データは語る

—統計的データ解析の考え方と実際—

成蹊大学工学部　岩崎　学

0．まえがき

本稿は，1997年4月に長野県松本市で行なった日本数学会主催市民講演会での話をまとめたものです．いくつかの実例を基に統計学の考え方を私なりに述べました．初めから読まれても，興味のありそうな例題から読んでいただいても結構です．

1．はじめに

2．データの集計とグラフ表示

　例2．1　冬季五輪における日本の獲得メダル数

　例2．2　学力試験の成績

3．統計的推測の実際

　例3．1　女子学生の比率

　例3．2　缶コーヒーの開発

　例3．3　新薬の開発と承認

4．最近の話題から

　例4．1　高次元超立方体の頂点の射影

5．おわりに

1．はじめに

「統計学」と聞くと何を思い浮かべますか．高校数学の一分野としての「確率・統計」でしょうか，それとも，新聞などで目にする数字の羅列でしょうか．

高校の授業では，サイコロを何度も振ったりコインを続けて投げたりの話がほとんどで，ランダムネス，不確定性の身近な例としてはよいのでしょうが，それが実際にどのように役に立つのか判然とはしません．さらに，計算が面倒なこともあって，あまり好きでなかったという感想をよく聞きます．新聞での経済統計，官庁統計などの

数字の羅列も，その数字の背後にある意味が分からない人にとっては少しも興味が持てないでしょう．また，大学で統計の講義を受けたことのある人の中にも，「とにかく面倒なことだけは覚えている」などという感想を持つ人は多いようです．

面倒な確率計算や数字の配列などは，もちろん統計学の守備範囲ですが，全体の中の一部でしかありません．では，統計学とは何なのでしょう．もちろん一言で定義するのは困難ですが，私の考えを知ってもらうために，私の勤務大学での私の研究室（統計学研究室）の紹介パンフレットの記述をご紹介しましょう．

「人文・社会科学，自然科学，工学などの分野を問わず，ある事柄を主張あるいは立証するためには客観的なデータの裏付けが欠かせない．そのためには，データを効率よく得るための計画から始まり，得られたデータを正しく解析して情報を最大限に引き出すための方法論が必要となる．その方法論の基礎を与えるのが統計学（statistics）である．」
（成蹊大学工学部経営工学科統計学研究室紹介パンフレットより）

抽象的な表現で分かりにくいでしょうから，以下では，私が直接関係した（そして現在も関係している）話題を例として取り上げ，統計学および統計的データ解析の考え方及び実際の使われ方をみていきたいと思います．

２．データの集計とグラフ表示

データの集計，そして適切なグラフ表示は，統計的データ解析の第一歩ですが，侮れない一歩です．身近な統計の例を二つ挙げて，データのグラフ化の勘どころを見ることにしましょう

例２．１．　冬季五輪における日本の獲得メダル数

1998年には長野で日本にとっては２回目の冬季オリンピックが開催されることになっています．過去に日本がどのくらいのメダルを獲得したのかの一覧表が表２．１で，それをグラフ化したものが図２．１です．表よりもグラフのほうが獲得メダル数の年ごとの推移がよく分かります．一昔前は（といっても10年くらい前に過ぎませんが），このようなグラフ化は手で行なっていましたが，現在では，パソコンの表計算ソフト

を使っていとも簡単にできるようになりました．数学の諸分野の中でも，統計学がパソコンの恩恵を一番受けているのではないでしょうか．

グラフからは，1972年大会のメダル３個が目を引きます．これは，ご存じの方はご存じでしょうが，札幌大会における70m級ジャンプでの金，銀，銅独占という「日の丸飛行隊」の快挙によるものです．また，最近メダル数が増加している傾向も読み取ることができます．

表２.１　冬季五輪における日本の獲得メダル数

回	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
年	24	28	32	36	48	52	56	60	64	68	72	76	80	84	88	92	94	98
金	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	?
銀	0	0	0	0	0	0	1	0	0	0	1	0	1	1	0	2	2	?
銅	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	4	2	?
計	0	0	0	0	0	0	1	0	0	0	3	0	1	1	1	7	5	?

