, 0 になる。尤度比が 1 よりも大のものを G_1 群に，1 未満のものを G_2 群に判別すれば，s_1, s_2, s_4, s_5 が G_1 群に判別され，s_3 と s_6 が G_2 群に判別される。尤度比 1 のものを判定保留とすることも多い。

一般に判別境界点は 1 を用いるが，この値をいろいろな水準で動かして ROC 曲線を描いて評価すればよい。

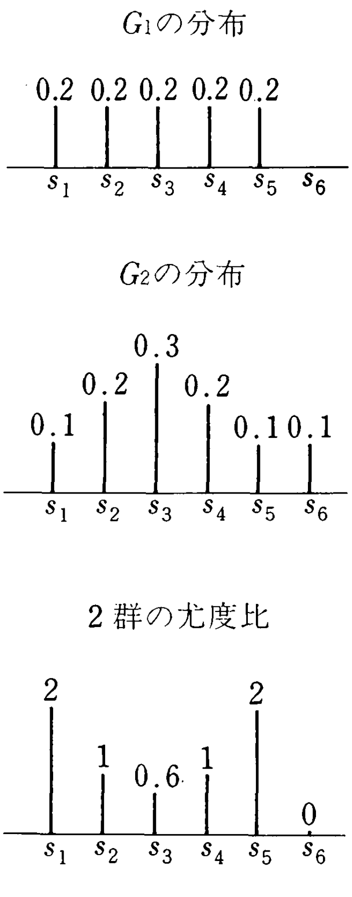


図1　離散確率分布

関連ページ

ベイズの定理を用いた判別　346

FUNCAT による判別　320

数量化Ⅱ類　190

ROC 曲線　14

決定行列　104

AID　32

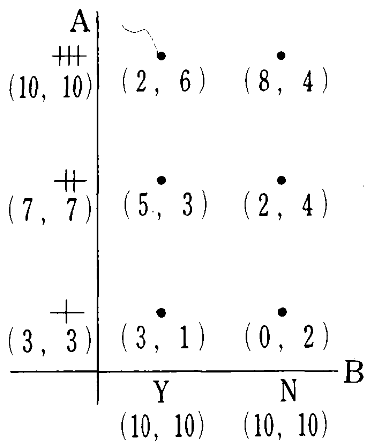


図2　データ例

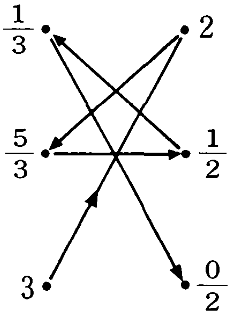


図3　尤度比による順位

閾値	検査	G_1		G_2		誤分類数
		定数	%	実数	%	
≧3	+Y	3	15	1	5	18
≧2	+++N	11	55	5	25	14
≧5/3	++Y	16	80	8	40	12
≧1/2	++N	18	90	12	60	14
≧1/3	+++Y	20	100	18	90	18
1/3>	+N	20	100	20	100	20

表1　尤度法の判別結果

■例題　検査A(+++, ++, +)と検査B(Y, N)が順序尺度と考えられる場合，図2のようにこれらの検査結果の組み合わせを平面上の点として配置できる。カッコの中の数字はG_1とG_2群の例数である。

両検査結果を単独に考えれば，いずれの検査結果でも両群の観測数が同じ(+++とYとNで10例ずつ，++で7例ずつ，+で3例ずつ)であり，この検査結果から判別できない。そこで，2検査の結果を同時に考えて尤度と尤度比を考える。たとえば，Aが+++でBがYの場合には，P(+++ Y$|G_1$)=2，P(+++ Y$|G_2$)=6なので，尤度比は1/3になる。この値を図3に示す。図中の矢印は，尤度比の大小順を示す。表1は，尤度比の判別境界点(閾値)を変えて集計したものである。たとえば，尤度比が3以上のものをG_1群，未満をG_2群とすれば，Aが+でBがYの検査結果がG_1群と判別される。この検査結果のうち，G_1群の3例が正しく判別され，G_2群の1例が間違ってG_1群と判別されたことになる。すなわち，誤分類数は18例になる。閾値を5/3にすれば，G_1群の4例とG_2群の8例，計12例が誤分類される。検査単独ではすべて等確率のため，1以上をG_1群とすれば，G_2群の20例が，1以下をG_2群とすればG_1群の20例が誤分類されることになる。以上から，尤度比方式による判別手法の成績のほうがよいことがわかる。

しかし，よい判別分析とは，説明変数の空間をできるだけ簡単な領域に分割し，その各領域をいずれかの群に割りふることである。図3の矢印や表1の検査順をみれば分かる通り，判別に規則性がみられない。検査BではYとNの繰り返しであり，検査Aでは+，+++，++の順序が次では逆転している。この意味で，尤度比方式による判別はデータに敏感に対応しすぎているといえる。FUNCATは，データに線形制約を課して，尤度比方式の無規則性を減少させている。　(新村)

輸送問題

いわゆる輸送問題は線形計画問題の特殊な型のもので，複数の供給地から複数の需要地への物資を運ぶ最適計画を求めようとする。一般の線形計画問題から区別されるのは，その構造の特殊性から，単体法よりもはるかに効率のよい解法が用意されているためである。

★解説

■ヒッチコックの輸送問題　最も古典的な問題で，一般の線形計画法に先立って1941年に発表されたという点でも名高い。m個の供給地(たとえば倉庫)のそれぞれにはs_iの供給量があり，n個の需要地(たとえば販売店)のそれぞれはd_jの需要がある。倉庫iから販売店jに単位量を輸送するのにc_{ij}の費用がかかる。総輸送費が最少になるように輸送計画を立てたい，というのがこの問題である。なお，この問題では，任意の倉庫と販売店の組み合わせに対して必ず輸送路があり，しかもその輸送路には輸送量の上限はないものとする。

■一般型輸送問題　上の問題で，任意のiとjの組に対し必ずしも輸送路は存在しなくてもよく，また中継地(たとえば配送デポ)を考慮してもよいものとする。もちろん，ここでは積み替えが許される。また，倉庫・販売店間だけでなく，倉庫・デポ間，デポ・販売店間の輸送可能なすべてのルートには，輸送費単価とともに輸送量の上限が与えられている。

■施設配置問題または固定費つき輸送問題　ヒッチコックの輸送問題では倉庫は定まっていた。しかし,販路の広がった場合など，新たに倉庫を用意することも含めて考えたいこともある。倉庫の候補地が複数個あって,そのうちの何個所かを開設した上での輸送費最少となる計画を作りたい。倉庫を作るには固定費がかかるので，表記のようによばれる。これは整数計画問題だが，輸送問題が利用される。

関連ページ
線形計画法　218
単体法　234
整数計画法　208

■輸送問題の解法　3カ所の倉庫から4販売店へ輸送する問題を例にして，ヒッチコックの輸送問題の解法を示す。表1の太枠内は倉庫iから販売店jへ単位量を輸送する輸送単価であり，枠の外には各倉庫の供給量及び各販売店の需要量が記されている。総供給量と総需要量は等しい。もし，等しくないときは架空の倉庫や販売店を考えれば，この形に帰着できる。

i ＼ j	1	2	3	4	供
1	3	2	2	6	8
2	4	1	5	2	15
3	1	4	2	3	9
需	5	8	7	12	32

表1　輸送単価と制約

■北西隅のルール　単体法のフェーズIに当たる最初の基底解を求めるため，北西隅(つまり左上の隅)から，需要もしくは供給量の制限いっぱい送るよう定めていく。①8と5の小さい方の値を左上の隅に書く。これは倉庫1から店1へ5単位送ることを意味する(表2)。②枠外の8と5から5を引いて，表2の点線部について，同様の割り当てを行う。

5				8
				15
				9
5	8	7	12	

表2

つまり，3が割り当てられる。③以下同様に進めて表3を得る。輸送量0のところは書いてない。これが始めの基底解で，この解を実行するときの総輸送費は表1の単価をかけて計算し，94となる。

5	3		
	5	7	3
			9

表3

■シンプレックス基準の計算　①基底変数に対応するc_{ij}を書き抜き，表4を作る。②枠外のu_i, v_jの値は$u_1=0$と置き，$u_i+v_j=c_{ij}$となるように定める。③表4では空白の個所で$\tilde{c}_{ij}=u_i+v_j-c_{ij}$の値を求める(表5)。

					u_i
	3	2			0
		1	5	2	−1
				3	0
v_j	3	2	6	3	

表4

■基底解の改善　$\tilde{c}_{ij}>0$の個所を基底変数に入れる。①表3の該当個所にθと書き，そこを出発点に縦横にθを増減して，縦横とも和に変化のないようにする(表6)。②輸送量の1つが0になるようθを定める。表6では7。③$\theta=7$として表6を書き直す(表7)。④このときの輸送費減少量は$4\times7=28$。

		4	−3
−2			
2	−2	4	

表5

5	3		
	5	$7-\theta$	$3+\theta$
		θ	$9-\theta$

表6

5	3		
	5		10
		7	2

表7

表7が第2の基底解であるから，これに表3以降の手続きを繰り返して更に基底解を改善する。結局，表8が最小解で，総輸送費は56となる。

	5	3	
	3		12
5		4	

表8

単体法と比べて見るとわかるように，輸送問題の特徴を生かし，計算はかなり簡単化されている。一般型問題の解法については，たとえば〔60〕(第2巻)。　(森村)

要因効果

因子の効果がその水準によって異なるとき主効果があるといい，また別の因子の水準がどのようなものであるかによって効果が異なるときにはこれらの因子間には交互作用があるという。この主効果と交互作用を総称して要因効果という。

実験を計画するにさいしては，とりあげる因子とともにこれらの因子間にどのような要因効果があるか，事前に十分検討することが実験を成功に導く重要な鍵となる。

★解説

関連ページ
因子　20

反応釜の収率向上の目的で，**因子**としてA(反応温度)，B(反応時間)の2つをとりあげて計4回の実験を行い，表1のデータが得られているとする。簡単にするため実験誤差は無視できるとする。このとき，A_1の効果は

因子	B_1	B_2
A_1	x_{11}	x_{12}
A_2	x_{21}	x_{22}

表1　データ

$x_{11}-x_{21}$ (B_1のとき)，$x_{12}-x_{22}$ (B_2のとき)

になるが，この平均の1/2をA_1の主効果といい，α_1で表すと，$\alpha_1=(x_{11}+x_{12}-x_{21}-x_{22})/4$ となる。同様にA_2，B_1，B_2の主効果α_2，β_1，β_2を定義できて，これらには $\alpha_1+\alpha_2=0$，$\beta_1+\beta_2=0$ の関係が成立する。

A_1とB_1の交互作用は，これらの組み合わせ効果であるから，全体の平均を $\mu=(x_{11}+x_{12}+x_{21}+x_{22})/4$ として，x_{11}のうちでA_1，B_1の主効果では説明できない項として，

$$(\alpha\beta)_{11}=x_{11}-(\mu+\alpha_1+\beta_1)=(x_{11}-x_{12}-x_{21}+x_{22})/4$$

と定義される。$(\alpha\beta)_{12}$，$(\alpha\beta)_{21}$，$(\alpha\beta)_{22}$も同様に定義されて，$(\alpha\beta)_{11}+(\alpha\beta)_{21}=0$，　$(\alpha\beta)_{11}+(\alpha\beta)_{12}=0$，　$(\alpha\beta)_{21}+(\alpha\beta)_{22}=0$，$(\alpha\beta)_{12}+(\alpha\beta)_{22}=0$ の関係が成り立つ。

要因効果を用いてデータを表すと，$x_{11}=\mu+\alpha_1+\beta_1+(\alpha\beta)_{11}$ となる。これをデータの**構造模型**とよぶ。実際には実験誤差ε_{11}を伴うから，$(\alpha\beta)_{11}$の検出には繰り返しが必要。

構造模型　110

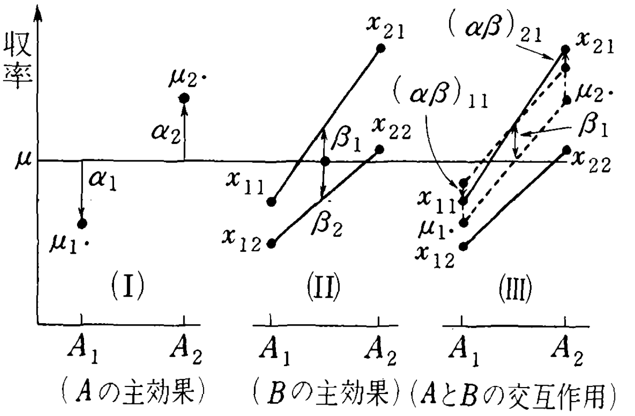


図1 主効果と交互作用

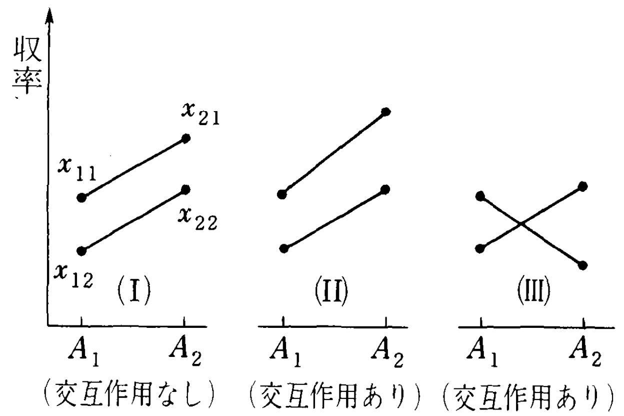


図2 交互作用の有無

■要因効果の図的表示 実験誤差 ε がないものとして前ページで定義した要因効果を，図に表示すると図1のようになる。図1の(I)は，$\mu_{1\cdot}$(A_1 における平均)，$\mu_{2\cdot}$ と μ (一般平均とよぶ)の差として定義される A の主効果 α_1, α_2 を示す。(II)には，$\mu_{\cdot 1}$, $\mu_{\cdot 2}$，と μ の差である B の主効果を示す。もし交互作用がなければ $\mu_{1\cdot}$, $\mu_{2\cdot}$ に，β_1 を加えたものが x_{11}, x_{21} になり，β_2 $(=-\beta_1)$ を加えたものが x_{12}, x_{22} になる。したがって，x_{11} と x_{21} を結ぶ線分と x_{12} と x_{22} を結ぶ線分は，互いに平行である。しかし，いまの例では交互作用が無視できないから，(III)に示されているようにこれら2つの線分は平行にならない。このことを用いて，得られたデータをグラフ表示することによって，図2のような形で交互作用の有無を判定することができる。

■交互作用の検出 一般には実験誤差 ε を伴うから，このことを考慮にいれて判断することが必要になる。たとえば，A_1B_1 条件下における収率のデータ x_{11} は，

$$x_{11}=\mu+\alpha_1+\beta_1+(\alpha\beta)_{11}+\varepsilon_{11}$$

となる。これをみると，交互作用と誤差の添字が同一であるから両者を区別することができない。これを交互作用と誤差が交絡しているという。交絡を防ぐためには，同じ条件下で少なくとも2回以上実験することが必要になる。A_1B_1 条件下での第1回，第2回目のデータを，それぞれ x_{111}, x_{112} とすると，これらは

$$x_{111}=\mu+\alpha_1+\beta_1+(\alpha\beta)_{11}+\varepsilon_{111}$$
$$x_{112}=\mu+\alpha_1+\beta_1+(\alpha\beta)_{11}+\varepsilon_{112}$$

となり，交互作用 $A\times B$ を検出することが可能になる。

■高次の交互作用 前項では2因子交互作用 $A\times B$ を考えたが，3つの因子 A, B, C をとりあげて実験を行うと，2因子交互作用として $A\times B$, $A\times C$, $B\times C$ の3つ，3因子交互作用として $A\times B\times C$ が考えられる。3因子以上の交互作用を高次の交互作用といい，これらの交互作用も2因子交互作用と同様に定義できるが，高次の交互作用は一般には大きくなく，技術的解釈が困難である場合が多い。 (宮村)

要因実験

特性値に影響すると思われる因子を2つ以上同時にとりあげて，それらの水準のすべての組み合わせについて行う実験。単一因子実験に比べると，同じ実験回数でより多くの情報を得ることができる。

★解説

単一因子実験を交互作用が存在する場合に適用すると，誤った結論を導く場合がある。この単一因子実験の欠点を克服するために考え出された実験法が要因実験である。

要因実験では，必要な**要因効果**すべてが推定可能になるから，正しい結論が得られることになる。いま，とりあげた因子がA，Bの2つで各因子とも2水準であるとする。要因実験ではすべての水準の組み合わせの実験を行うから，A_1B_1，A_1B_2，A_2B_1，A_2B_2の組み合わせで少なくとも計4回の実験を行うことになり，それぞれの組み合わせにおいてデータ(特性値) x_{11}，x_{12}，x_{21}，x_{22}が観測されることになる。もし実験誤差がなければ，Bの水準がB_1のときにAをA_1からA_2に変えたときの効果をx_{11}とx_{21}より，BがB_2のときのAの効果をx_{12}，x_{22}より求めることができる。したがって，相手の因子Bの水準が変わることによりAの効果がどのように変わるか，すなわち交互作用を推定することができる(実験誤差がある場合には，実験誤差と交互作用とを分離するために，同じ組み合わせ条件での実験の繰り返しが必要となる)。しかし，単一因子実験では一部のデータが欠けるためにAとBの交互作用を推定することができない。要因実験のデータの解析法は**分散分析**参照。