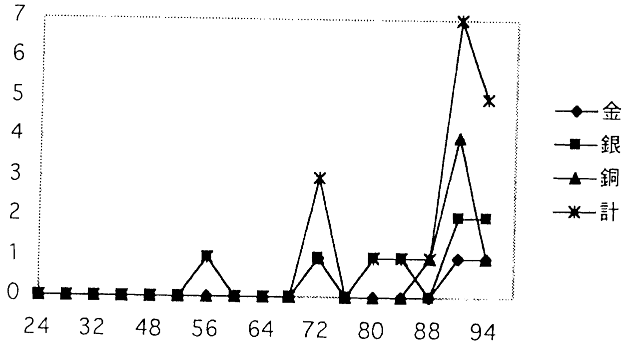
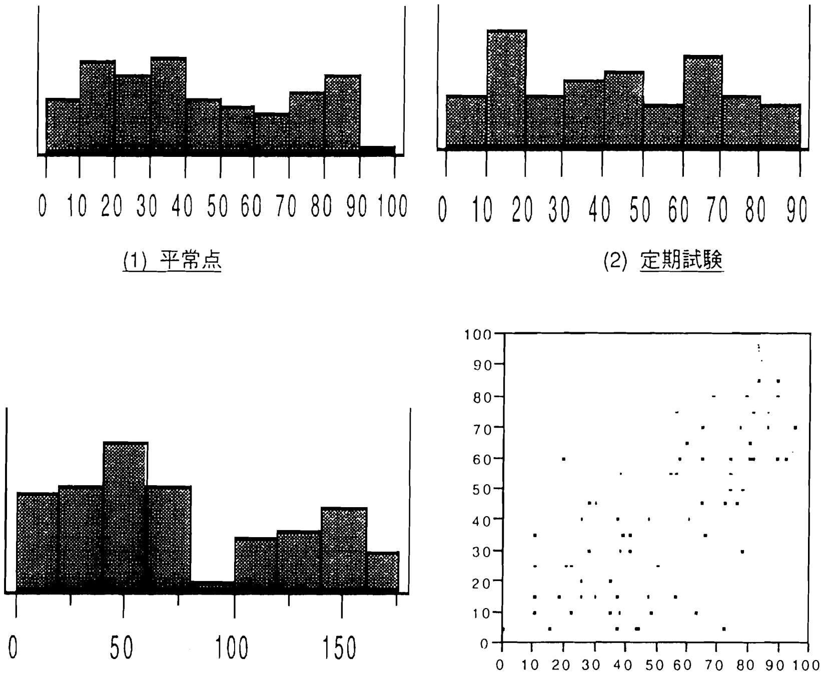


図２.１．冬季五輪における日本の獲得メダル数のグラフ

獲得メダル数の話は第５節でもう一度触れることにして，次にもっと身近な授業における成績評価について見てみましょう．

例２．２．　学力試験の成績

図２．２（1），（2）は，私の勤務する大学での「確率統計」の授業における「平常点」と「定期試験」の点数をヒストグラム化したものです．「平常点」は，授業への出席回数，小テストの点数，レポートの評価などを総合したもので，「定期試験」の点数と合わせて最終的な成績評価としています．「平常点」，「定期試験」とも正規分布（第４節で後述）という感じもしませんが，かといって際立った特徴も見られません．



(1) 平常点　(2) 定期試験

(3) 合計点（平常点＋定期試験）　(4) 平常点と定期試験の散布図

図2.2．ある授業科目における成績評価

ところが，それらの合計点のヒストグラム（図2．2（3））を見ると，明らかに二山あることが分かります．こういう結果になると教師は楽で，安心して点数の低い学生を落とすことができます（実際は，それでは落ちる学生が多すぎるので，少し甘く評価を付けましたが）．

「平常点（x_1）」と「定期試験（x_2）」の2次元の散布図を描いたのが図2．2（4）です．散布図から，どうやら2群ありそうだということが見て取れます．合計点を y とし，x_1，x_2 を要素にもつ2次元ベクトルを $x=(x_1, x_2)$ とすると，$a=(1, 1)$ として，

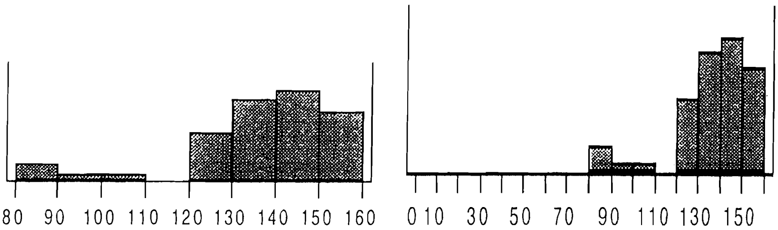
$$y = x_1 + x_2 = (a, x) \qquad (2.1)$$

ですから（ここで，(,) は内積を表わす記号），合計点は，ベクトル a 方向への各データの射影と見ることができます．x_1，x_2 各々の分布（統計の言葉では周辺分布）を見ていただけでは分からない構造が，a 方向に射影することによって明らかになりました．