要因実験では，因子の数が増えるとその水準の組み合わせの数は加速度的に増加するので，そのすべての組み合わせの実験を行うことは難しくなる。そのため，1/2，1/4の組み合わせだけを実施する一部実施法という計画がある。この計画は，**直交表**を用いることにより容易に構成することができる。

関連ページ
単一因子実験　232
要因効果　394
因子　20
分散分析　336
直交表　238

■**要因実験と単一因子実験**　温度，反応時間を因子A，Bとし，これらの収率への影響を調べて最適条件を求めるため，つぎのような実験を計画した。

(1)　A，Bそれぞれの因子ごとに，他方の因子を固定して，1因子実験を行う(表1)。

(2)　A，Bすべての組み合わせについて実験を行う(表2)。

(1)，(2)の実験とも実験回数は18回で同じである。2つの方法を比較するため，各水準の組み合わせにおける真の収率が表3のように与えられるとする。(1)の単一因子実験では，はじめBの水準B_0の選び方によって結果が異なる。$B_0=B_1$とすると，表3よりAがA_3のとき収率は最大となるから，$A_0=A_3$としてBの水準を変えて実験を行うと，BがB_2のとき収率は最大になる。よってA_3B_2が最適条件になるが，表3よりA_2B_2が実は最適条件である。もしB_0としてB_2またはB_3を選んでいれば，最適条件A_2B_2に到達する。このように単一因子実験では，最初のBの選択の違いにより結論が異なるとともに，真の最適条件に到達しない可能性もある。ところが，(2)の要因実験では9個の組み合わせすべてで実験を行うことになるから，最適条件を見逃す危険性は少ない。すなわち，(2)の方法で実験を行うと，A(B)の効果へのB(A)の影響を検出することができる。また，(2)ではA_1について計6回の実験を行うことになるが，(1)では3回にすぎないから，(2)によって得られるデータは(1)の2倍の情報を含んでいることになる。

A_1 (80°)	A_2 (100°)	A_3 (120°)
○	○	○
○	○	○
○	○	○

(B_0に固定)

B_1 5分	B_2 10分	B_3 15分
○	○	○
○	○	○
○	○	○

(A_0に固定)

表1　2つの1因子実験

B \ A	A_1	A_2	A_3
B_1	○ ○	○ ○	○ ○
B_2	○ ○	○ ○	○ ○
B_3	○ ○	○ ○	○ ○

表2　2因子実験

B \ A	A_1	A_2	A_3
B_1	30	50	60
B_2	60	85	80
B_2	65	70	65

表3　A, B組み合わせにおける収率(%)

■**要因実験と一部実施法**　要因実験を行えば必要な要因効果をすべて推定可能になるが，因子の数が増えると水準の組み合わせが増えて実験回数が多くなる。たとえば，各因子とも2水準，因子の数が6ならば$2^6=64$，7ならば$2^7=128$の組み合わせがある。しかし，いくつかの因子間(とくに3つ以上の因子)の交互作用は検出する必要がないと判明している場合には，全体の実験の大きさの1/2あるいは1/4実施で十分な情報が得られることがある。これを一部実施法という。(宮村)

予測

将来のことに関しては，すべてについてその予測がありうるが，ビジネスの場では，技術予測，経済予測，需要予測などが主なものであろう。予測の結果は多かれ少なかれ意思決定に利用されるから，遠い将来の予測は，多少外れても修正のチャンスがあるので粗い精度で十分であろう。そして計画策定や注意の喚起といった点に重点が置かれるので，当てることは二の次であることがふつうである。これに対し，生産計画立案のための需要予測などでは，高い精度を要求され，当てることにウェートがかかりがちになる。

★解説

■技術予測　10年先にどういう技術が使いものになっているか，といった予測を指す。デルファイ法などが利用される。発展の道すじを予測するシナリオ書きの方法もある。

■経済予測　国や国際機関などによる経済予測は，国の施策や企業の長期計画立案などのために利用される。手法的には，計量経済分析(連立方程式モデル)，時系列分析，回帰分析，産業連関分析などが主なものである。

■需要予測　ある製品の比較的近い将来における需要量を予測するには，時系列分析や回帰分析などがよく用いられる。特に，市場が比較的安定していて，長期にわたる傾向や，季節，曜日などの影響を考慮すれば，従来から蓄積されている自社の販売実績から予測ができるような場合には，時系列分析が主流になろう。期間的にはやや長期を見通した予測で，しかも目標となる計画性を定めたいようなときには，相関分析やその他の経済予測の手法が用いられることも多い。いずれにせよ，予測の期間をできるだけ短くし，偶然変動の余地を少なくすることが望ましい。そのため，情報の早期収集システムを作ることが重要であろう。

関連ページ
発想的OR　294
移動平均法　18

■偶然変動と予測　予測が当たらないのは偶然変動に原因があるから，過去のデータから偶然変動を除くことが予測技術の第一歩である。このため，移動平均や指数平滑法などが用いられる。

独立に変動する偶然変動を加え合わせると，一般には**相殺効果**が働いて，個々の変動の和よりも少ない変動幅を示す。59ページに示すように，2つの確率変数が独立であるとき，それらの和の分散は，個々の分散の和に等しい。それで，もし，これらが同じ分布に従っていて，その分散が σ^2 であるならば，和の分散は $2\sigma^2$ である。同様に，n 個の和の分散は $n\sigma^2$ となり，標準偏差は $\sqrt{n}\,\sigma$ となる。$n=4$ のとき $\sqrt{4}=2$ であるから，標準偏差は2倍にしかならない。

移動平均は，n 個加えて n で割るから，平均後の系列の標準偏差は $\sigma/\sqrt{n}$ となる。

■平滑法　このように移動平均をとったデータの系列は，もとの系列に比べて偶然変動の幅は小さくなるので，一般に変化がおとなしく，滑らかになる。それで**平滑化**とよばれる。移動平均のように，適当な範囲で平均をとると，季節変動を消して平滑化した系列が得られる。

ある程度偶然変動が消えれば，平滑後に得られた系列を延長して，予測値が得られるから，平滑化はそのまま予測のための手段となる。

■フィルタ　フィルタを通して特定の周波数の波を取り出すことは，他の周波数の波を雑音と思えば，偶然変動の除去に相当する。したがって，フィルタの理論も予測の手法と見なせる。特にカルマンによるフィルタの理論は有効な方法の1つと考えられている。

■指数平滑法　移動平均は過去 m 個のデータに，重み $1/m$ を平等に掛けて加えている。しかし，その先に対しては重みは0となってしまっているわけで，突如としてある時点から先は無関係というのも不自然であろう。むしろ，1時点前は α，その前は α^2 というように，だんだん小さくなる重みを掛けて平均する方が自然である。この重みのかけ方から指数平滑法の名がつけられた。時点 n におけるデータを x_n とし，時点 $n-1$ までのデータから推定したそれを $\hat{x}_n$ で表すとき，

$$\hat{x}_{n+1}=(1-\alpha)\hat{x}_n+\alpha x_n$$

という関係式が成り立つ。このように，時点 n のデータとその推定値があれば時点 $n+1$ の推定値は得られるから，計算機で計算する際，x_n と $\hat{x}_n$ とを格納するメモリを用意しさえすれば，次々の時点の推定値は得られることになる。これはメモリの節約につながり，実用上有効な性質である。このようないくつかの理由から，特に需要予測を日常業務の中で行う場合には，指数平滑法がよく用いられる。 参考文献〔46〕

（森村）

ラテン方格法

n個の異なる数字1，2，…，nをn行n列の方形に並べて，各行各列にどの数字もちょうど1回ずつ現れるようにした方格を $n\times n$ のラテン方格といい，これを用いて割り付けを行う実験配置法。

★解説

乱塊法では1種類のブロック**因子**をとりあげたが，2種類のブロック因子をとりあげた実験配置法がラテン方格法である。

乱塊法の項では，A(反応温度)を制御因子(3水準),日をブロック因子(3水準)として，表1の実験順序で実験を行った。この実験では，1日の中で比較したいAの3水準A_1，A_2，A_3が1回ずつ実験されていて，日の影響はこれらに対して平等であるが，1日の中での順序に関しては，A_2水準は3日間のうち2日間は第1回目に実験が行われているが，A_3は1度も第1回目には行われていないので，平等ではない。しかし，1日の中での実験順序が結果に影響を与えるのであれば，この実験の方法では不公平になる。

この1日の中の実験順序についてもバランスをとる実験法がラテン方格法である。このためには，各行各列に1，2，3の数字がちょうど1回ずつ現れる 3×3 ラテン方格を用いて，行に日を，列に1日の中での実験順序を，方格の中の数字にAの水準を対応させて，表2のように割り付ければよい。

関連ページ
乱塊法　402
因子　20

日＼順序	1	2	3
第1日	A_2	A_1	A_3
第2日	A_2	A_3	A_1
第3日	A_1	A_3	A_2

表1　乱塊法での実験順序

日＼順序	1	2	3
第1日	A_1	A_2	A_3
第2日	A_2	A_3	A_1
第3日	A_3	A_1	A_2

表2　ラテン方格法での実験順序

1	2	3	4
2	1	4	3
3	4	2	1
4	3	1	2

1	2	3	4
2	3	4	1
3	4	1	2
4	1	2	3

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

表 3　4×4 標準ラテン方格

■ラテン方格　2×2 のラテン方格は 2 個，3×3 のラテン方格は 12 個，4×4 のラテン方格は 576 個…あることがわかっているが，それらのうちで第 1 行と第 1 列に 1，2，3，……の数字がこの順番に並んでいる方格を標準方格という。4×4 のラテン方格で標準方格を表 3 に示す。なおラテン方格法では，ラテン方格を用いて割り付けを行うので，2 種類のブロックの大きさは等しくなければならないという制約がある。

■無作為化の方法　ラテン方格の配置を無作為に選ぶには，数多くのラテン方格から 1 つを無作為に選ぶのが原則であるが，実際には 1 つの標準方格を適当に選び，その行および列の順序を無作為に並べ替えればよい。

■ラテン方格法の実験データの分散分析　因子 A の l 個の水準 $A_1, A_2, \cdots, A_l$ を選び，各水準の比較を $l\times l$ のラテン方格を用いて実験を行ったとする。2 種類のブロック因子を B（行），C（列）で示すと，得られるデータは表 4 のようになる。表では A の水準が明記されていないが，これらのデータから，

$$T_{\cdot\cdot(k)}=\text{水準 } A_l \text{ のデータの和}$$

B \ C	C_1	C_2	……	C_l	和	平均
B_1	x_{11}	x_{12}	……	x_{1l}	$T_{1\cdot}$	$\bar{x}_{1\cdot}$
B_2	x_{21}	x_{22}	……	x_{2l}	$T_{2\cdot}$	$\bar{x}_{2\cdot}$
⋮	⋮	⋮		⋮	⋮	⋮
B_l	x_{l1}	x_{l2}	……	x_{ll}	$T_{l\cdot}$	$\bar{x}_{l\cdot}$
和	$T_{\cdot 1}$	$T_{\cdot 2}$	……	$T_{\cdot l}$	$T_{\cdot\cdot}$	
平均	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	……	$\bar{x}_{\cdot l}$		$\bar{x}_{\cdot\cdot}$

表 4　ラテン方格法による実験データ

$$\bar{x}_{\cdot\cdot(k)}=T_{\cdot\cdot(k)}/l$$

を計算できる。このときのデータの構造模型は，

$$x_{ij}=\mu+\alpha_i+\beta_j+\gamma_k+\varepsilon_{ij}$$

ここで，

μ：一般平均

α_i：B_i の効果，$\alpha_1+\cdots+\alpha_l=0$

β_j：C_j の効果，$\beta_1+\cdots+\beta_l=0$

γ_k：A_k の効果，$\gamma_1+\cdots+\gamma_l=0$

となる。これらの要因 A，B，C，E（誤差）の平方和 S_A，S_B，S_C，S_E，自由度 ϕ_A，ϕ_B，ϕ_C，ϕ_E は，

$$CT=T_{\cdot\cdot}{}^2/l^2$$
$$S_T=(x_{11}{}^2+x_{12}{}^2+\cdots+x_{ll}{}^2)-CT$$
$$S_A=(T_{\cdot\cdot}{}^2{}_{(1)}+\cdots+T_{\cdot\cdot}{}^2{}_{(l)})/l-CT$$
$$S_B=(T_{1\cdot}{}^2+\cdots+T_{l\cdot}{}^2)/l-CT$$
$$S_C=(T_{\cdot 1}{}^2+\cdots+T_{\cdot l}{}^2)/l-CT$$
$$S_E=S_T-(S_A+S_B+S_C)$$
$$\phi_A=\phi_B=\phi_C=l-1,\ \phi_E=(l-1)(l-2)$$

となる。さらに平方和を自由度で除して各要因の不偏分散を求め，これを誤差要因の不偏分散と比較することにより，要因効果の有無を判定することができる。詳しくは分散分析の項 （→336 ページ）参照。　（宮村）

乱塊法

実験の場全体をいくつかのできるだけ均一なブロックに分け，各ブロック内での実験順序を無作為化して行う実験。これにより，実験誤差からブロック効果を分離できるため，実験精度がよくなる。

★解説

ある化学反応工程の収率向上のため，**因子**として反応温度(A)をとりあげ，最適条件を得るため3水準A_1(80°)，A_2(100°)，A_3(120°)を選び，繰り返し3回,計9回の実験を計画した。ところが，1日の実験は3回しか行えないこと，および天候などの環境条件などの違いによる影響があるかもしれないため，計9回の実験の場全体を**無作為化**するかわりに，1日の中でのA_1，A_2，A_3の3回の実験の順序を無作為化してこれを3日間にわたって行った。

もし日間に大きな違いがあるのに，9回の実験全体を無作為化すると日間の変動が実験誤差に含まれることになり，A_1，A_2，A_3水準間の比較の精度が悪くなる。これを防ぐために，実験全体の場をいくつかの均一な場(ブロックとよぶ)にわけ，各ブロック内で比較したい，ひとそろいの処理A_1，A_2，A_3の実験を行い，その実験順序を無作為化する。ここで各ブロック内ではA_1，A_2，A_3の3水準の実験が行われているから，これを1つの因子(ブロック因子とよぶ)とみることもできる。このようにブロック因子を1つ導入して，フィッシャーの3原則の1つである局所管理を実現するのが乱塊法である。

ブロック因子は，因子の各水準間の比較をして最良の水準を選ぶのではなく，実験誤差よりブロック効果を分離して，実験の精度をあげるために導入した因子である。ブロック因子としては，日，装置，原料ロット，作業者，地区などがあり，これを導入した乱塊法はよく用いられる。

関連ページ
因子　20
無作為化　384
実験計画法　160

第1日(B_1)	第2日(B_2)	第3日(B_3)
A_2 A_1 A_3	A_2 A_3 A_1	A_1 A_3 A_2

表 1 乱塊法の実験順序

順序	1	2	3	4	5	6	7	8	9
水準	A_2	A_1	A_1	A_3	A_2	A_3	A_1	A_3	A_2

表 2 完全無作為化による実験順序

A \ B	B_1	B_2	……	B_m	和	平均
A_1	x_{11}	x_{12}	……	x_{1m}	$T_{1\cdot}$	$\bar{x}_{1\cdot}$
A_2	x_{21}	x_{22}	……	x_{2m}	$T_{2\cdot}$	$\bar{x}_{2\cdot}$
⋮	⋮	⋮		⋮	⋮	⋮
A_l	x_{l1}	x_{l2}	……	x_{lm}	$T_{l\cdot}$	$\bar{x}_{l\cdot}$
和	$T_{\cdot 1}$	$T_{\cdot 2}$	……	$T_{\cdot m}$	$T_{\cdot\cdot}$	
平均	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	……	$\bar{x}_{\cdot m}$		$\bar{x}_{\cdot\cdot}$

表 3 乱塊法による実験データ

■実験のやり方 前ページの乱塊法の実験をするためには，まず第1日目，第2日目，第3日目の A_1, A_2, A_3 の実験順序を乱数あるいはクジを用いて無作為化しなければならない。このことにより，日内の実験順序による影響を偶然誤差と考えることができる。実験順序を無作為化すると，たとえば表1のようになる。ここで，日をブロック因子として，記号 B で示し，その水準を B_1, B_2, B_3 と書く。表1と表2を比較して，乱塊法と実験の場全体を無作為化する完全無作為化法の相違に注意しておこう。

■乱塊法の実験データの分散分析 この場合の実験データは表3のようになる。このとき，実験データの構造模型は，

$$x_{ij}=(\text{一般平均})+(A_i\ \text{水準の効果})+(\text{ブロック}\ B_j\ \text{の効果})+(\text{誤差})$$
$$=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$$

ここで

$$\alpha_1+\alpha_2+\cdots+\alpha_l=0$$
$$\beta_1+\beta_2+\cdots+\beta_m=0$$

となる。ブロック因子 B は実験精度を上げるために導入したものであるから，これと制御因子 B の交互作用を考えても意味がない。この実験データからは，

(1) A の水準間に差があるかどうか，

(2) ブロック(B)間に差があるかどうか，

の検定を行うことができる。これを分散分析表にまとめると表4になる。ここで，

$$CT=T_{\cdot\cdot}^2/lm\ (T_{\cdot\cdot}=\text{全データの和})$$
$$S_T=(x_{11}^2+x_{12}^2+\cdots+x_{lm}^2)-CT$$
$$S_A=(T_{1\cdot}^2+T_{2\cdot}^2+\cdots+T_{l\cdot}^2)/m-CT$$
$$S_B=(T_{\cdot 1}^2+T_{\cdot 2}^2+\cdots+T_{\cdot m}^2/l-CT$$
$$S_E=S_T-(S_A+S_B)$$

これより，有意水準 α で

$$V_A/V_E>F^{\phi_A}_{\phi_E}(\alpha)\Rightarrow A\ \text{は有意}$$
$$V_B/V_E>F^{\phi_B}_{\phi_E}(\alpha)\Rightarrow B\ \text{は有意}$$

と判断する。 (宮村)

要因	平方和	自由度	不偏分散	不偏分散比
A	S_A	$\phi_A=l-1$	$V_A=S_A/\phi_A$	V_A/V_E
B(ブロック)	S_B	$\phi_B=m-1$	$V_B=S_B/\phi_B$	V_B/V_E
E(誤差)	S_E	$\phi_E=(l-1)\times(m-1)$	$V_E=S_E/\phi_E$	
計	S_T	$\phi_T=lm-1$		

表 4 分散分析表

離散分布

ある期間に発生した計算機の故障回数をXで表すと，この確率変数Xのとりうる値は 0, 1, 2, … という**離散的**な値である。このような離散的な値をとる確率変数の分布を総称して**離散分布**という。その代表的なものに**二項分布**，**ポアソン分布**がある。

★解説

二項分布，ポアソン分布については別に項目があるので，それ以外の離散分布について紹介する。

■**単位分布**　Xが一定値aしかとりえないとき，その分布を**単位分布**とか**一定分布**とよぶ。つまり，$\mathrm{P}\{x=a\}=1$。

■**一様分布**　Xが 1, 2, …, n の値を同じ確率でとるとき，その分布を$(1, 2, \cdots, n)$上の**一様分布**とよぶ。つまり，

$$\mathrm{P}\{x=k\}=1/n \qquad (k=1,\ 2,\ \cdots,\ n)$$

で，平均値は $(n+1)/2$，分散が $(n^2-1)/12$ である。さいころを投げたとき，でる目の数の分布がこれにあたる。

■**幾何分布と負の二項分布**　1 回の**試行**において事象Eが生起する確率がpである**ベルヌイ試行**を考える。このとき，Eがr回生起するまでにかかる試行回数S_rの分布は，

$$\mathrm{P}\{S_r=k\}=\binom{-r}{k-r}(-q)^{k-r}p^r \quad (k=r,\ r+1,\ \cdots;\ q=1-p)$$

となり，**負の二項分布**とよばれる。その平均は r/p，分散は rp/p^2 である。$\binom{-r}{m}$ は $(1+x)^{-r}$ を展開したときのx^n項の係数で，負の二項係数という。$r=1$ のときは，

$$\mathrm{P}\{S_1=k\}=q^{k-1}p \qquad (k=1,\ 2,\ 3,\ \cdots)$$

と簡単になる。これを**幾何分布**とよぶ。

幾何分布の特徴は，**指数分布**と同じくメモリレスということである。図 1 で見るとおり，S_rは幾何分布に従うr個の互いに独立な確率変数 X_i $(i=1,\ 2,\ \cdots,\ r)$ の和として表される。幾何分布が連続的な場合の指数分布に対応するように，負の二項分布は**アーラン分布**と対応する。

関連ページ
確率変数　56
確率分布　56
二項分布　276
ポアソン分布　352
平均値　344
分散　332
試行　56
ベルヌイ試行　276
指数分布　156
アーラン分布　66

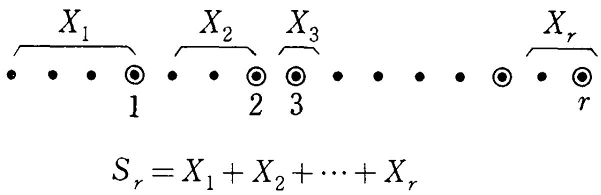


図1　幾何分布と負の二項分布の関係
（⊙印は事象Eの発生した試行，各X_iは幾何分布に従う）

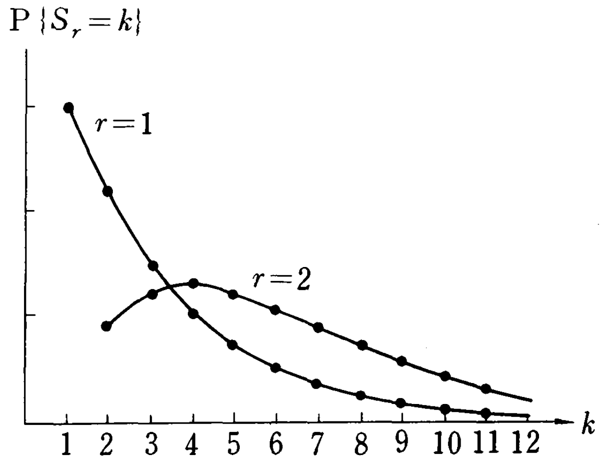


図2　負の二項分布

k	0	1	2	3	4	5	6	7以上
超幾何分布	.330	.408	.202	.052	.008	.001	.000	—
二項分布	.349	.387	.194	.057	.011	.002	.000	—

表1　二項分布による超幾何分布の近似例（$N=100$, $M=10$, $n=10$のとき）

■超幾何分布　全体でN個からなる製品ロットの中に，M個の不良品が混じっているとする。この中からn個をランダムに選んで検査したときに見つかる不良品の数Xの分布は，

$$P\{X=k\} = \binom{M}{k}\binom{N-M}{n-k} \Big/ \binom{N}{n}$$

$$(k=0,\ 1,\ 2,\ \cdots,\ \min(n,\ M))$$

となり，これを**超幾何分布**とよぶ。平均値はnp，分散は$np(1-p)(N-n)/(N-1)$である。ただし，pは全体の中での不良品の割合で，$p=M/N$である。

上の例では不良品総数Mが既知であるとして確率を計算した。実際は，上のような実験の結果からMを推定したい。そのとき，この確率モデルを用いて**最尤法**(→134ページ)によりMを推定することができる。

この例のようなサンプルの抜き取り方を**非復元抽出**(→54ページ)という。これに対し，1個とって良・不良を調べ，それを元に戻してまたランダムに1個とる，これをn回繰り返す場合もある。このやり方を**復元抽出**という。復元抽出のとき，n個の中に含まれる不良品数をX'とすると，その分布は$p=M/N$とした二項分布になる。Nが大きいとき，復元抽出でも同じものが2度も抜き取られる可能性は小さいので，非復元抽出と同じと考えてよいだろう。実際$N>10n$のとき，超幾何分布は二項分布でよく近似される（表1）。　（森）

利回り法と差額投資

利回り法は，投資案の利回り r（→268ページ）と計算利率 i の値を比較し，有利さを判断する方法である。すなわち，$r>i$ なら有利，$r<i$ なら不利となる。

差額投資とは，代替的な2つの投資案の差のキャッシュ・フローでとらえられるものである。追加投資ともいう。

代替的（相互に排反的）な投資案を利回りで比較する場合は，差額投資の利回りに着目しなければならない。

★解説

投資案の**利回り**は，投資の価値とリターンの価値が等しくなる利子率のことである。たとえば，表1のA案をみてみよう。この案の利回りを r とすると，100万円投資することにより1年後には $100\times(1+r)$ 万円の果実が得られ，それが120万円になっているのだから，

$$100\times(1+r)=120$$

となる。この式から r の値は，$r=20\%$ となり，$i=10\%$ では利回りの方が大きいので，この案はペイする。ちなみに**正味終価**を求めると，次式に示すように，1年後には10万円の正味の得が得られる。

$$S=120-100\times(1+0.1)=10 \ (\text{万円})$$

次に表1のB案についてみると，利回りは18%になる。ここでA案の利回りがB案の利回りより大きいので，A案の方が有利と判断してはいけない（B案の正味終価は16万円になる）。差額投資（B−A）をみると，投資が100万円，リターンが116万円で利回りは $16\%>10\%$ となり，A案に（B−A）案を追加する案（B案）が有利になる。

案	投　資	1年後のリターン
A	100万円	120万円
B	200万円	236万円

表1　代替的投資案

関連ページ
利回り　268
投資案の経済性指標　268
正味終価　268
効率指標の使い方　116
追加効率指標の使い方　240

案	投　資	リターン	寿命
A B	1000万円 2000万円	400万円 700万円	5年 5年

表2　2つの投資案

表2に示すような2つの投資案AとBがある。Aは1000万円の投資で年間の人件費400万円を削減できるもの，Bはさらに投資を1000万円増して合計2000万円投資すると，マテハンの費用が年間300万円削減でき，合計700万円の効果があがるものである。

利回りによる比較は，次のようにして行うことになる。。

(1)　投資額の小さい順に案を並べる。

ここでは表2がそれにあたる。

(2)　差額投資を求める。

差額投資	投　資	リターン	寿命
0→A	1000	400	5
A→B	1000	300	5

この表で0→Aは，何もしない案(0)とAの差額投資を意味している。

(3)　差額投資の利回りの値を求める。

［0→Aの利回り］

$1000[P\to M]_5^r=400$

より，

$$[P\to M]_5^r=\frac{400}{1000}=0.4$$

資本回収係数の数表から，利回りはおよそ29%になることがわかる。

［A→Bの利回り］

$1000[P\to M]_5^r=3000$

より

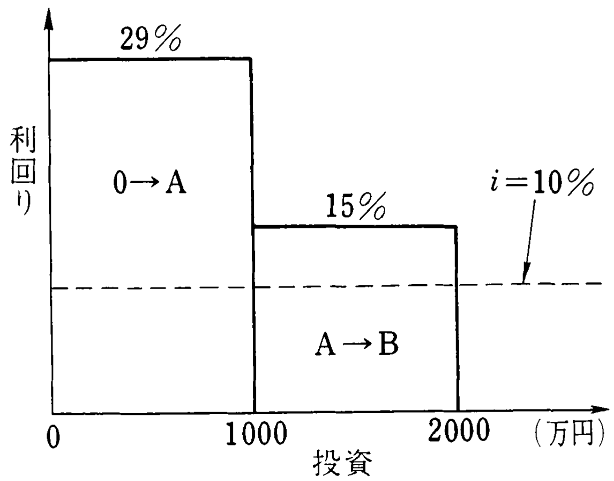


図1　利回りによる比較

$$[P\to M]_5^r=\frac{300}{1000}=0.3$$

同様に，利回りはおよそ15%になる。

差額投資	利 回 り
0→A	29%
A→B	15%

(4)　求めた利回りを計算利率と比較して判断する。

たとえば $i=10\%$ のとき，差額投資0→Aは有利，同様に A→Bも有利，したがって，0からAへ進み，さらにAからBへ進むこと，すなわちB案が有利な案になることがわかる(図1参照)。

B案の固有の利回り22%とA案の29%を比較すると，A案が有利になってしまい，誤りをおかすことになる。

利回りは初期投資の効率を示す指標である。一般に，効率指標は独立案の選択に使うもので(→116ページ)，上記のような代替案(排反案)の比較には差額投資を用いなければならない。(中村)

流体近似と拡散近似

到着率がサービス能力を上回るか，もしくはほぼ等しい場合に待ち行列の挙動を表現する近似的なモデル。平均値だけに注目し，待ち行列の変動は確定的であるとみなすのが流体近似，それに分散の影響まで考慮し確率的変動をするとみなすのが拡散近似である。当然のことながら，流体近似はかなり簡単で，拡散近似の方がずっと難しくなるが精確さは増し，到着率が低くてもよい近似を与える。

★解説

■離散と連続　待ち行列モデルでは，客は1人1人区別できることを念頭に置いている。このため，客の到着時点も1つ1つ区別され，待ち行列長や系内数は非負の整数値しかとらない。しかし，到着率がサービス能力を上回る，いわゆるラッシュ・アワーのときには，待ち行列の長さは一般に長くなるため，1人ぐらいの違いはあまり問題にならなくなる。

　また，客の到着や退去の時点も頻繁に現れるため，連続的に変化が起こると考えても差しつかえない。このため，変化時点も変化量もごくわずかずつ変化するとして得られる連続型のモデルで待ち行列を近似することが可能になる。

■流体近似　トラヒック密度（→ 374ページ）が1を超えるとき，到着とサービスの割合がともに確定的であるとみなし，流体モデルで近似する。到着率を，時々刻々の平均到着数とするか，それを幾分大き目に見積もって，369ページの所論に類似の議論をすればよい。

■拡散近似　待ち行列の窓口が暇になることがなければ，n人目の客の待ち時間は独立で同一分布にしたがう確率変数の和に等しい。これは中心極限定理によって近似的に正規分布にしたがう。この事実を巧みに利用して拡散近似が得られる。数学的議論はかなり難解である。

関連ページ
待ち行列　366

待ち行列とラッシュ・アワーの問題　368

中心極限定理　226

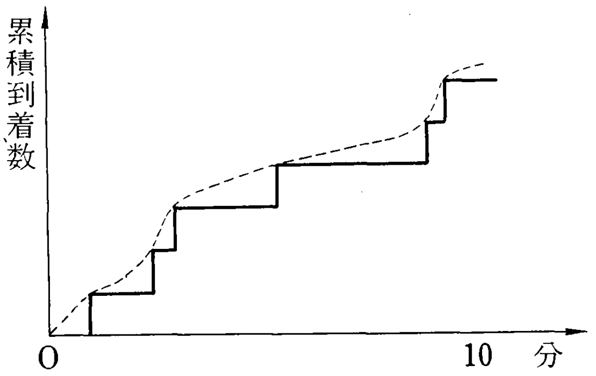


図1　累積到着数と連続近似

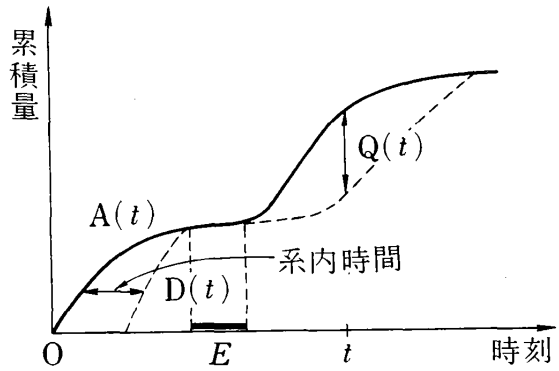


図2　累積数による待ちの表現

■累積到着数の図示　待ち行列は客の到着とサービスの時間的変動によって生じたり消滅したりするから，この両者をなんらかの方法で表現することが大切である。ふつうの待ち行列の理論では，それらを数学的に表現する。しかし多くの実用上の問題では，それらを図示することによって直観的処理をすると効果が大きい。この方法は，流体近似を考える基礎でもあるが，流体近似ということを離れても，それ自体で有効である。

客の到着順に到着時刻を時間軸上にプロットし，それらの時点で1つずつ上方にジャンプする階段関数をかくと図1のようになる。この階段関数の縦座標は，その時点までに到着した客の累積数を表している。

図1の状況よりも，もっと頻繁に客が到着すると，階段関数は滑らかな曲線に近くなる。

■待ちの表現　累積到着数を連続関数$A(t)$で近似し，同様に，サービスを終わって退去する累積数$D(t)$を図2のように並べてかくと，この2つの曲線の間の縦方向の差$Q(t)$が系内数を，横方向の差が系内時間を示している。

この図をかく上で注意しなければならないことは，$D(t)$は$A(t)$を追い越して上に出ることはないという点である。図の区間Eの間，系内数は0で，サービスをする客がいないためである。

累積到着数の曲線は，実績値についてかくのは簡単であるが，本来は確率的な変動をするものであるから，待ち時間の分布を求めたりしたいと考えるときには，この方法は向いていない。しかし，大まかな予測に基づいて，$A(t)$を仮定し，ある方策に基づいて$D(t)$の見当をつけるならば，$A(t)$が時間的に変化する場合も容易に扱えるという利点もある。ふつうの待ち行列のような確率論的扱いをこのような場合に行うのはとても大変で，実用的な立場からは，この方法が優れている。

■定性的な解析　$A(t)$の予測も大まかなので，ここの方法は，定量的に正確な計算をするというより，いろいろな方策のよさの見当をつける，定性的な結論を得るのに適している。（森村）