今の場合はもともとの分布が2次元ですから別に1次元の空間に射影しなくても2

次元の散布図を吟味することによってデータの構造を実際に目で見ることができます．しかし，もっと高次元のデータですと，直接グラフに描いて目で見ることはできませんから，適当な低次元の空間を選んでその方向にデータを射影し，元の高次元データの構造を捉えようとします．例えば，国語，数学，理科，社会，英語の５科目の試験ですと，データは５次元ということになり，それらの合計点は（1，1，1，1，1）方向への１次元の射影と見なすことができます．この方向が意味を持つかどうかは議論の余地のあるところでしょうが，少なくともその方向が元の５次元データの特徴をうまく捉えているものであるかどうかの確認はなされてしかるべきでしょう（多変量解析と呼ばれる手法が使われます）．

さて，現在私は，自分の大学の他に，ある女子大の非常勤講師として統計学を教えています．全然別の授業内容を講義するのも面倒なので，上記の科目と同じ内容の講義をし，平常点も同じように付けて，同じ問題の試験をしました．その女子大の学生の「合計点」をパソコンを使ってヒストグラム化したのが図２．３（1）です．図２．２（3）と見比べてみてください．図２．３（1）でも二山ありますが，下のほうの山がぐんと小さくなっていることが分かります．



(1) コンピュータの出力　　(2) スケールを調整

図２.３　ある女子大における成績

ここで，重要な事に気が付きます．それは，図２．３（1）では横軸のスケールが８０から始まっていることです．最近のパソコンのソフトウェアはよくできていて，グラフが見やすいようにコンピュータが適当にスケールを調整してくれます．女子大のほうでは80点以下の学生がいなかったため，コンピュータが自動的にグラフの始まりを 80 としていたのです．これでは比較の際に困りますので，スケールを 0 から始まるように調整し直しました．それが図２．３（2）です．これを先ほどの図２．２（3）と比べなくてはなりません．いかがでしょうか．図２．３（2）の女子大のほうの二山は「できる」と「できない」ではなく「できる」と「ふつう」であったことが分かり

ます．こちらでは全員が単位を取得しました．

上の比較から，私の勤務する大学の学生と非常勤講師をしている女子大の学生とでは成績の差が大きいことが実証されたと思われるでしょう．同じ試験問題で試験を実施して点数の差が歴然と現われた，という客観的なデータが得られている訳ですから．しかしそう早合点はできません．実のところ，私の大学ではノート類の持ち込みは不可で試験したのに対し，女子大のほうでは，種々の理由もあり，ノート，配布プリント，参考書など何でも持ち込み可で試験を行ないました．点数がいいのは当然だったのです．

ここでの教訓は，データがどのように取られたのかに関する情報がない限り，結果の数値の解釈をしてはならないということです．統計の理論は，データの解析だけでなく，データをどのように取るのかの計画段階がきわめて重要であることを教えてもくれます．

3．統計的推測

前節では，得られたデータをどのようにグラフ化するか，そしてそれをどう解釈するかの例を見ました．しかし，統計学，特に近代の推測統計学では，現に取られたデータそのものの加工やグラフ化だけでなく，そのデータを生み出す真のメカニズム，あるいはデータが取られた母集団全体に関する情報を得ることが解析の目的であるとされます．研究対象の全体像をその一部でしかないデータから知ろうというのですから，そのためには，いかにデータを取るかの計画段階から，得られたデータから最大限の情報を抽出する技術，そして，データ解析の基となる哲学，思考法がきわめて重要なものとなります．それら全てに統計学が関わっているのです．

いくつかの例により統計的データ解析の実際を見ることにします．

例３．１． 女子学生の比率

私の勤務する成蹊大学では，在学生全体の約 1/3 が女子学生です．私は朝 1 時限目の授業が好きなので，好んで朝早くから授業をしますが，学生の出足が鈍いのが悩みの種です（自分も学生時代早起きが苦手だったことは既に棚に上げてしまっています）．そして，経験から，男子学生よりも女子学生の方が熱心に授業に出てきているという印象があります．これを統計的検定の枠組みで考えてみましょう．

調べたいことは，朝9時から授業に出てくる女子学生の比率を p としたとき，それが 女子学生の在籍比率 1/3 よりも多いか否かです．$p > 1/3$ であれば女子学生の方が授業への出席に関しては熱心ということになります．統計的検定では，まず帰無仮説と呼ばれる仮説（通常 H_0 と書く）をたて，それが正しいかどうかをデータを用いて検証します．また，帰無仮説が否定されたときに採用される仮説を対立仮説といい，通常 H1 と書きます．ここでの帰無仮説および対立仮説は，

$$H_0 : p = 1/3 \quad \text{vs.} \quad H_1 : p > 1/3 \qquad (3.1)$$

です．

校門のところに立ち，朝登校してくる学生の性別を調べることにします．このとき，次々に登校してくる学生が互いに独立である，すなわち，女子ばかりが固まってくるようなことはないという仮定が重要になります．小学校ならばいざ知らず，大学では集団登校ということもあまりありませんから，この仮定は満たされているとしても差し支えないでしょう．