連続分布

ヒトの身長やある装置の寿命の長さは，ある範囲の値を連続的にとりうる。このように，連続的に値をとる確率変数の分布を総称して**連続分布**という。代表的なものに，正規分布や指数分布がある。

★解説

正規分布，**指数分布**，**ガンマ分布**，**ワイブル分布**，および **χ^2 分布**などの**標本分布**については別に項があるので，ここではそれ以外の連続分布を紹介する。

■ピアソン（Pearson）型の分布　密度関数 $f(x)$ が適当なパラメータ a，c_0，c_1，c_2 に対して，

$$\frac{d}{dx}\log f(x)=-\frac{a+x}{c_0+c_1x+c_2x^2}$$

を満たすとき，その分布を**ピアソン型の分布**という。a や c_i の値によって，$f(x)$ はいろいろな形をとる。たとえば，$c_1=c_2=0$ のときは正規分布 $\mathrm{N}(-a,\ c_0)$ となる。また，$(c_0+c_1x+c_2x^2)=0$ の根のタイプで密度関数 $f(x)$ の形を分類できる。$c_2=0$，$c_1\fallingdotseq 0$ のときはガンマ分布となり，タイプⅢと称せられる。以下に紹介するベータ分布や逆ガウス分布，t 分布などもピアソン型の分布である。

■対数正規分布　$\log X$ が正規分布 $\mathrm{N}(\mu,\ \sigma^2)$ に従うとき，X の分布を**対数正規分布**とよぶ。その密度関数は，

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma x}e^{-\frac{(\log x-\mu)^2}{2\sigma^2}}\qquad (x>0)$$

となる（図1参照）。平均は $e^{\mu+\frac{\sigma^2}{2}}$，分散は $e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$ である。別の書き方をすると，Z が $\mathrm{N}(\mu,\ \sigma^2)$ に従うとき，$X=e^Z$ の分布が対数正規となるわけであるが，Z が少し変化しても X の値は大きく変わりうることがわかるだろう。この分布はバラツキの大きい寿命時間の分布によくあてはまり，信頼性解析に利用される。また，所得分布などの経済関係や生態系における集団のサイズの分布などにもよく用いられる。

関連ページ
確率分布　56
正規分布　200
指数分布　156
ワイブル分布　418
ガンマ分布　66
χ^2 分布　314
標本分布　314

信頼性　184

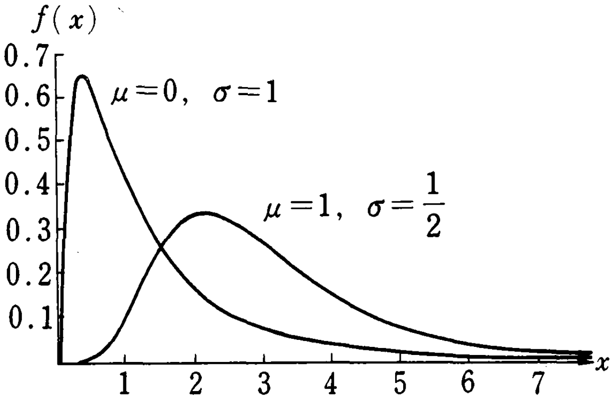


図 1　対数正規分布の密度関数

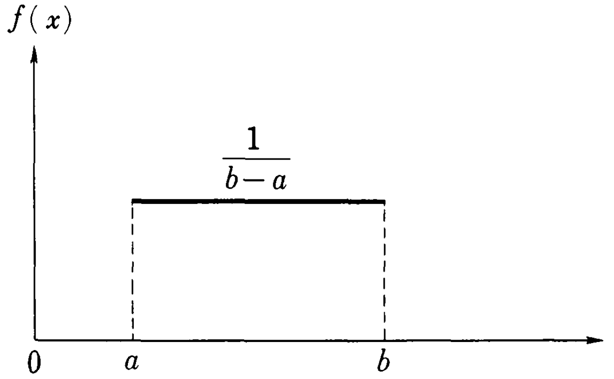


図 2　一様分布の密度関数

■**一様分布**　密度関数が

$$f(x)=1/(b-a) \qquad (a \leqq x \leqq b)$$

で与えられるとき，これを区間〔a, b〕上の**一様分布**あるいは**矩形分布**とよぶ。平均は $(b-a)/2$, 分散は $(b-a)^2/12$ である。

■**ベータ分布**　密度関数が

$$f(x)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

$$\alpha,\ \beta>0 \ (0 \leqq x \leqq 1)$$

で与えられる分布を**ベータ分布**という。平均は $\alpha/(\alpha+\beta)$ である。PERT (→292 ページ) における各作業の所要時間の分布などに用いられる。

■**パレート (Pareto) 分布**　分布関数が次のように与えられる分布をいう。

$$F(x)=1-(k/x)^{\alpha} \qquad (x \geqq k)$$

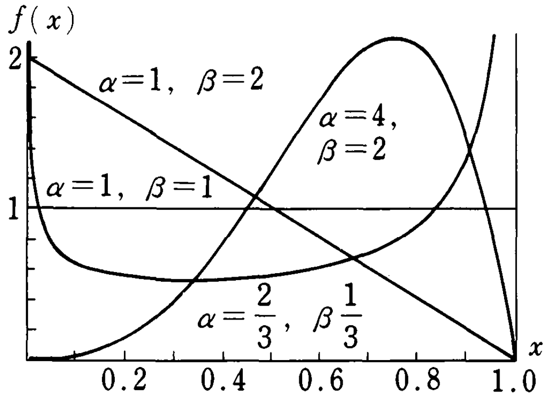


図 3　ベータ分布の密度関数

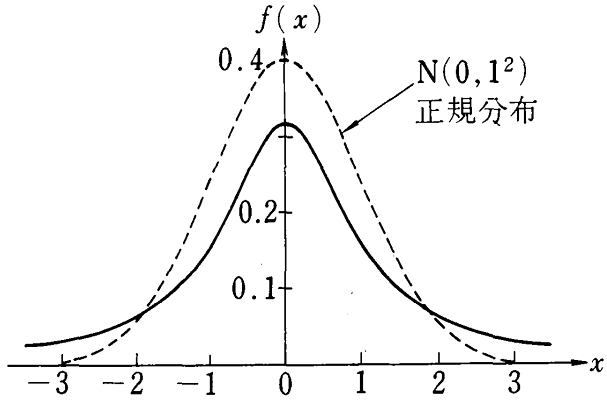


図 4　コーシー分布の密度関数

ここで $k, \alpha>0$ である。$\alpha>1$ のとき，平均は $\alpha k/(\alpha-1)$, 分散は $\alpha k^2/(\alpha-1)^2(\alpha-2)$ である。パレート (1848−1923) が，所得 x 以下の人の割合がこれによくあてはまる，として提唱した。この離散型を $Zipf$ 分布という。

■**コーシー (Cauchy) 分布**　X, Y が独立でともに正規分布 N(0, 1) に従うとき，X/Y の確率密度関数は，

$$f(x)=\frac{1}{\pi(1+x^2)} \qquad (-\infty<x<\infty)$$

となる。これは原点に関して対称であるが，平均値は存在しない。自由度 1 の **t 分布**と同じ。図 4 で見るように正規分布と比べ裾野が広い。各 X_1, X_2 が独立でコーシー分布に従うとき，その和もコーシー分布となる。

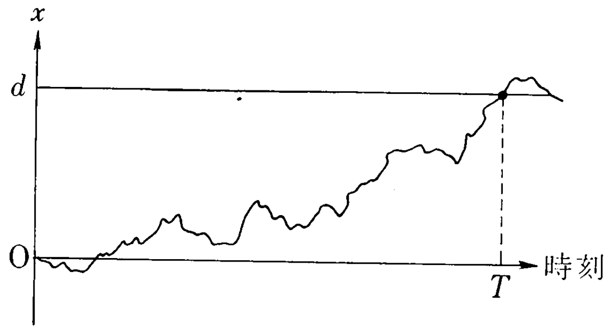


図5　一次元ブラウン運動と初通過時間T

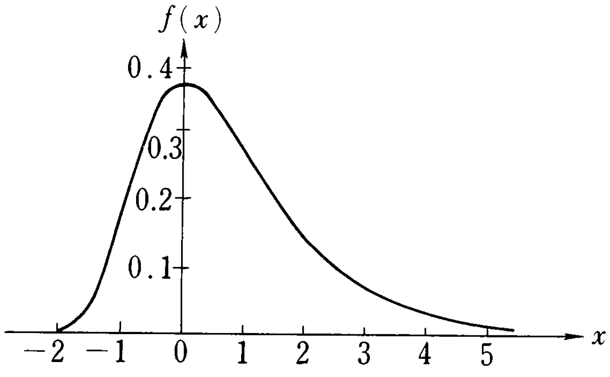


図7　極値分布の密度関数($\theta=1,\ \xi=0$)

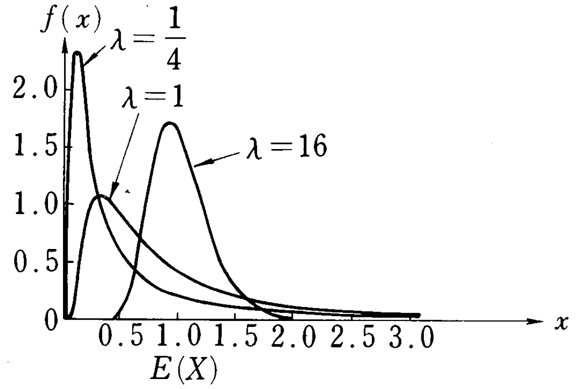


図6　逆ガウス分布の密度関数
($E(T)=d/a=1,\ \lambda=d^2/b$とおいてある)

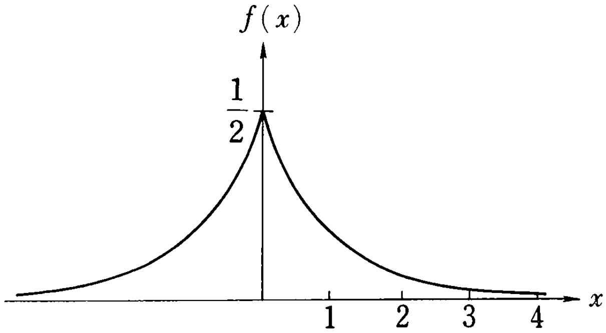


図8　ラプラス分布の密度関数

■逆ガウス分布　水面に浮かべた花粉が，水の分子運動によりランダムに移動する。これをブラウン運動という。この一方向への運動，つまり1次元ブラウン運動において，dを初めて越える時点をTとすると，Tの密度関数は，

$$f(t)=\frac{d}{\sqrt{2\pi b t^3}}\,e^{-\frac{(at-d)^2}{2bt}}\qquad (t>0)$$

となる(図5，6)。aはブラウン運動の単位時間当たりの平均変位を表し，bは拡散係数とよばれる。Tの平均はd/a，分散はbd/a^3である。この分布を**逆ガウス分布**とよぶ。逆に，観測時間長をτと固定したとき，その間の変位量は正規分布$N(a\tau,\ b\tau)$に従う。逆ガウスの名はこの関係に由来する。

■極値分布　分布関数が，

$$F(x)=e^{-e^{-\frac{x-\xi}{\theta}}}\qquad (-\infty<x<\infty)$$

で与えられるとき，これを(タイプ1の)**極値分布**，あるいは**2重指数分布**とよぶ(図7)。この分布は互いに独立で同一分布に従う確率変数X_1，X_2，…，X_nの中の最大値の分布を，適当に正規化してnを大きくしたときの極限の分布として導かれる。それでこの名がある。平均値は$\xi+0.5772\theta$，分散は$\pi^2\theta^2/6$である。また，Xが上の分布に従うとき，$Z=e^{-X-\xi}$とおくと，Zの分布はワイブル分布(→418ページ)となる。

■ラプラス(Laplace)分布　密度関数が，

$$f(x)=\frac{1}{2}e^{-\lambda|x|}\qquad (-\infty<x<\infty)$$

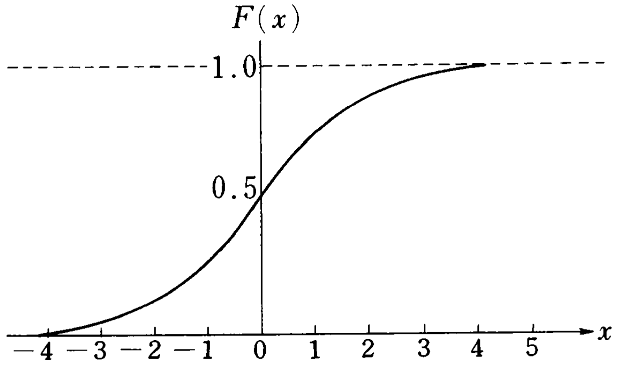


図 9　ロジスティック分布の分布関数
（$\alpha=0$, $\beta=1$ のとき）

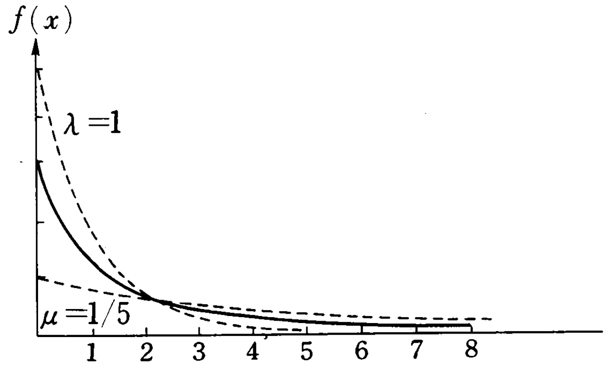


図 10　超指数分布の密度関数
（$\lambda=1.0$, $\mu=1/5$, $p=1/2$ のとき）

で与えられる(図8)。**両側指数分布**ともいう。X_1, X_2 が独立でともに平均 $1/\lambda$ の指数分布に従うとき，X_1-X_2 の分布はラプラス分布となる。

■ロジスティック分布　分布関数が

$$F(x)=1/(1+e^{-\frac{x-\alpha}{\beta}})\quad(-\infty<x<\infty)$$

で与えられる分布をいう。この分布関数の形は(x を時間と考えたときの)，成長曲線を表現するのに用いられ，そのときは**ロジスティック曲線**とよぶ。この曲線は，

〔時刻 x の成長率〕＝〔現在量－初期量〕×〔最終到達量－現在量〕

という関係から導かれ，ある製品の市場の成長過程の記述や化学反応のモデルとして用いられる。また，任意の母集団からの n 個の標本の最大値と最小値の平均(範囲の中央値)の分布は，n が大きいときこの分布に近づく。

■レイライ (Rayleigh) 分布　密度関数が

$$f(x)=\alpha x e^{-\frac{\alpha x^2}{2}}\qquad \alpha>0,\ (x\geqq 0)$$

で与えられる分布をいう。自由度 2 の χ^2 分布に従う確率変数の平方根がこの分布に従う。つまり，X, Y が独立でそれぞれ正規分布 $\mathrm{N}(0,1)$ に従うとき，(2 次元)座標 (X, Y) の原点からの距離 $\sqrt{X^2+Y^2}$ の分布がレイライ分布となる。この分布の故障率は，

$$r(x)=\alpha x$$

と直線的に増加している。

■超指数分布　計算機の処理時間を調べると，比較的短いジョブと長いジョブとが，ある割合(たとえば，p と $1-p$ の割合)で混じり合っている。短いジョブの処理時間の分布が平均 $1/\lambda$ の指数分布，長い方が平均 $1/\mu$ の指数分布に従うとき，それらが混合したジョブ全体の処理時間の分布の密度関数は，

$$f(x)=p\cdot\lambda e^{-\lambda x}+(1-p)\cdot\mu e^{-\mu x}\quad(x\geqq 0)$$

となり，(2 次の)超指数分布とよぶ。変動係数(→59, 348 ページ)が 1 より大きいバラツキの大きな場合によくあてはまる。一般に，上のようにパラメータの違う分布に重みをつけて合わせたものを**混合分布**という。ほかにも，アーラン分布を混合した Luchak 分布など，状況に合わせていろいろ用いられている。　　（森）

労働統計

雇用・失業，労働市場，賃金，労働時間，労働者福利，労働災害，労働争議など，労働問題として考えられるすべての事項を観察対象とする統計を総称して労働統計という。これに労働生産性，労働者家計及び労働者生活に大きな影響を及ぼす物価統計(消費者物価)を加えて，広義の労働統計とすることもある。

労働統計は主として労働省が実施しているが，雇用・失業統計の一部，家計，消費者物価は総務庁統計局が，労働生産性については日本生産性本部が実施している。

★解説

雇用・失業に関する統計は，人口，就業，失業，非労働力(家事・通学など)の各総量や割合，就業している者の産業，職業，従業上の地位(自営業主，家族従業者及び雇用者)の別，雇用者の入職，離職の状況，あるいは，部門間などの労働移動，労働者の採用から退職にいたるまでの企業の雇用制度などについて明らかにしている。

労働市場統計は，公共職業安定所で取り扱った求人，求職，紹介及び就職の状況を統計化したもので，各公共職業安定所，都道府県及びわが国全体の別に集計されている。

賃金統計には，①平均賃金の月々の動きを明らかにする賃金水準統計(毎月勤労統計調査)，②個々の労働者の賃金を労働者の属性(男女，学歴，年齢など)別に明らかにする構造統計(賃金構造基本統計調査)，③賃金の構成，制度及び労働費用の実態を明らかにした制度統計，がある。

労働時間統計は，実労働時間と労働時間制度に関する統計からなり，賃金に関する統計とあわせて調査されている。

労働災害統計は，労働災害の発生状況を明らかにした統計と安全・衛生，労働者の健康状況などを対象とした安全・衛生統計とがある。

関連ページ
日本の統計　382

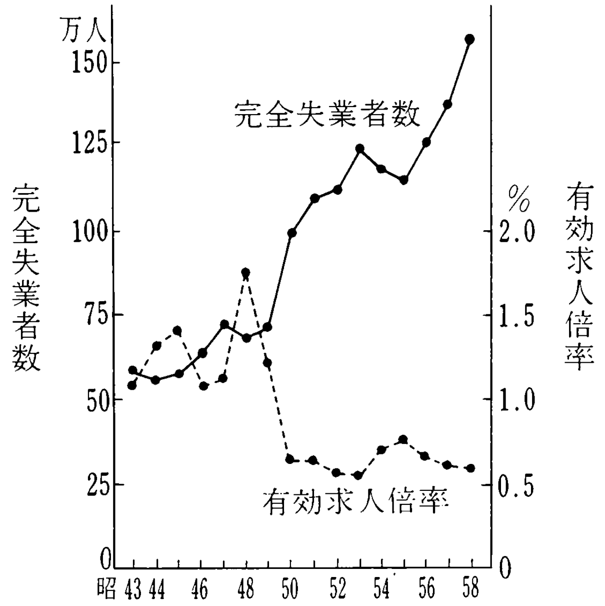


図 1　失業と求人倍率

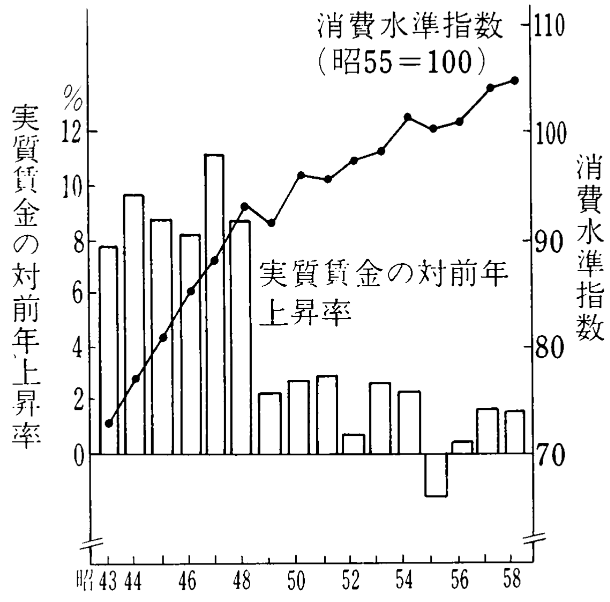


図 2　実質賃金と消費水準

昭和26～48年にいたる高度経済成長期には，労働力過剰の状態が急速に改善され，43年ごろからは人手不足の経済に入っていった。しかし，48年秋の石油危機に伴う物価上昇を沈静化するためにとられた経済調整政策は，その後の低成長経済に変わり，人手不足の基調を大きく転換させた。

労働力の需給関係を示す完全失業者数は，48年までは年平均60万人前後の低水準であった。しかし，49年以後急速に増加し，50年に100万人の大台を超え，53年には122万人，57年には136万人となり，さらに58年には156万人と，ここ30年来の最高水準となった。公共職業安定所における有効求人倍率(求人数を求職者数で割ったもの，前月からの繰越を含む)は，48年の1.76倍から一転して下降に転じ，52年には0.56倍に低下した。その後も0.6～0.7倍で推移している。

雇用，失業，労働市場の状況は，経済の成長と景気の動向との関係から観察すると理解しやすい。

労働力需給の悪化は，企業業績の停滞も加わって賃金や消費水準(物価変動の影響を取り除いた実質消費の水準)の向上を妨げる要因となった。昭和40～48年間では，年率9.1％の上昇率で高まってきた実質賃金の改善スピードは，その後の10年間では，わずかに年率1.4％の上昇率に落ちた。また，生活水準の指標となる消費水準指数(55年＝100)は，38年の56.7から48年には93.4と，10年間に65％高まったが，58年には104.8と，その後の10年間はわずか12％の改善にとどまっている。

賃金，家計の改善は，経済全般，労働力需給，物価動向の関連からとらえるとよい。　(市野)

ワーク・デザインとKJ法

ワーク・デザインは目的中心の問題解決法としてナドラー(G.Nadler)によって創始された方法。対象とする問題にとっての目的は何かを掘り下げることを手がかりに理想システムを描き出し，現実にとるべき行動を導き出す手順を整理したものである。

KJ法は，問題の構造を断片的に得られた事実の集約から浮かび上がらせることを目標に，川喜田二郎氏によって創案された創造的問題解決の技法体系と思想である。

これらは問題を明確にするための2つの方向の代表的手法と見なされ，相互に表裏一体の関係を持っているとも考えられている。

★解説

トップ・ダウンとボトム・アップという見方で見れば，ワーク・デザインは前者で，KJ法は後者になる。機能展開型と事実収束型という分類をすれば，ワーク・デザインは前者で，KJ法は後者になる。「新QC7つ道具」の1つとして，QC運動の中で利用されている系統図法も前者に属するものと見ることができる。

これらの手法の修得のために権威づけられた専門の研修コースが設けられている点は共通で，それらへの参加を通じてこれら手法を体得することがすすめられている。

ワーク・デザインの目的　問題として取り上げている対象を，それを取り上げた当事者の持っている固有の目的を達成させるシステムとしてとらえ，コスト対効果の面からも，人間性尊重の面からも，できるだけ理想に近いシステムを企画・開発・設計・実施することである。このため，関係者を集め，現状にとらわれず理想システムをまず描き出し，順次現実に近づけていく。その間に改善の途を探っていく手順がこの手法の根幹となっている。

関連ページ

■ワーク・デザインの理念　理念として次の5つがあげられている。

(1)　生産性向上と人間性の尊重。

(2)　対象システムの改善は，不満の解消にとどまらず,将来のよりよい状況実現のための開発まで含めて考える。

(3)　制約条件の許す限り，理想システムに近いものを設計する。

(4)　対象はすべてシステムとして扱い，その特性をすべて考慮する。

(5)　関係者の参画意欲と全人格的能力の発揮。

■理想システム　最終目標の理想システムは，時間0，コスト0で目的を達成できるシステムである。この状態までは，何らかの意味で改善の余地ありと考えられるからである。これは無理としても，研究開発によりいずれは実現可能と考えられる，いわば開発目標の理想システム，更には現在の技術で実現可能なシステム，というように目標を下げて適当な目標を設定する。

システムとは，目的，インプット，アウトプット，変換順序，環境，設備，人間的要素，情報キャタリスト(インプットをアウトプットに変換する個々の情報的手段)の8つの特性から成ると考え，当面するシステムでは，これらの特性が具体的にどうなっているかを考えながら分析を進めるのである。

■ワーク・デザインの手順　10段階にわたる手順が定められていて，中には更に数段階のステップに分けられているものもある。また，それを進める上での基本原則が明示されているものもある。この作業を進める上で，定量的データを扱うOR手法が役立つこともある(参考文献〔13〕など)。

■KJ法　現場における観察・調査から得られた知見には，個々に見ると一見脈絡のないものも多いが，それを総合するとそこに構造が浮かび上がることが多い。そのような構造を見出すために,個々の事実をまずカードに書く。カードができ上がったら，それらを目の前に並べ，それらの間に何らかの意味で近いものを探し出す。近いものどうしをセットにし，その要点を一行見出しにまとめ，グループ化していく。その際，まず小さなグループに分け，更にそのグループ間に類似性があれば，それらを大きなグループへと集約していく。次に,こうしてできたグループ及びその内部のカード間の関係に注目し，配置を行う。因果関係，相互関係，対立・矛盾などさまざまな関係が見出され，順序関係があれば1列に並ぶし，循環するものは円形に配列が可能であろう。この図解ができれば，それを文章化あるいは口頭で説明する。こうして問題や状況の本質を見抜き，解決策の立案を行う手法。このKJ法の名は,創始者の頭文字からつけられ,登録商標である。研修機関はKJ法本部・川喜田研究所（03−493−5625）〔16，17，18〕(森村)

ワイブル分布

信頼性工学において寿命の分布をあらわす代表的な分布として知られている。スウェーデンのワイブル(W. Weibull)によって1950年ごろに初めて提唱されたので，この名が冠せられた。

★解説

ワイブル分布の確率密度関数は，

$$f(t)=\frac{m}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{m-1}\exp\left[-\left(\frac{t-\gamma}{\eta}\right)^{m}\right],\quad \gamma\leqq t<\infty$$

$$\text{or}\quad =\frac{m}{t_0}(t-\gamma)^{m-1}\exp\left[-\frac{(t-\gamma)^m}{t_0}\right],\quad \gamma\leqq t<\infty$$

によってあらわせる。ここで，

$$\text{平均 (MTTF)}=\eta\cdot\Gamma(1+1/m)$$

となるので，ηを**尺度パラメータ**という。また，図1より分かるように，γは**位置パラメータ**の役割をもつ。さらに，図1より分かるように，mは**形状パラメータ**とよばれる。

この分布に対応する**故障率関数**は，

$$\lambda(t)=m(t-\gamma)^{m-1}/\eta$$

となる。したがって図2より分かるように，mの値によって故障のパターンが異なり，

$0<m<1$ ……**初期故障**，$m=1$ ……**偶発故障**，

$m>1$ ……**摩耗故障**

のようになる。

関連ページ

MTTF 40

故障率 156

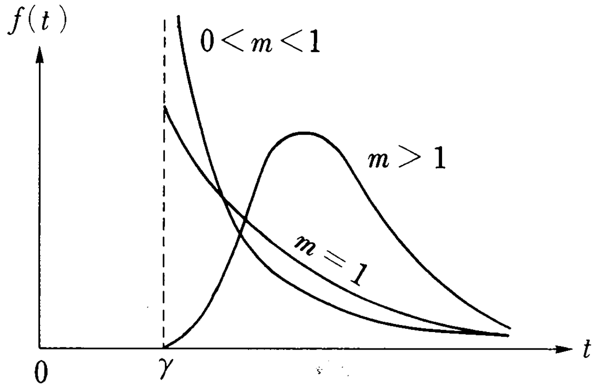


図1 $f(t)$の図

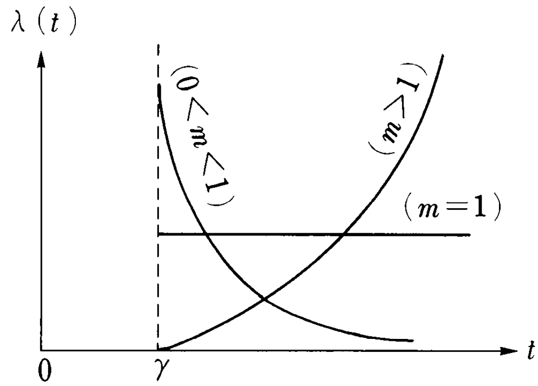


図2 $\lambda(t)$の図

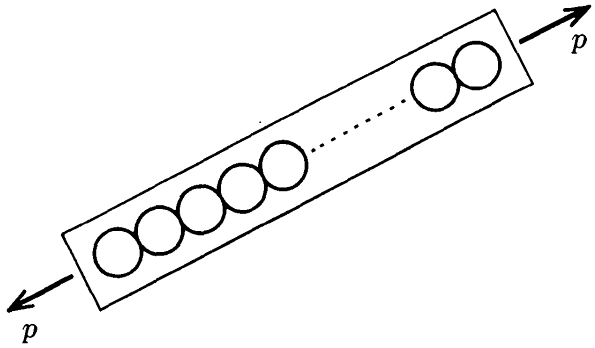


図3　最弱リンクモデル

■ワイブル分布の意味　ワイブル分布はワイブルの発案によるもので，ワイブルのこの分布の発想のよりどころは，最弱リンクモデルである。最弱リンクモデルは図3のような環のチェーンを考え，これをpなる力で引っ張ると，最弱強度の環が切れるとしたものである。

いま，1つの環の強度分布関数を

$$F(p)=P(S<p)$$

（ここで，Sはリンクの強度）

とすれば，この環の信頼度は，

$$R(p)=1-F(p)$$

となり，さらに，n個の環によって構成されるリンクの信頼度 $\boldsymbol{R}_s(p)$ は，

$$\boldsymbol{R}_s(p)=\{R(p)\}^n$$

となる。このリンクの中の最弱強度の環が切れる確率がこれである。

このとき，$R(p)$と$\boldsymbol{R}_s(p)$が同じ形の関数形であるためには，

$$R(p)=e^{cp^k}$$

の形となる。この信頼度関数の形を

$$R(p)=e^{-p^m/p_0}$$

とおいたものがワイブル分布の場合の信頼度関数である。

■ワイブル分布の特長　ワイブル分布は次のような特長を持っている。

(1)　いろいろな機器の寿命分布によく合っている。

(2)　この分布はmの値が1より大，1より小，または1に等しいときに，それぞれ，故障のパターンが摩耗，破壊，および偶発型となっているので，種々の故障パターンを説明するのに便利である。

(3)　次にのべるようなワイブル確率紙を用いると，データ解析が簡便となる。

■ワイブル確率紙　ワイブル分布の分布関数を

$$F(t)=e^{-t^m/t_0}$$

とおけば，

$$\log\log\frac{1}{1-F(t)}$$
$$=m\log t-\log t_0$$

となる。これを書き直して

$$Y=mX-\log t_0$$

とおけば，これは直線の方程式となる。このことによって，横軸を対数目盛りで t を，縦軸を $\log\log\{1-F(t)\}^{-1}$ 目盛りで$F(t)$を目盛った確率紙をワイブル確率紙という。　（真壁）

割当問題

n 人の人に n 個の仕事を割り当てるとき，最も効率のよい割り当てをしたい，という形の問題をいう。下で見るように，輸送問題の特別な形とも見られるが，整数計画問題の1つとも見られる。ハンガリー法など特別な解法があるので解きやすい。

★解説

たとえば，表1のように，A，B，C，Dの4人が，a，b，c，d の4つの仕事を終える予想時間が与えられているとしよう。誰にどの仕事を割り当てるのが最も都合がよいか。仕事で見ると a はB，b はD，c はCがやるのがよいが，d もCがよいので，c と競合する。こういう事情を勘案して割当案を定めるのである。

職＼人	A	B	C	D
a	5	4	7	6
b	6	7	3	2
c	8	11	2	5
d	9	8	6	7

表1

各人への仕事の割り当ては (c, a, b, d) のように順列の形で表せるから，あらゆる割り当て方は $n!$ 通りであり，それぞれについて所要時間を $8+4+3+7=22$ のように計算してその最小値を探せば，この問題は解ける。しかし，この方法は n が10でもとても実用にはならないほどの計算時間を必要とする。

人 i に仕事 j が割り当てられたら $X_{ij}=1$，割り当てられなかったら $X_{ij}=0$ という値をとる変数 X_{ij} を用いると，

$$
\begin{array}{ll}
x_{11}+x_{12}+\cdots+x_{1n}=1, & x_{11}+x_{21}+\cdots+x_{n1}=1 \\
x_{21}+x_{22}+\cdots+x_{2n}=1, & x_{12}+x_{22}+\cdots+x_{n2}=1 \\
\vdots & \vdots \\
x_{n1}+x_{n2}+\cdots+x_{nn}=1, & x_{1n}+x_{2n}+\cdots+x_{nn}=1
\end{array}
$$

となるから，この制約条件は輸送問題の制約式の右辺がすべて1という場合になっている。

一方 x_{ij} が0か1ということなので整数条件になり，整数計画問題とも見られる。輸送問題のアルゴリズムで解いてもよいが，もっと容易なアルゴリズムがある。

関連ページ

輸送問題　392

整数計画法　208

■ハンガリー法　ハンガリーのケーニッヒ(König)の業績をしのんでつけられた名まえといわれている。表1の例について説明する。

段階①　各行の各要素からその行の最小値を引き，更に各列から最小値を引く。結果は表2。

0	0	3	2
3	5	1	0
5	9	0	3
2	2	0	1

表2

段階②　表2の0を各行各列から1つずつ選ぶことができれば，その0の位置の組が割当案になる。表2では，それができないので次へ進む。

段階③　表2のすべての0をできるだけ少ない数の縦または横の線で覆う。表3では，縦2本，横1本の線で覆ったが，縦1本，横2本で覆っても，計3本であることに変わりはない。