全部で n 人調査し，その中での女子学生の人数を X とします．H_0 が正しいとしたとき，$n = 10$ として，$X = k$ となる確率 $P(X = k)$ および $X \geq k$ となる確率 $P(X \geq k)$ を示したのが表３．１です．

表３．１　10 人中の女子学生の人数の確率（p = 1/3 の場合）

k	P(X = k)	P(X ≧ k)
10	0.0001	0.0001
9	0.0003	0.0004
8	0.0030	0.0034
7	0.0163	0.0197
6	0.0569	0.0766
5	0.1366	0.2131
4	0.2276	0.4407
3	0.2601	0.7009
2	0.1951	0.5960
1	0.0867	0.9827
0	0.0173	1.0000

実際調査し，10 人中 5 人が女子学生だった場合，得られたデータでの比率は 0.5 と明らかに 1/3 よりも大きいのですが，それをもって $p > 1/3$ と結論するのは，やはりためらわれるでしょう．$p = 1/3$ であっても 10 人中たまたま 5 人が女子学生であるというのは十分考えられるからです．実際，表３．１より $P(X \geq 5) = 0.2131$ ですので，5 回に 1 回は任意に選んだ 10 人中女子学生が 5 人以上となることが分か

ります．しかし，調査した 10 人中 8 人が女子学生だった場合はどうも $p > 1/3$ らしいと結論できるでしょう．$P(X \geqq 8) = 0.0034$ と確率が小さいからです．

このように，帰無仮説が正しいとしたときにデータが実際の実現値よりも極端な値をとる確率を P-値（P-value）といいます．P-値が小さい時，帰無仮説とデータには整合性がなく，帰無仮説は誤りであると判断されます（これを帰無仮説を棄却するといいます）．統計的検定のロジックは背理法に似ています．数学では矛盾は100%矛盾ですが，統計では100%ということは期待できません．95%ならばよしとしようというのが多くの統計家の合意です．

10 人中 5 人では $p = 1/3$ は否定できませんでしたが，その理由は調査人数が少ないためです．$p = 1/3$ の下で，$k/n = 0.5$ を一定として n を増やして $P(X \geqq k)$ を求めたのが表３．２です．

表３．２　p＝1/3 のとき、いくつかの n に対して k/n ≧ 0.5 となる確率

n	k	P(X ≧ k)
10	5	0.2131
20	10	0.0919
30	15	0.0435
50	25	0.0108
100	50	0.0004

$n = 30$ とすると $P(X \geqq 15) = 0.0435$ と P-値が5%未満になりますので，30 人中 15 人が女子学生であれば帰無仮説 $H0 : p = 1/3$ は棄却されます．同じ比率 0.5 を与える場合でも，データ数が多いほうが証拠能力が高まり，仮説の真偽の判断が明確になります．その意味で，たくさんデータをとることが奨励されます．しかし，データ数を増やすと，測定の精度が低下したり種々のノイズが混入したりして，データの質が悪くなることが多いのには注意を払う必要があります．統計学の数学的な扱いでは，よくデータ数 n を大きくしたときの極限ないし漸近理論が問題としますが，実際問題では，データ数の増加に伴う種々の影響を無視した議論は無意味です．

例３．２． 缶コーヒーの開発

経済の大幅な右上がり成長が期待できない今，多くの企業にとって，ローコストかつ短期間で魅力ある新製品が開発できるかどうかは死活問題となっています．ここでは，統計手法が新製品開発に寄与した例をご紹介します．

190ml 入りの缶コーヒーの新製品を開発したいという話が私の許に持ち込まれまし

た．ちょうどその頃，品質管理，実験計画に興味を持っていたこともあり，お手伝いすることにしました．缶コーヒーの味を決める要因にはいくつかありますが，そのときは

Ａ：コーヒー豆の量（多，中，少）
Ｂ：砂糖の量　　　（多，中，少）
Ｃ：ミルクの量　　（多，中，少）
Ｄ：ある添加物　　（ある，なし）
Ｅ：豆の煎り方　　（深煎り，浅煎り）

という実験条件を設定し，それらの最適な組み合わせの決定を目的としました．５つの要因の異なる水準の組み合わせは全部で 3 × 3 × 3 × 2 × 2 ＝108 通りあります．しかし，これだけの数の試料を全部用意するのは時間とコストの制約上不可能で，実験実施グループからは，何とか20種類くらいに抑えたいとの要求がありました．また，コーヒー豆の量を多くして深煎りすると苦くて飲めたものではないのでその組み合わせの実験は行なわない，などという制約条件も付きました．その様なわがまま（というより現場の要求）を取り入れた実験計画を立て，得られたデータを適切に解析するのがここでの目的です．