0	0	3	2
3	5	1	0
5	9	0	3
2	2	0	1

表3

段階④　表3の線で消されていない要素から，それらの最小値(ここでは2)を引き，縦横の線の重なっている要素に加える(表4)。

0	0	5	4
1	3	1	0
3	7	0	3
0	0	0	1

表4

段階⑤　段階②に戻る。表4では，0の要素を各行各列に1個ずつ割り当てて表5または表6を作ることができるからこれで終わる。もしだめなら段階3へ進む。最適解の総費用は17である。

1			
			1
		1	
	1		

表5

	1		
			1
		1	
1			

表6

■アルゴリズムの基本的考え方　割当問題は，費用の和が最小になる組を各行各列から1つずつ選び出して作るのであるから，各行，各列の費用が，どの要素についても同一量だけ多くともあるいは少なくとも，総費用には全く影響がない。それで段階1の操作を行ったのである。費用0の要素だけで割り当てができれば，最低の費用の割り当てが完了したことを意味する。

段階③，④での計算は，ケーニッヒの定理に基づくが，いささか厄介であるため，ここでは触れない。これはまた，ネットワークの上の流れの問題とも密接に関係している。

(注意)なお，前述のアルゴリズムの段階①では，先に行の最小値を引いたが，先に列の最小値を引いてもよい。こうすると表7が得られる。ここから，行の最小値を引くと表4になり，最初の段階で解が得られたことになる。

0	0	5	4
1	3	1	0
3	7	0	3
4	4	4	5

表7

■割当問題になる問題　列車やバス，あるいは航空機の定期運行におけるダイヤ作成には，当然乗務員の割り当てが付随する。乗務員は1乗務を終わると，定められた休養時間の後，別の便に乗務する。制約を満たす範囲で，できるだけ遊休時間を減らすのが目的になる。この種の例題については，たとえば〔14〕参照。　　（森村）

付表・参考文献

1．国の主な統計機構

昭和59年８月現在

総務庁	統計局	統計基準部
		統計調査部
経済企画庁	調査局	
大蔵省	大臣官房	調査企画課
	関税局	輸出課
	証券局	資本市場課
文部省	大臣官房	調査統計課
厚生省	大臣官房	統計情報部
農林水産省	経済局	統計情報部
通商産業省	大臣官房	調査統計部
運輸省	大臣官房	情報管理部
労働省	大臣官房	政策調査部
建設省	計画局	調査統計課

2．主な指定統計一覧

昭和59年８月現在

指定統計番号	名簿	作成期間	調査期日
1	国勢調査	総務庁統計局	昭和22年10月１日，昭和25年以後５年ごとの10月１日
2	事業所統計	総務庁統計局	昭和23年以後３年ごと
7	毎月勤労統計調査	労働省	昭和22年８月以後毎月
10	工業統計調査	通商産業省	昭和22年以後毎年12月31日
11	通商産業省生産動態統計	通商産業省	昭和23年以後毎月
13	学校基本調査	文部省	昭和23年以後毎年５月１日
14	住宅統計調査	総務庁統計局	昭和23年以後５年ごと
23	商業統計	通商産業省	昭和24～27年毎年，昭和29～51年隔年，昭和54年以降３年ごと
26	農林業センサス	農林水産省	昭和25年以後５年ごとの２月１日及び10年ごとの８月１日
30	労働力調査	総務庁統計局	昭和25年４月以後毎月
56	家計調査	総務庁統計局	昭和27年11月以後毎月

57	個人企業経済調査	総務庁統計局	昭和27年10月以後毎四半期
60	厚生行政基礎調査	厚生省	昭和28年から毎年，最近は6月の第1木曜日
94	賃金構造基本統計	労働省	昭和33年以降毎年，最近では6月30日
110	法人企業統計	大蔵省	昭和45年6月以後毎四半期及び毎半年
113	特定サービス産業実態統計	通商産業省	昭和48年以後毎月11月（または12月）1日
114	社会生活基本統計	総務庁統計局	昭和51年10月，昭和56年10月

3．主な経済指数一覧

指標の指　名	主管部課	沿　　革	主な内容	ウェート・算式	公表周期時期・方法
消費者物価指数	総務庁統計局統計調査部消費統計課	昭和21年8月以降につき毎月算出されている。数次の改正が行われたが現行は昭和55年基準による。	512品目（食料201，被服76，光熱7，住居77，雑費151）価格は小売物価統計調査によって得た価格を用いる。全国平均のほか都道府県庁所在都市別等にも公表されている。	ウェート：家計調査によって得た昭和55年中の品目別支出金額をウェートとする。 算式：ラスパイレス方式	毎月 消費者物価指数（速報・月報） 消費者物価指数年報
家計調査消費水準指数	同上	昭和26年を基準の消費水準を作成し，以後数度の基準年次の改正が行われた。現行は昭和55年基準による。	全国及び人口5万以上都市それぞれについて非農家全世帯及び勤労者世帯の指数が作られている。作成区分は消費者物価指数の類別と同じ。	算式：家計調査による1ヵ月1世帯当たりの消費支出金額を世帯人員4人月間日数30.4日（365/12）の支出額に調整し，これを世帯人員，	毎月 家計調査報告

				月間日数調整済55年月別支出額の年間平均値で除して調整支出金額指数を作成し，この指数を消費者物価指数で除して消費水準を算出する。	
貿易価格指数	大蔵省関税局輸出課	昭和23年以降，経済企画庁が貿易指数を作成していたが，昭和30年以降大蔵省が統一して作成している。現在は昭和55年を基準年とした指数である。	輸出1109品目，輸入721品目，価格は採用商品グループの単価による。	算式：フィッシャー式	毎月，年 外国貿易概況
貿易数量指数	同上	昭和23年以降，経済企画庁が貿易指数を作成していたが，昭和30年以降大蔵省が統一して作成している。現在は昭和55年基準による。		算式：フィッシャー式 （金額指数を貿易価格指数で除して作成）	同上
鉱工業生産指数	通商産業大臣官房調査統計部　統計解析課	昭和25年以降5年ごとに基準年を設け作成している。現在のものは，	534品目，業種別のほか，用途別分類によっても作成されている。	ウェート：昭和55年における付加価値額及び生産額の2種類あり工業統計調査	毎月 生産・出荷・在庫統計（速報）

		昭和55年基準による。		の付加価値と生産額により算出している。 算式：ラスパイレス式	鉱工業生産指数年報など
雇用指数	労働大臣官房 政策調査部統計調査第1課	昭和27年から算定されている。 現在は，昭和55年基準による。	毎月勤労統計調査の常用雇用規模30人以上の事業所を対象。 産業別に作成されている。	算式： $\frac{\text{毎月の実数値}\times 100}{\text{55年平均実数値}}$ 3年に1回の調査対象事業所の抽出替えに伴い，時系列の連続性を維持するために過去に遡って指数の修正がなされる。	毎月 毎月勤労統計調査報告 毎月勤労統計調査総合報告書 労働統計年報
賃金指数	同上	同上	現金給与総額・定期給与・所定内給与について作成。 産業別・事業所規模別に作成されている。	同上	同上
総合輸送活動指数	運輸大臣官房 情報管理部情報処理課	昭和40年6月以来毎月作成している。 現行指数は昭和55年を基準年次としている。	鉄道，自動車，船舶，航空機の各輸送機関による輸送量（原則としてトンキロ・人キロ）を対象とする。（基本22系列）	ウェート：昭和55年における各輸送機関別の粗付加価値額及び生産額 算式：ラスパイレス式	毎月 運輸統計季報など

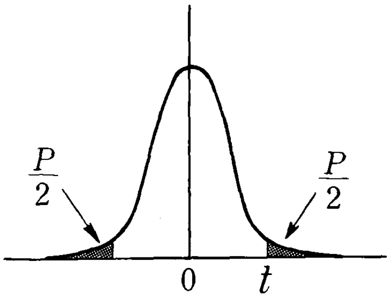
卸売物価指数	日本銀行調査統計局	明治20年以来作成されており数回の改正が行われた。現在は昭和55年基準指数が作成されている。	819品目（工業製品757，鉱産物11，電気・ガスなど5）	ウェート：指数対象品目の総取引額 算式：ラスパイレス式	毎月 卸売物価指数月報 物価指数年報
料金指数（東京）	同上	（基準年次：昭和9～11年）	17項目（電灯，ガス，水道，郵便，電話，電報，貨物，鉄道，電車，バス，新聞，テレビ，映画，入浴，理髪，クリーニング，宿泊）		経済統計月報・年報

4. 正規分布表 $\phi(t)=\dfrac{1}{\sqrt{2\pi}}e^{-t^2/2}$

t	$\int_0^t \phi(t)dt$	$\phi(t)$	t	$\int_0^t \phi(t)dt$	$\phi(t)$	t	$\int_0^t \phi(t)dt$	$\phi(t)$
.00	.00000	.39894	.45	.17364	.36053	.90	.31594	.26609
.01	.00399	.39892	.46	.17724	.35889	.91	.31859	.26369
.02	.00798	.39886	.47	.18082	.35723	.92	.32121	.26129
.03	.01197	.39876	.48	.18439	.35553	.93	.32381	.25888
.04	.01595	.39862	.49	.18793	.35381	.94	.32639	.25647
.05	.01994	.39844	.50	.19146	.35207	.95	.32894	.25406
.06	.02392	.39822	.51	.19497	.35029	.96	.33147	.25164
.07	.02790	.39797	.52	.19847	.34849	.97	.33398	.24923
.08	.03188	.39767	.53	.20194	.34667	.98	.33646	.24681
.09	.03586	.39733	.54	.20540	.34482	.99	.33891	.24439
.10	.03983	.39695	.55	.20884	.34294	1.00	.34134	.24197
.11	.04380	.39654	.56	.21226	.34105	1.01	.34375	.23955
.12	.04776	.39608	.57	.21566	.33912	1.02	.34614	.23713
.13	.05172	.39559	.58	.21904	.33718	1.03	.34850	.23471
.14	.05567	.39505	.59	.22240	.33521	1.04	.35083	.23230
.15	.05962	.39448	.60	.22575	.33322	1.05	.35314	.22988
.16	.06356	.39387	.61	.22907	.33121	1.06	.35543	.22747
.17	.06749	.39322	.62	.23237	.32918	1.07	.35769	.22506
.18	.07142	.39253	.63	.23565	.32713	1.08	.35993	.22265
.19	.07535	.39181	.64	.23891	.32506	1.09	.36214	.22025
.20	.07926	.39104	.65	.24215	.32297	1.10	.36433	.21785
.21	.08317	.39024	.66	.24537	.32086	1.11	.36650	.21546
.22	.08706	.38940	.67	.24857	.31874	1.12	.36864	.21307
.23	.09095	.38853	.68	.25175	.31659	1.13	.37076	.21069
.24	.09483	.38762	.69	.25490	.31443	1.14	.37286	.20831
.25	.09871	.38667	.70	.25804	.31225	1.15	.37493	.20594
.26	.10257	.38568	.71	.26115	.31006	1.16	.37698	.20357
.27	.10642	.38466	.72	.26424	.30785	1.17	.37900	.20121
.28	.11026	.38361	.73	.26730	.30563	1.18	.38100	.19886
.29	.11409	.38251	.74	.27035	.30339	1.19	.38298	.19652
.30	.11791	.38139	.75	.27337	.30114	1.20	.38493	.19419
.31	.12172	.38023	.76	.27637	.29887	1.21	.38686	.19186
.32	.12552	.37903	.77	.27935	.29659	1.22	.38877	.18954
.33	.12930	.37780	.78	.28230	.29431	1.23	.39065	.18724
.34	.13307	.37654	.79	.28524	.29200	1.24	.39251	.18494
.35	.13683	.37524	.80	.28814	.28969	1.25	.39435	.18265
.36	.14058	.37391	.81	.29103	.28737	1.26	.39617	.18037
.37	.14431	.37255	.82	.29389	.28504	1.27	.39796	.17810
.38	.14803	.37115	.83	.29673	.28269	1.28	.39973	.17585
.39	.15173	.36973	.84	.29955	.28034	1.29	.40147	.17360
.40	.15542	.36827	.85	.30234	.27798	1.30	.40320	.17137
.41	.15910	.36678	.86	.30511	.27562	1.31	.40490	.16915
.42	.16276	.36526	.87	.30785	.27324	1.32	.40658	.16694
.43	.16640	.36371	.88	.31057	.27086	1.33	.40824	.16474
.44	.17003	.36213	.89	.31327	.26848	1.34	.40988	.16256

t	$\int_0^t \phi(t)\,dt$	$\phi(t)$	t	$\int_0^t \phi(t)\,dt$	$\phi(t)$	t	$\int_0^t \phi(t)\,dt$	$\phi(t)$
1.35	.41149	.16038	1.80	.46407	.07895	2.25	.48778	.03174
1.36	.41309	.15822	1.81	.46485	.07754	2.26	.48809	.03103
1.37	.41466	.15608	1.82	.46562	.07614	2.27	.48840	.03034
1.38	.41621	.15395	1.83	.46638	.07477	2.28	.48870	.02965
1.39	.41774	.15183	1.84	.46712	.07341	2.29	.48899	.02898
1.40	.41924	.14973	1.85	.46784	.07206	2.30	.48928	.02833
1.41	.42073	.14764	1.86	.46856	.07074	2.31	.48956	.02768
1.42	.42220	.14556	1.87	.46926	.06943	2.32	.48983	.02705
1.43	.42364	.14350	1.88	.46995	.06814	2.33	.49010	.02643
1.44	.42507	.14146	1.89	.47062	.06687	2.34	.49036	.02582
1.45	.42647	.13943	1.90	.47128	.06562	2.35	.49061	.02522
1.46	.42786	.13742	1.91	.47193	.06439	2.36	.49086	.02463
1.47	.42922	.13542	1.92	.47257	.06316	2.37	.49111	.02406
1.48	.43056	.13344	1.93	.47320	.06195	2.38	.49134	.02349
1.49	.43189	.13147	1.94	.47381	.06077	2.39	.49158	.02294
1.50	.43319	.12952	1.95	.47441	.05959	2.40	.49180	.02239
1.51	.43448	.12758	1.96	.47500	.05844	2.41	.49202	.02186
1.52	.43574	.12566	1.97	.47558	.05730	2.42	.49224	.02134
1.53	.43699	.12376	1.98	.47615	.05618	2.43	.49245	.02083
1.54	.43822	.12188	1.99	.47670	.05508	2.44	.49266	.02033
1.55	.43943	.12001	2.00	.47725	.05399	2.45	.49286	.01984
1.56	.44062	.11816	2.01	.47778	.05292	2.46	.49305	.01936
1.57	.44179	.11632	2.02	.47831	.05186	2.47	.49324	.01889
1.58	.44295	.11450	2.03	.47882	.05082	2.48	.49343	.01842
1.59	.44408	.11270	2.04	.47932	.04980	2.49	.49361	.01797
1.60	.44520	.11092	2.05	.47982	.04879	2.50	.49379	.01753
1.61	.44630	.10915	2.06	.48030	.04780	2.51	.49396	.01709
1.62	.44738	.10741	2.07	.48077	.04682	2.52	.49413	.01667
1.63	.44845	.10567	2.08	.48124	.04586	2.53	.49430	.01625
1.64	.44950	.10396	2.09	.48169	.04491	2.54	.49446	.01585
1.65	.45053	.10226	2.10	.48214	.04398	2.55	.49461	.01545
1.66	.45154	.10059	2.11	.48257	.04307	2.56	.49477	.01506
1.67	.45254	.09893	2.12	.48300	.04217	2.57	.49492	.01468
1.68	.45352	.09728	2.13	.48341	.04128	2.58	.49506	.01431
1.69	.45449	.09566	2.14	.48382	.04041	2.59	.49520	.01394
1.70	.45543	.09405	2.15	.48422	.03955	2.60	.49534	.01358
1.71	.45637	.09246	2.16	.48461	.03871	2.61	.49547	.01323
1.72	.45728	.09089	2.17	.48500	.03788	2.62	.49560	.01289
1.73	.45818	.08933	2.18	.48537	.03706	2.63	.49573	.01256
1.74	.45907	.08780	2.19	.48574	.03626	2.64	.49585	.01223
1.75	.45994	.08628	2.20	.48610	.03547	2.65	.49598	.01191
1.76	.46080	.08478	2.21	.48645	.03470	2.66	.49609	.01160
1.77	.46164	.08329	2.22	.48679	.03394	2.67	.49621	.01130
1.78	.46246	.08183	2.23	.48713	.03319	2.68	.49632	.01100
1.79	.46327	.08038	2.24	.48745	.03246	2.69	.49643	.01071