「実験計画法」は，統計学の一応用分野でもある品質管理（Quality Control ＝ QC）の中で重要な部分を占める理論体系・方法論です．日本製品の世界に冠たる品質のよさは，企業の長年の品質管理，品質向上運動の賜物であるというのが世界の通説です，日本流の品質管理は，全社的な活動として，TQC（Total Quality Control）と呼ばれています．多大の成功を収めた日本の品質管理ですが，従業員ならば誰でも使えることを目指したため（悪いことではありません，いいことです），紙と鉛筆と電卓を前提とした方法論でした．ところが現在では，職場のそこここにパソコンが導入され，誰でもがコンピュータを使える環境が整ってきました．電卓とパソコンでは計算能力に天と地ほどの差がありますから，用いられる方法論も違って当然です．しかし，従来までの方法論でうまくやってきたという成功体験があるものですから，なかなか旧態然とした手法から抜け出せないでいます．そうこうしているうちに，特に米国では日本に学べとばかりに TQM（Total Quality Management）の名前の下，品質向上運動に力を入れ出しました（日本でも最近は TQC から TQM と呼び名が変わってきています）．

コンピュータを用いた品質管理技法の一つが，最適実験計画の自動的探索です．実験に取り入れる因子数，各因子の異なる水準数，実験回数，制約条件などをコンピュータに入力し，ある最適基準を満たすような計画をコンピュータに自動的に作らせるというものです．1960年代に理論的な結果が得られ，70年代にアルゴリズムの提案があり，80年代にソフトウェアとして実現され，90年代に入ってようやく実用化されて普及しつつあるものです．計算手段と統計手法とは密接な関係があります．

表３．３がコンピュータを使って得た「最適」計画です（岩崎，1992）．

表３．３　缶コーヒーの嗜好の実験のための計画

実験番号	コーヒー豆	砂糖	ミルク	添加物	煎り方
1	少	少	多	無	浅
2	少	多	少	無	浅
3	中	中	中	無	浅
4	多	少	少	無	浅
5	多	多	多	無	浅
6	少	少	少	有	浅
7	少	多	多	有	浅
8	中	少	多	有	浅
9	多	中	多	有	浅
10	多	多	少	有	浅
11	多	少	中	有	浅
12	少	多	多	無	深
13	少	少	少	無	深
14	中	多	中	無	深
15	中	少	多	無	深
16	中	中	少	無	深
17	少	少	多	有	深
18	少	多	少	有	深
19	中	少	少	有	深
20	中	多	多	有	深

さすがに一から十までコンピュータに任せるという訳にもいきませんから，出力してきた計画を専門的見地から検討しましたが，なかなかよい計画であるということになり，若干の調整後実際に実験に入りました．実験結果の解析についての詳細はここでは述べられませんが，なかなかの成功をみせ，従来よりも短時間に十分満足がいく結果が得られました．私は研究者ですので，論文を書く義務があります．そこで，この実験結果を論文にまとめることを提案しましたが，「こういういい結果を社外に出すべきでない」という実験の担当者の上司の判断があり，私は論文を一つ損しました．しかし，一統計家としては大いに満足しています．

例３．２． 新薬の開発と承認

最近の悲惨な薬害エイズ事件にみられるように，過去にも何度となく薬害事件が起こり，社会問題化しています．その際，薬剤の審査体制の甘さが批判の対象となりますが，どのようなプロセスによって薬剤が認可されるかは案外知られていないようです．ましてや，審査段階で統計が果たす大きな役割を知っている人はほぼ皆無でしょう．

新薬の申請認可では，動物を使った毒性試験，薬効薬理試験，また，患者に投与する臨床試験成績の結果等のデータを統計的に解析する必要があります．製薬会社側は自社の開発薬剤の有効性を主張したいでしょう．しかし，厚生省側はその効力に疑問を呈するかもしれませんし，医療保険支払側の言い分，医師の言い分，患者の言い分，果てはライバル会社の言い分など，実に多くの利害がぶつかり合います．そのようなときに判断の基準となるのは，「声の大きさ」ではなく，実際のデータとそこから導かれる客観的な結論です．データ解析の方法論たる統計学の重要性は論を俟ちません．実際，欧米では医薬分野で働く統計・生物統計の博士号を持つ専門家が数多くいます．日本でもその数は増えつつありますが，まだ決定的に不足している現状です．

新薬を開発したメーカーは，毒性試験，臨床試験成績などをまとめた資料を厚生省に提出します（申請段階では「薬」とは呼べず単なる「物質」ですが）．厚生省のヒアリングをすませたのち，中央薬事審議会の新薬調査会に審査が付託されます．調査会で専門の委員が審査し，疑問点などを開発メーカーとやり取りして認可すべきかどうかを判断します．調査会段階をパスしますとさらに特別部会に上程され，再度審査を受けます．このように何段階もの審査を経たのちようやく認可され，発売にこぎ着けるという訳です（1997年7月から厚生省の組織変えに伴い審査体制が若干変更になりました）．各審査段階で却下される薬剤も少なくありません．却下されますと薬として販売することはできませんから今までの苦労が水の泡となります．

新薬の調査会は扱う病態ごとにいくつかのものに分かれています．例えば第二調査会は循環器系の疾患に対する薬剤を専門に審査し，その委員構成は表３．４のようになっています．

表 3.4　中央薬事審議会第二調査会の構成

基礎部門：製剤　1，毒性　2，薬理　1，吸収排泄　1，統計　1

臨床部門：大学教授クラスの臨床医師　9

事務局　：厚生省技官など　4

統計の専門家がメンバーに加わっている点に注目して下さい．調査会の仕事は，審査

薬剤の有効性，安全性の科学的評価です．よく，政府関係の審議会は役人の作文の承諾機関と化しているという悪口を聞きますが，新薬の調査会ではそんなことは一切なく，専門の委員の意見が尊重され，厚生省側は事務局に徹しています．その意味で，決定に関する全ての責任は各委員が負うことになります．製薬メーカー側の誤った統計解析を見過ごしたりしますと，当然統計の専門委員の責任が厳しく問われます．

次に，１つの薬剤を開発するに当たり，どのような段階を踏み，どのくらいの時間がかかるのかを示したのが表３．５です．

表3.5　ある薬剤の開発経過

新規物質の発見・合成（1981）
規格安定性試験（1984～1993）
毒性試験（1985～1992）
薬理試験（1984～1993）
吸収・排泄試験（1985～1992）
各種臨床試験
　第Ⅰ相臨床試験（健常男子）（1987～1988）
　第Ⅱ相臨床試験（患者投与）（1988～1991）
　第Ⅲ相臨床試験（二重盲検比較試験）（1991～1993）
　一般臨床試験（1991～1993）
申請（1993）
審査（1993～1995）計６回
認可・発売（1996）

きわめて多くの段階と長い時間がかかることがお分かりでしょう（薬剤となり得る新規物質の発見までの時間を加味するとさらに期間は長くなります）．開発費用は一説に100億円とも150億円ともいわれますが，うなずける数字です．

そのような長時間と莫大な費用をかけた開発の最終段階として表３．６のような結果が出ます．これを「科学的に」判断しなければなりません．

表3.6　ある薬剤の臨床試験の結果

有効性	有効	無効	計	有効率
治験薬	１０４	２４	１２８	８１．３％
対照薬	１０６	３１	１３７	７７．４％
計	２１０	５５	２６５	

安全性	安全	問題	計	安全率
治験薬	１０９	１９	１２８	８５．２％
対照薬	１２４	１３	１３７	９０．５％
計	２１０	５５	２６５	

表3．6は全部で 265 人分のデータです．治験薬というのは，新しく開発した薬，対照薬というのは，現在市販されていて臨床的評価が確立しているものです．新薬は何らかの意味で対照薬に比べてよくなくては認可に至りません．有効性をみますと，有効率が対照薬の 77.4% に比べて治験薬は 81.3% ですから，勝っているといえるかもしれません．しかし，安全性は，対照薬の 90.5% に対し治験薬は 85.2% ですから負けています．「この」臨床試験の結果をどう解釈すべきでしょうか．

この臨床試験の他に他の臨床試験もありますから，実際にこの薬剤が投与された患者数はもう少し多くなります．しかし，この薬剤は高血圧の人の血圧を下げる降圧剤で，高血圧と診断される患者数は日本全体で何百万人もいますから，それらの人達に投与したらどうなるかを，多くてもわずか数百人のデータから判断しなくてはなりません．統計的方法が重要な役割を果たすことがお分かりでしょう．

統計的検定を適用すると，

$$H_0：薬剤間に差はない \quad vs. \quad H_1：新薬は対照薬に勝る$$

の検定を行なうことになります．例3．1と同じような検定を施してみますと，実現値以上に極端な値をとる確率，すなわちP-値は，有効性が 0.638，安全性が 0.474 と，「差はない」という帰無仮説を棄却するだけの根拠に欠け，有効性においても安全性においても両薬剤間に有意の差は認められませんでした．検定で帰無仮説が棄却されなければその帰無仮説は採択されしたがって両薬剤は同等，という結論を導かなくてはならないというのも統計学の教えですが，薬剤に関する基礎，臨床の様々な面を考慮した結果，本薬剤は既存薬に比較して同等以上であるとの判断がなされ，認可になりました．