5. t 分布表

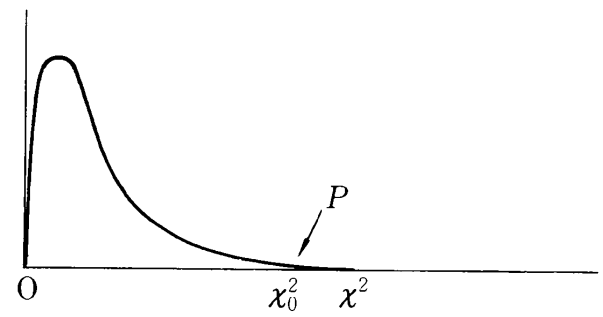


$$P=2\int_{t_0}^{\infty}\frac{\Gamma\left(\frac{\phi+1}{2}\right)}{\sqrt{\phi\pi}\,\Gamma\left(\frac{\phi}{2}\right)\left(1+\frac{t^2}{\phi}\right)^{(\phi+1)/2}}dt$$

$(\phi\cdot P)\rightarrow t$

ϕ \ P	.40	.30	.20	.10	.05	.02	.01	.001
1	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.941
4	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859
6	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405
8	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

6. χ^2(カイ２乗)分布表



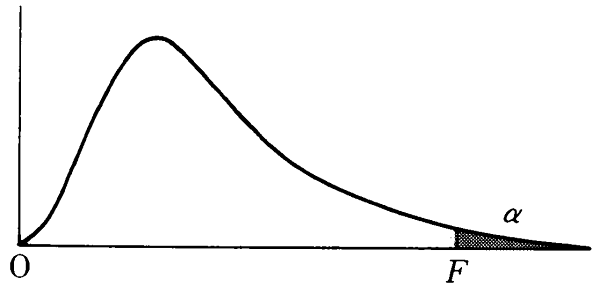
$$P=\int_{\chi_0^2}^{\infty}\frac{2^{-\phi/2}}{\Gamma\left(\frac{\phi}{2}\right)}(\chi^2)^{\phi/2-1}e^{-\chi^2/2}d\chi^2$$

$(\phi\cdot P)\rightarrow\chi^2$

ϕ \ P	.995	.990	.975	.950	.900	.500	.250	.100	.050	.025	.010	.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^2392	0.016	0.455	1.323	2.71	3.84	5.02	6.63	7.88
2	0.010	0.020	0.051	0.103	0.211	1.386	2.77	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	2.37	4.11	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	3.36	5.39	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.610	4.35	6.63	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	2.20	5.35	7.84	10.64	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.17	2.83	6.35	9.04	12.02	14.07	16.01	18.48	20.3
8	1.344	1.646	2.18	2.73	3.49	7.34	10.22	13.36	15.51	17.53	20.1	22.0
9	1.735	2.09	2.70	3.33	4.17	8.34	11.39	14.68	16.92	19.02	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	9.34	12.55	15.99	18.31	20.5	23.2	25.2

11	2.60	3.05	3.82	4.57	5.58	10.34	13.70	17.28	19.68	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	11.34	14.85	18.55	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	12.34	15.98	19.81	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	13.34	17.12	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	14.34	18.25	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	15.34	19.37	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.09	16.34	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.86	17.34	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.12	11.65	18.34	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.85	12.44	19.34	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.28	11.59	13.24	20.34	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	10.98	12.34	14.04	21.34	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.20	11.69	13.09	14.85	22.34	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.86	12.40	13.85	15.66	23.34	28.2	33.2	36.4	39.4	43.0	45.6
25	10.52	11.52	13.12	14.61	16.47	24.34	29.3	34.4	37.7	40.6	44.3	46.9
26	11.16	12.20	13.84	15.38	17.29	25.34	30.4	35.6	38.9	41.9	45.6	48.3
27	11.81	12.88	14.57	16.15	18.11	26.34	31.5	36.7	40.1	43.2	47.0	49.6
28	12.46	13.56	15.31	16.93	18.94	27.34	32.6	37.9	41.3	44.5	48.3	51.0
29	13.12	14.26	16.05	17.71	19.77	28.34	33.7	39.1	42.6	45.7	49.6	52.3
30	13.79	14.95	16.79	18.49	20.60	29.34	34.8	40.3	43.8	47.0	50.9	53.7
40	20.7	22.2	24.4	26.5	29.1	39.3	45.6	51.8	55.8	59.3	63.7	66.8
50	28.0	29.7	32.4	34.8	37.7	49.3	56.3	63.2	67.5	71.4	76.2	79.5
60	35.5	37.5	40.5	43.2	46.5	59.3	67.0	74.4	79.1	83.3	88.4	92.0
70	43.3	45.4	48.8	51.7	55.3	69.3	77.6	85.5	90.5	95.0	100.4	104.2
80	51.2	53.5	57.2	60.4	64.3	79.3	88.1	96.6	101.9	106.6	112.3	116.3
90	59.2	61.8	65.6	69.1	73.3	89.3	98.6	107.6	113.1	118.1	124.1	128.3
100	67.3	70.1	74.2	77.9	82.4	99.3	109.1	118.5	124.3	129.6	135.8	140.2
y_p	−2.58	−2.33	−1.96	−1.64	−1.28	0.000	0.674	1.282	1.645	1.960	2.33	2.58

7. F分布表(1)（5％点）



$$\alpha=\frac{n_1^{n_1/2}n_2^{n_2/2}}{B\left(\frac{n_1}{2},\ \frac{n_2}{2}\right)}\int_{F_\alpha}^{\infty}\frac{F^{n_1/2-1}}{(n_1F+n_2)^{(n_1+n_2)/2}}dF,\qquad F_\alpha=F_{n_2}^{n_1}(\alpha)\geqq 0$$

$$B\left(\frac{n_1}{2},\ \frac{n_1}{2}\right)=\frac{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}+\frac{n_2}{2}\right)}$$

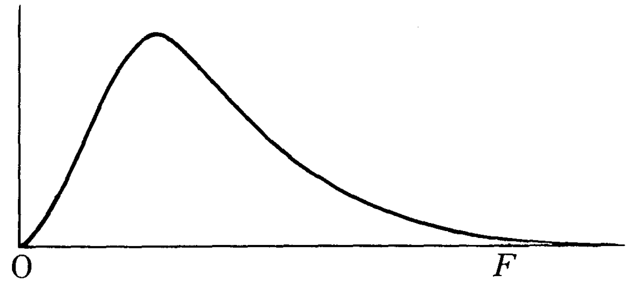
$(n_1, n_2)\ \to F(0.05)$

n_2 \ n_1	1	2	3	4	5	6	8	12	24	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	238.88	243.91	249.05	254.32
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.85	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54

11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25
∞	3.84	2.99	2.60	2.37	2.21	2.09	1.94	1.75	1.52	1.00

（注：R.A.Fisher の数表より転載）

8. F分布表(2)(1%点)



$(n_1, n_2) \to F(0.01)$

n_2 \ n_1	1	2	3	4	5	6	8	12	24	∞
1	4052.2	4999.5	5403.3	5624.6	5763.7	5859.0	5981.6	6106.3	6234.6	6366.0
2	98.50	99.00	99.17	99.25	99.30	99.33	99.37	99.42	99.46	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	9.89	9.47	9.02
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91

11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.59	3.16
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.17	2.80	2.36
22	7.94	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00

（注：R.A.Fisher の数表より転載）

9. 資金の時間的価値の換算係数

終 価 係 数，$S=P\times(1+i)^n$

n＼i	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	i／n
1	1.060	1.080	1.100	1.120	1.150	1.200	1.250	1.300	1.400	1.500	1
2	1.124	1.166	1.210	1.254	1.322	1.440	1.562	1.690	1.960	2.250	2
3	1.191	1.260	1.331	1.405	1.521	1.728	1.953	2.197	2.744	3.375	3
4	1.262	1.360	1.464	1.574	1.749	2.074	2.441	2.856	3.842	5.062	4
5	1.338	1.469	1.611	1.762	2.011	2.488	3.052	3.713	5.378	7.594	5
6	1.419	1.587	1.772	1.974	2.313	2.986	3.815	4.827	7.530	11.391	6
7	1.504	1.714	1.949	2.211	2.660	3.583	4.768	6.275	10.541	17.086	7
8	1.594	1.851	2.144	2.476	3.059	4.300	5.960	8.157	14.758	25.629	8
9	1.689	1.999	2.358	2.773	3.518	5.160	7.451	10.604	20.661	38.443	9
10	1.791	2.159	2.594	3.106	4.046	6.192	9.313	13.786	28.925	57.665	10
11	1.898	2.332	2.853	3.479	4.652	7.430	11.642	17.922	40.496	86.498	11
12	2.012	2.518	3.138	3.896	5.350	8.916	14.552	23.298	56.694	129.746	12
13	2.133	2.720	3.452	4.363	6.153	10.699	18.190	30.288	79.371	194.620	13
14	2.261	2.937	3.797	4.887	7.076	12.839	22.737	39.374	111.120	291.929	14
15	2.397	3.172	4.177	5.474	8.137	15.407	28.422	51.186	155.568	437.894	15

現 価 係 数，$P=S\times\dfrac{1}{(1+i)^n}$

n＼i	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	i／n
1	0.9434	0.9259	0.9091	0.8929	0.8696	0.8333	0.8000	0.7692	0.7143	0.6667	1
2	0.8900	0.8573	0.8264	0.7972	0.7561	0.6944	0.6400	0.5917	0.5102	0.4444	2
3	0.8396	0.7938	0.7513	0.7118	0.6575	0.5787	0.5120	0.4552	0.3644	0.2963	3
4	0.7921	0.7350	0.6830	0.6355	0.5718	0.4823	0.4096	0.3501	0.2603	0.1975	4
5	0.7473	0.6806	0.6209	0.5674	0.4972	0.4019	0.3277	0.2693	0.1859	0.1317	5
6	0.7050	0.6302	0.5645	0.5066	0.4323	0.3349	0.2621	0.2072	0.1328	0.0878	6
7	0.6651	0.5835	0.5132	0.4523	0.3759	0.2791	0.2097	0.1594	0.0949	0.0585	7
8	0.6274	0.5403	0.4665	0.4039	0.3269	0.2326	0.1678	0.1226	0.0678	0.0390	8
9	0.5919	0.5002	0.4241	0.3606	0.2843	0.1938	0.1342	0.0943	0.0484	0.0260	9
10	0.5584	0.4632	0.3855	0.3220	0.2472	0.1615	0.1074	0.0725	0.0346	0.0173	10
11	0.5268	0.4289	0.3505	0.2875	0.2149	0.1346	0.0859	0.0558	0.0247	0.0116	11
12	0.4970	0.3971	0.3186	0.2567	0.1869	0.1122	0.0687	0.0429	0.0176	0.0077	12
13	0.4688	0.3677	0.2897	0.2292	0.1625	0.0935	0.0550	0.0330	0.0126	0.0051	13
14	0.4423	0.3405	0.2633	0.2046	0.1413	0.0779	0.0440	0.0254	0.0090	0.0034	14
15	0.4173	0.3152	0.2394	0.1827	0.1229	0.0649	0.0352	0.0195	0.0064	0.0023	15

資本回収係数，$M=P\times\dfrac{i(1+i)^n}{(1+i)^n-1}$

n＼i	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	i／n
1	1.0600	1.0800	1.1000	1.1200	1.1500	1.2000	1.2500	1.3000	1.4000	1.5000	1
2	0.5454	0.5608	0.5762	0.5917	0.6151	0.6545	0.6944	0.7348	0.8167	0.9000	2
3	0.3741	0.3880	0.4021	0.4163	0.4380	0.4747	0.5123	0.5506	0.6294	0.7105	3
4	0.2886	0.3019	0.3155	0.3292	0.3503	0.3863	0.4234	0.4616	0.5408	0.6231	4
5	0.2374	0.2505	0.2638	0.2774	0.2983	0.3344	0.3718	0.4106	0.4914	0.5758	5
6	0.2034	0.2163	0.2296	0.2432	0.2642	0.3007	0.3388	0.3784	0.4613	0.5481	6
7	0.1791	0.1921	0.2054	0.2191	0.2404	0.2774	0.3163	0.3569	0.4419	0.5311	7
8	0.1610	0.1740	0.1874	0.2013	0.2229	0.2606	0.3004	0.3419	0.4291	0.5203	8
9	0.1470	0.1601	0.1736	0.1877	0.2096	0.2481	0.2888	0.3312	0.4203	0.5134	9
10	0.1359	0.1490	0.1627	0.1770	0.1993	0.2385	0.2801	0.3235	0.4143	0.5088	10
11	0.1268	0.1401	0.1540	0.1684	0.1911	0.2311	0.2735	0.3177	0.4101	0.5058	11
12	0.1193	0.1327	0.1468	0.1614	0.1845	0.2253	0.2684	0.3135	0.4072	0.5039	12
13	0.1130	0.1265	0.1408	0.1557	0.1791	0.2206	0.2645	0.3102	0.4051	0.5026	13
14	0.1076	0.1213	0.1357	0.1509	0.1747	0.2169	0.2615	0.3078	0.4036	0.5017	14
15	0.1030	0.1168	0.1315	0.1468	0.1710	0.2139	0.2591	0.3060	0.4026	0.5011	15

年金現価係数，$P = M \times \dfrac{(1+i)^n - 1}{i(1+i)^n}$

n \ i	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	i / n
1	0.943	0.926	0.909	0.893	0.870	0.833	0.800	0.769	0.714	0.667	1
2	1.833	1.783	1.736	1.690	1.626	1.528	1.440	0.361	1.224	1.111	2
3	2.673	2.577	2.487	2.402	2.283	2.106	1.952	1.816	1.589	1.407	3
4	3.465	3.312	3.170	3.037	2.855	2.589	2.362	2.166	1.849	1.605	4
5	4.212	3.993	3.791	3.605	3.352	2.991	2.689	2.436	2.035	1.737	5
6	4.917	4.623	4.355	4.111	3.784	3.326	2.951	2.643	2.168	1.824	6
7	5.582	5.206	4.868	4.564	4.160	3.605	3.161	2.802	2.263	1.883	7
8	6.210	5.747	5.335	4.968	4.487	3.837	3.329	2.925	2.331	1.922	8
9	6.802	6.247	5.759	5.328	4.772	4.031	3.463	3.019	2.379	1.948	9
10	7.360	6.710	6.145	5.650	5.019	4.192	3.571	3.092	2.414	1.965	10
11	7.887	7.139	6.495	5.938	5.234	4.327	3.656	3.147	2.438	1.977	11
12	8.384	7.536	6.814	6.194	5.421	4,439	3.725	3.190	2.456	1.985	12
13	8.853	7.904	7.103	6.424	5.583	4.533	3.780	3.223	2.469	1.990	13
14	9.295	8.244	7.367	6.628	5.724	4.611	3.824	3.249	2.478	1.993	14
15	9.712	8.559	7.606	6.811	5.847	4.675	3.859	3.268	2.484	1.995	15