4．最近の話題

ここでは，私自身が最近興味を持っている話題についてお話しします．まず，統計でよく出てくる正規分布について．正規分布（normal distribution）は平均値 μ と分散 σ^2 で特徴づけられる連続型の分布で，通常 $N(\mu, \sigma^2)$ と表わされます．特に $N(0, 1)$ は標準正規分布と呼ばれ，その確率密度関数は

$$f(x) = (2\pi)^{-1/2} \exp(-x^2/2) \qquad (4.1)$$

で与えられ，その形は図4．1のようになります．

正規分布に限らず連続型の分布では，ある区間内の値を取る確率は，確率密度関数の

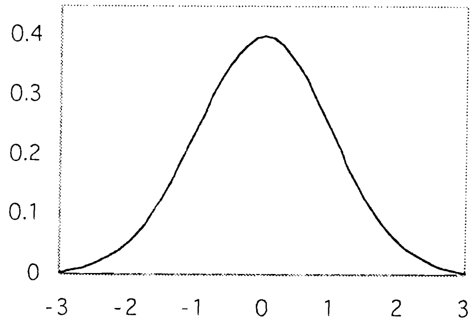


図４．１　標準正規分布の確率密度関数

その区間での積分で与えられ，例えば表４．１のようになります．（4.1）の関数は不定積分が初等関数で表現できない典型的な例で，したがって積分値を求めるのに数表が用いられてきました．しかし，最近ではパソコン用の表計算ソフトにも正規確率を計算する関数が組み込まれていて，簡単に積分値を得ることができます．時代の進歩というほかはありません．

表４．１　標準正規分布の確率

区間	-1.5 以下	(-1.5, -0.5)	(-0.5, 0.5)	(0.5, 1.5)	1.5 以上
確率	0.07	0.24	0.38	0.24	0.07

表４．１で示された各区間の確率に見覚えはありませんか．これは中学校などでの５段階評価の成績表の１～５の割合になっています．このことは，人間の能力のばらつきは正規分布に従うと暗黙裡に仮定されていることを意味しています．

人間の能力は１次元的なものでなく，多くの要素が複雑に絡みあった多次元的なものと解釈するのが妥当でしょう．そのような多次元的なものを１次元の尺度で表わすと（１次元空間に射影すると）正規分布になるということに他なりません．では，その多次元的なものは，その多次元空間でどのように分布しているのでしょうか．統計学では，多変量正規分布というものがあって，多変量正規分布の１次元空間への射影はまた正規分布になることが証明されています．

しかし，射影して正規分布になるのは多変量正規分布に限りません．Diaconis and Freedman（1984）は，高次元空間にある点が，次の２条件

条件１：原点からの距離がほとんど同じ

条件２：互いにほぼ直交する

を満たすとき，任意の方向への１次元の射影は，もとの点の次元が十分高くかつ点の

数が多いとき概ね正規分布になることを示しました．すなわち，得られたデータが正規分布であっても，その元となる高次元では，多変量正規分布にとらわれないもっと自由なモデルを考えてもよいという可能性を表わしているといえます（岩崎，1997）．

例４．１． 高次元超立方体の頂点の射影

上で述べた Diaconis and Freedman の条件を満たすもっとも簡単な例が高次元超立方体の頂点の集合です．図４．１は２次元超立方体（という大げさなものでなく正方形）のある方向への射影を表わすものです．

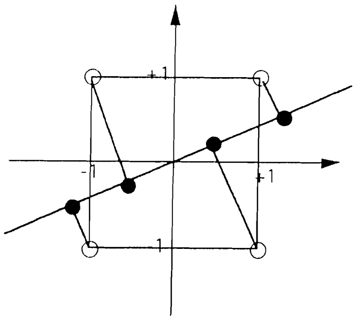
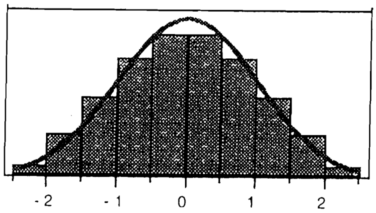
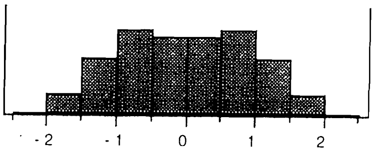


図４.２　２次元立方体（正方形）の各頂点の直線への射影

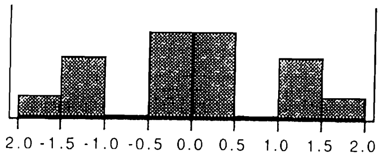
この場合は点が４個しかありませんから正規分布かどうか判定することはできません．図４．２は７次元超立方体の128個の頂点を別の方向に射影したものです．（1）の射影はほとんど正規分布です（曲線は正規分布の確率密度関数）．射影の方向を変えると（2）や（3）のような射影が得られます．



(1) 正規分布に近い射影



(2) 中心がややへこんだ射影



(3) ３つのグループに分かれる射影

図４.３　７次元超立方体の各頂点（128個）の直線への射影のいくつか

私はまだ，どうして図４．２（3）のように頂点が３グループに分かれるのか，その方向はどのような方向であるのか，はっきり理解ができていません．並の人間の能力では，高次元区間を思い浮かべるのは大変に困難です（高次元空間が見えないといい数学者にはなれないのでしょうけれども）．

高次元立方体の頂点はアンケートなどの質問紙への回答と密接な関係があります．例えば，

Q１：あなたは数学が得意ですか？　　　　Yes，No

Q２：あなたは黒板の字がきれいですか？　Yes，No

Q３：あなたは安室奈美恵が好きですか？　Yes，No

という３つの質問への回答を考えてみましょう．Yes を +1 に，No を -1 に置き換えれば，質問への回答者は，原点を中心にして１辺が 2 の立方体の 8 個の頂点のいずれかと１対１対応がつきます．そして，各頂点のある直線への射影は，その方向を表わすベクトルを $a = (a^1, a^2, a^3)$ とすると，+1 あるいは -1 の値をとる変数 x^1，x^2，x^3 の重み付き和 $a^1 x^1 + a^2 x^2 + a^3 x^3$ となります．各頂点の射影の分布は多くの場合正規分布になる，というのが Diaconis and Freedman の結果です．多くの学力試験などではそうなっているのでしょう．あるいは，例２．２で示したように，その射影は正規分布とは似ても似つかないようなものになるかもしれません．いずれにせよ，このような観点から個人差とその全体の分布を捉えるのは興味ある研究といえます．ちょっと考えてみようと思っています．

５．統計学の今後

これまでみてきた例は，統計学の幅広い応用分野のほんの一部です．標語的にいえば，データのあるところ統計学有りです．実際，計量経済学（econometrics），計量心理学（psychometrics），計量生物学（biometrics），計量社会学（sociometrics）など，計量・・学（・・metrics）と呼ばれる研究分野は多岐に渡っていて，それら全てに統計的方法論が重要な役割を果たします．これらに加え，環境問題などでも不可欠なものとなっています．適用分野によって個々の方法論には若干の違いはありますが，良質のデータをきちんととり，そこから最大限の情報を引き出すという目的は同一です．

冬季五輪のメダル数の推移を示した図２．１をもう一度見てください．このグラフから長野オリンピックでのいくつくらいメダルが取れそうかという予測は，コンピュー

タに数値を入力し，ある種の統計手法を使えばできないことはありません．しかし，そのようなコンピュータによる予測では，例えば1972年の札幌オリンピックでの金，銀，銅を独占したときの興奮，最近のオリンピックでの感動的な選手達の偉業などは一切反映されていません．

「データ」と「数値」とは全然違います．データには表面的に現われる数値以外のもろもろの情報を併せ持っています．そのようなデータの背後にある多くの事柄を読み取らない限りよいデータ解析はできません．コンピュータにできるのは，「データ解析」ではなく，単なる「数値解析」です．コンピュータは確かに有力なツールですが，コンピュータさえあれば統計解析は OK というのは大きな間違いです．

このように重要な役割を果たす統計学ですが，悩みの種は，どこの分野でもそうでしょうが，人材不足です．私見では，統計学に向く人材とは

(1) 何にでも興味をもつ旺盛な好奇心がある

(2) 論理的思考に富むが，突飛な発想もできる

(3) 数学が得意といわないまでも苦手でない

といった性向を併せもつ人です（そういう贅沢をいうから駄目なんだ，そもそも自分はどうかという陰の声が聞こえてきそうです）．日本人全体の中で数学が得意という人は，残念ながら，少数派です．そして，数学が得意という人の中で統計に興味をもつ人も，さらに残念ながら，少数派です．しかし，私はあきらめていませんし，あきらめてはならないと思っています．我こそはと思う若い人が出てくることを期待しています．

参考文献

Diaconis, P. and Freedman, D. (1984) Asymptotics of graphical projection pursuit. Annals of Statistics, 12, 793-815.

岩崎　学 (1992) コンピュータ指向型データ解析の新手法．行動計量学，19，37-49.

岩崎　学 (1997) 多変量正規分布と評価尺度の非直交性．科学研究費シンポジウム講演予稿集．