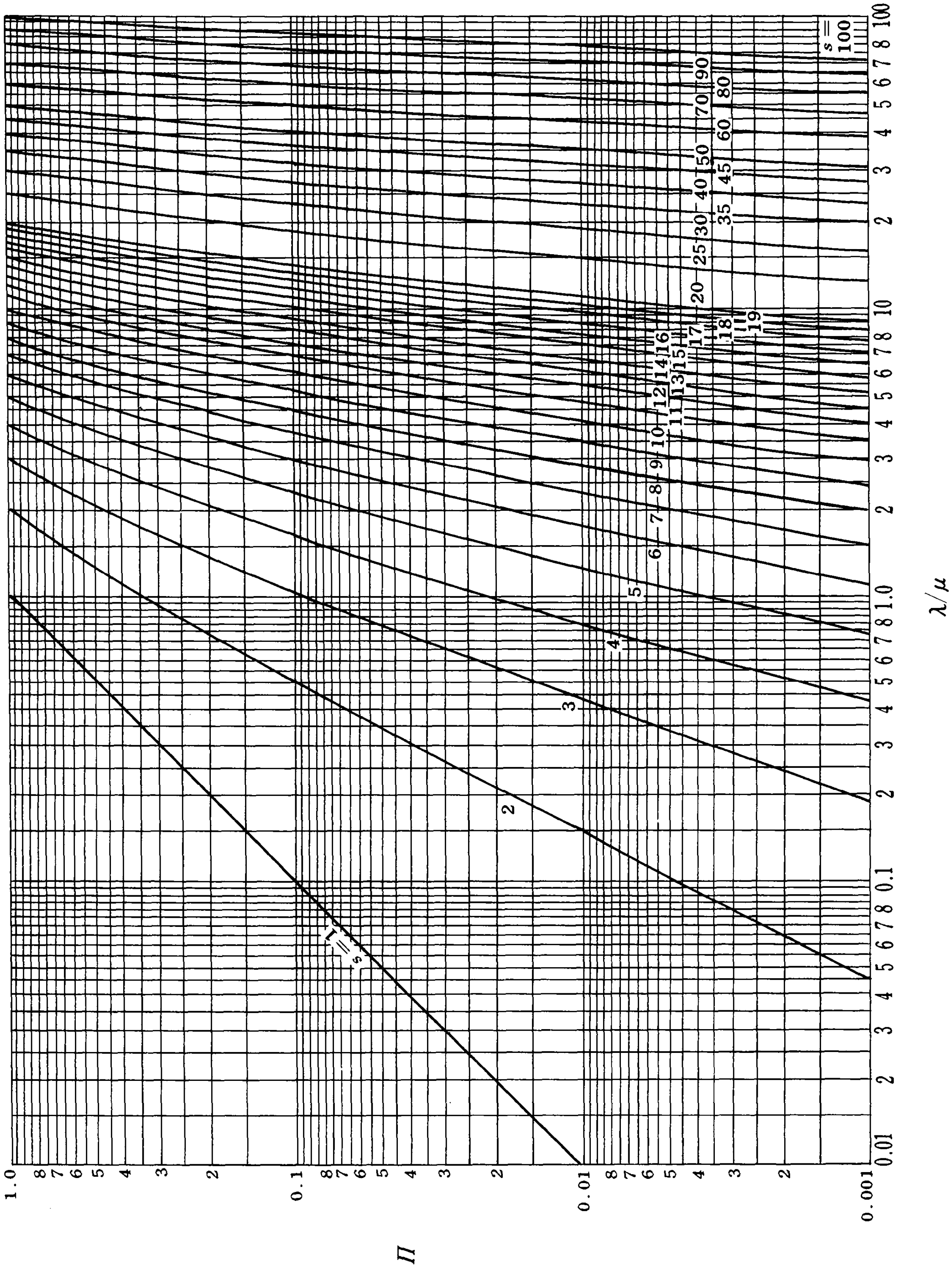
減債基金係数，$M = S \times \dfrac{i}{(1+i)^n - 1}$

n \ i	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	i / n
1	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1
2	0.48543	0.48076	0.47619	0.47169	0.46511	0.45454	0.44444	0.43478	0.41666	0.40000	2
3	0.31410	0.30803	0.30211	0.29634	0.28797	0.27472	0.26229	0.25062	0.22935	0.21052	3
4	0.22859	0.22192	0.21547	0.20923	0.20026	0.18628	0.17344	0.16162	0.14076	0.12307	4
5	0.17739	0.17045	0.16379	0.15740	0.14831	0.13437	0.12184	0.11058	0.09136	0.07582	5
6	0.14336	0.13631	0.12960	0.12322	0.11423	0.10070	0.08881	0.07839	0.06126	0.04812	6
7	0.11913	0.11207	0.10540	0.09911	0.09036	0.07742	0.06634	0.05687	0.04192	0.03108	7
8	0.10103	0.09401	0.08744	0.08130	0.07285	0.06060	0.05039	0.04191	0.02907	0.02030	8
9	0.08702	0.08007	0.07364	0.06767	0.05957	0.04807	0.03875	0.03123	0.02034	0.01335	9
10	0.07586	0.06902	0.06274	0.05698	0.04925	0.03852	0.03007	0.02346	0.01432	0.008823	10
11	0.06679	0.06007	0.05396	0.04841	0.04106	0.03110	0.02349	0.01772	0.01012	0.005848	11
12	0.05927	0.05269	0.04676	0.04143	0.03448	0.02526	0.01844	0.01345	0.007182	0.003883	12
13	0.05296	0.04652	0.04077	0.03567	0.02911	0.02062	0.01454	0.01024	0.005103	0.002582	13
14	0.04758	0.04129	0.03574	0.03087	0.02468	0.01689	0.01150	0.007817	0.003632	0.001718	14
15	0.04296	0.03682	0.03147	0.02682	0.02101	0.01388	0.0091168	0.005977	0.002587	0.001144	15

年金終価係数，$S = M \times \dfrac{(1+i)^n - 1}{i}$

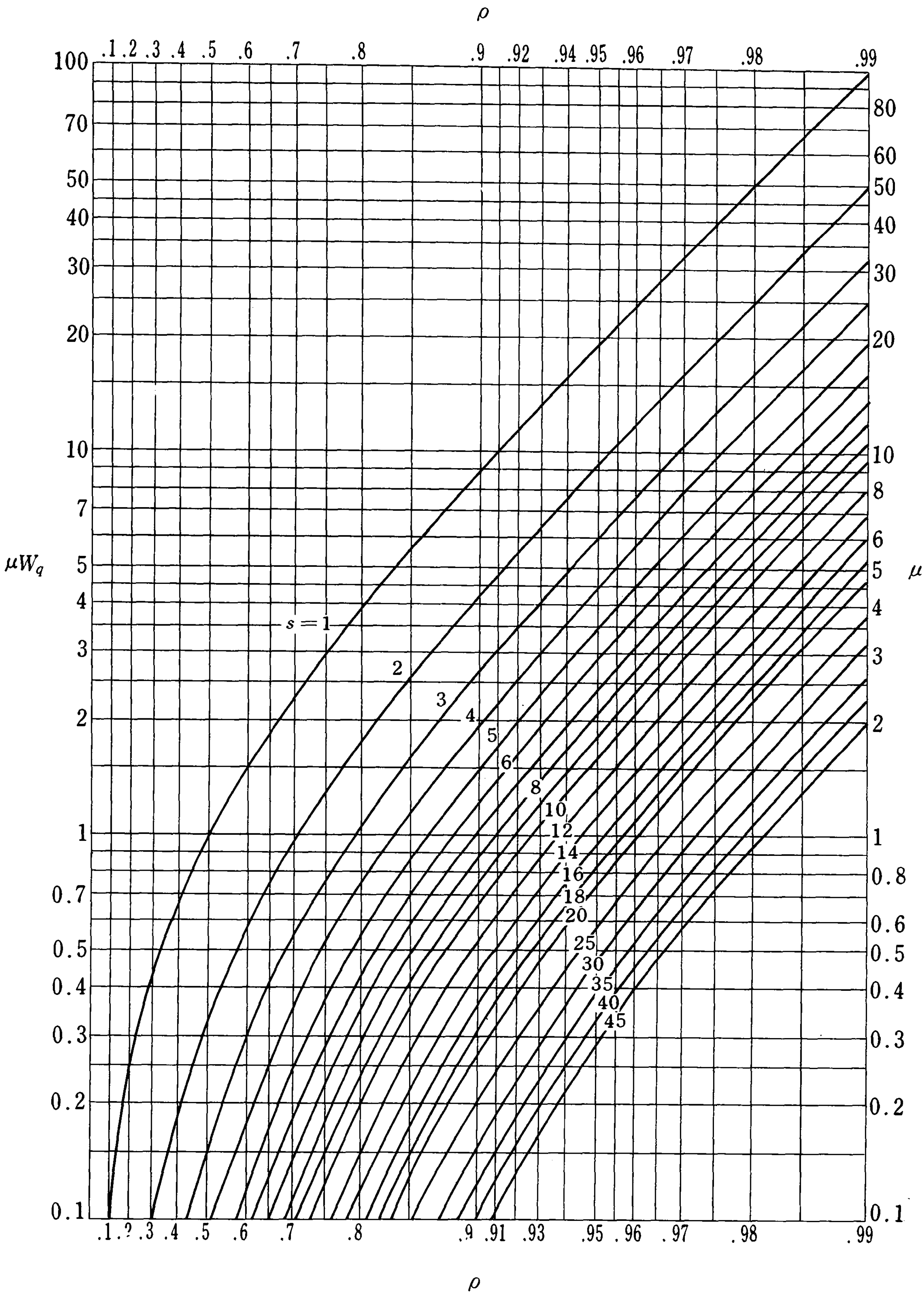
n \ i	6%	8%	10%	12%	15%	20%	25%	30%	40%	50%	i / n
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1
2	2.060	2.080	2.100	2.120	2.150	2.200	2.250	2.300	2.400	2.500	2
3	3.184	3.246	3.310	3.374	3.472	3.640	3.812	3.990	4.360	4.750	3
4	4.375	4.506	4.641	4.779	4.993	5.368	5.766	6.187	7.104	8.125	4
5	5.637	5.867	6.105	6.353	6.742	7.442	8.207	9.043	10.946	13.187	5
6	6.975	7.336	7.716	8.115	8.754	9.930	11.259	12.756	16.324	20.781	6
7	8.394	8.923	9.487	10.089	11.067	12.916	15.073	17.583	23.853	32.172	7
8	9.897	10.637	11.436	12.300	13.727	16.499	19.842	23.858	34.395	49.258	8
9	11.491	12.488	13.579	14.776	16.786	20.799	25.802	32.015	49.153	74.887	9
10	13.181	14.487	15.937	17.549	20.304	25.959	33.253	42.619	69.814	113.330	10
11	14.972	16.645	18.531	20.655	24.349	32.150	42.566	56.405	98.739	170.995	11
12	16.870	18.977	21.384	24.133	29.002	39.581	54.208	74.327	139.235	257.493	12
13	18.882	21.495	24.523	28.029	34.352	48.497	68.760	97.625	195.929	387.239	13
14	21.015	24.215	27.975	32.393	40.505	59.196	86.949	127.913	275.300	581.859	14
15	23.276	27.152	31.772	37.280	47.580	72.035	109.687	167.286	386.420	873.788	15

10. 待ち合わせ率Πの図($M/M/s$)



(注) ポアソン到着, 指数型サービス時間, 順序無関係

11. 平均待ち時間を求める図($M/M/s$)



参考文献

〔1〕R. L. Schultz・D. P. Slevin 編「Implementing Operations Research/Management Science」American Elsevier　1975

〔2〕R. M. サイヤート・J. G. マーチ(松田武彦・井上恒夫訳)「企業の行動理論」ダイヤモンド社　1967

〔3〕朝尾正・安藤貞一・楠正・中村恒夫「最新実験計画法」日科技連　1973

〔4〕朝香鉄一・石川馨編「品質保証ガイドブック」日科技連　1974

〔5〕朝香鉄一編「品質管理」日本規格協会　1980

〔6〕阿部統・国沢清典・藤井光昭・三国一義編「近代統計学小辞典」春秋社　1968

〔7〕石川馨「品質管理入門」日科技連　1964

〔8〕石川馨・藤森利美・久米均「化学者および化学技術者のための実験計画法」東京化学同人　1967

〔9〕上田尚一「統計用語辞典」東洋経済新報社　1981

〔10〕エヴェリット(山内光哉監訳)「質的データの解析」新曜社　1980

〔11〕O. W. Wight (吉谷龍一訳)「MRP による生産管理」日刊工業新聞社　1979

〔12〕奥野忠一・芳賀敏郎「実験計画法」培風館　1969

〔13〕海部不二雄監修・小泉知義「KJ 法とワークデザイン方式による問題解決ワークブック」ダイヤモンド社　1974

〔14〕カウフマン・フォール(阿部統・加藤茂夫訳)「OR への招待」好学社　1972

〔15〕仮谷太一「予測の知識」森北出版　1971

〔16〕川喜田二郎「発想法」中央公論社　1967

〔17〕川喜田二郎「続発想法」中央公論社　1970

〔18〕川喜田二郎・牧島信一「問題解決学－KJ ワークブック」講談社　1970

〔19〕クラスカル・ウィッシュ(高根芳雄訳)「多次元尺度法」朝倉書店　1980

〔20〕小林竜一「数量化理論入門」日科技連　1981

〔21〕駒澤勉「多次元データ分析の基礎」朝倉書店　1978

〔22〕近藤次郎「オペレーションズ・リサーチ」日科技連　1973
〔23〕今野浩・山下浩「非線形計画法」日科技連　1978
〔24〕今野浩・鈴木久敏編「整数計画と組合せ最適化」日科技連　1982
〔25〕塩見弘「信頼性工学入門」丸善　1967
〔26〕杉山高一・牛沢賢二「統計データの読み方」東洋経済新報社　1982
〔27〕杉山高一「多変量データ解析入門」朝倉書店　1983
〔28〕杉山高一「統計学入門」絢文社　1984
〔29〕杉山昌平「動的計画法」日科技連　1976
〔30〕鈴木光男「ゲーム理論入門」共立出版　1981
〔31〕千住鎮雄他「経済性分析」日本規格協会　1979
〔32〕千住鎮雄・伏見多美雄「経済性工学の基礎」日本能率協会　1982
〔33〕千住鎮雄・伏見多美雄「経済性工学の応用」日本能率協会　1983
〔34〕田口玄一他「確率・統計」日本規格協会　1981
〔35〕田口玄一「実験計画法，上・下」丸善　1976
〔36〕中山一郎「統計学辞典」東洋経済新報社　1957
〔37〕T. W. Anderson「An Introduction to Multivariate Analysis」
〔38〕T. W. Anderson「The Statistical Analysis of Time Series」John Wiley　1958
〔39〕西里静彦「質的データの数量化」朝倉書店　1982
〔40〕日本 IE 協会編「IE 技法ハンドブック」丸善　1968
〔41〕日本規格協会編「新版品質管理便覧」　1977
〔42〕ニューエル（森村英典・森雅夫共訳）「待ち行列論の応用」サイエンス社　1974
〔43〕ハワード（関根智明他訳）「ダイナミックプログラミングとマルコフ過程」培風館　1971
〔44〕フェラー（河田龍夫監訳）「確率論とその応用」紀伊國屋書店　1960
〔45〕藤井光昭「時系列解析」コロナ社　1974
〔46〕ブラウン（関根智明訳）「在庫管理における需要予測」紀伊國屋書店　1961
〔47〕ベルマン（小田中敏男他訳）「ダイナミック・プログラミング」東京図書　1973
〔48〕真壁肇「品質管理」朝倉書店　1974
〔49〕真壁肇「品質保証と信頼性」日科技連　1984

〔50〕真壁肇編「オペレーションズ・リサーチ」日本規格協会　1980
〔51〕牧野鉄治・野中保雄「信頼性工学」日科技連　1982
〔52〕牧野都治「格差・パレート図・ABC分析」日本評論社　1984
〔53〕武蔵野通研「待ち行列数表」電電公社
〔54〕森村英典・大前義次「応用待ち行列理論」日科技連　1975
〔55〕森村英典・高橋幸雄「マルコフ解析」日科技連　1979
〔56〕森村英典「確率」共立出版　1984
〔57〕安田三郎「社会統計学」丸善　1972
〔58〕ユックス(後藤昌司訳)「二値データの解析」朝倉書店　1980
〔59〕吉谷龍一他「MRP システム」日本工業新聞社　1977
〔60〕ワグナー(森村英典・伊理正夫監訳)「オペレーションズ・リサーチ入門」2(真鍋龍太郎訳)〈ネットワーク・モデル〉培風館　1976
〔61〕鷲尾泰俊「実験計画法入門」日本規格協会

索引

あ

い

こ

さ

信頼区間　confidence interval　82
区間推定の結果得られた区間。
信頼係数　confidence coefficient　82
信頼度に同じ。
信頼性　reliability　184
要求された機能を果たすことができる性質。
信頼度　confidence level　82
信頼区間内に母数がはいる確率。

す

推移確率行列　transition probability matrix　**382**, 383
マルコフ連鎖を規定する基本量の1つ。
推移図式　transition scheme　383
状態間を矢印で結んで推移を示す図。
水準の選定　21
実験計画における水準選定についての注意。
推測統計学　inferential statistics　75, **261**, 272
統計的推測の理論や方法を研究する学問。
推定値　estimate　134, **254**
標本から計算された推定量の値。
推定量　estimator　254
母数の推定のために考えられた統計量。
数理計画法　mathematical programming　186
制約式のもとで目的関数を最適化する手法。
数量化Ｉ類　quantification method of the first type　**188**, 190, 252, 298
質的データから外的基準を予測する方法。
数量化Ⅱ類　quantification method of the second type　15, **190**, 298
外的基準が名義尺度のときの数量化。
数量化Ⅲ類　quantification method of the third type　**194**, 299
外的基準をもたない質的データの分類手法。
数量化Ⅳ類　quantification method of the forth type　42, **196**, 299
類似度を利用したデータの位置づけ法。
スピアマンの順位相関係数　Spearman's rank correlation coefficient　290
2つの選択順位の相関を測る尺度。

せ

正規確率紙　normal probability paper　198
正規分布に従うか否かを確かめるのに便利。
正規分布　normal distribution　**200**, 226
最も基本的な確率分布の1つ。
正規母集団　normal population　314
母集団分布が正規分布の母集団。
制御因子　20
結果をみて最適水準を採択できる因子。
制御装置　127
プログラムに従って計算機を動かす装置。
政策反復解法　policy iteration procedure　380
マルコフ決定過程の定常最適政策を得る法。
生産国民所得　121
生産の側面からとらえた国民所得。
生産者危険　producer's risk　**63**, 109
抜取検査での第1種の過誤。
生産統計　204
生産の実態を明らかにしている統計。
生産能力指数　205
固定設備がフル生産したときの生産量。
整数計画法　integer programming　186, **208**, 392, 420

そ

ち

つ

な

に

ほ

ま

D

E

F

G

H

I

J

N

O

P

Q

R

S

T

U

V

W

X

Y

Z

統計・OR活用事典

昭和59年9月27日第1版第1刷発行　　定価4,800円

編　者　森村英典／牧野都治／
　　　　真壁　肇／杉山高一

発行者　東京書籍株式会社
　　　　代表者　小髙民雄

印刷所　東京書籍印刷株式会社

発行所　東京書籍株式会社
　　　　東京都台東区台東1-5-18

2534	516061	5313