

医学統計解析

柏木 力 著

朝倉書店

医学統計解析

柏 木 力

朝 倉 書 店

著者略歴

大正12年 横浜に生まれる
昭和21年 東京大学医学部卒業
昭和21年 東京大学医学部附属病院
昭和22年 気象研究所衛生気象研究室
昭和36年 名古屋大学環境医学研究所、現在
に至る。教授・医学博士

統計ライブラリー

医学統計解析

昭和54年2月10日 初版第1刷
昭和55年11月10日 第3刷

著者 かしわ 柏 き 木 つとむ 方
発行者 朝 倉 鑛 造
東京都新宿区新小川町2-10

発行所

株式
会社

朝 倉 書 店

東京都新宿区新小川町 2-10
郵便番号 1 6 2
電話 東京 (260) 0141(代)
振替口座 東京 6-8673 番
自然科学書協会会員

© 1979

政弘印刷・渡辺製本

無断複写・転載を禁ず

3047-400010-0032

序

本書は統計学の概説のための入門書ではない。そうではなくて、私が医学の研究にたずさわって以来の個人的な体験と、そこから得られた知識を基準にして、日常の医学研究において最も繁用されると判断された統計学の手法と、そこにおける基礎の考え方を述べたものである。

私が大学を終えて入局した東京大学の物療内科には、当時増山元三郎先生が講師として、気象と生体、病気との関係を研究しておられ、折にふれ、故鳥居敏雄教授やすぐれた先輩である高橋暁正、土肥一郎両氏の驥尾に付して、統計学を学ぶ機会に恵まれた。もっとも当時の増山先生は、難しい理論的な側面に流れすぎたという意味の、先生ご自身の反省を、ごく最近東京大学衛生学教室の畏友山本俊一君を通して間接的に聞いて、私の理解のたどたどしさと、不十分さとにわずかながら一種の安堵感を覚えたことも否定できない。

がともかく、この小著はそのような私が私なりに統計学を理解して行ったあとをなぞったものにすぎない。その意味で、不必要と思われたところは思いきって省略し、また専門の統計学者からは大胆すぎるモデルの単純化と現実の状況へのあてはめと非難されるかもしれない。また多分に实际的 (practical) に過ぎる恐れを感じないでもないし、何よりも高橋、土肥両先輩の二つの著書に何ほどのものを付け加えることができたかという危惧の念を禁じ得ない。

同時に、D. Mainland は統計学の cooking cards を厳しく排除し、統計的思想の普及をこそ常に心がけていたように、増山先生の教えもそこにあると思うのだが、本書を通じて私の意図したところも同じであって、統計的な考え方、思考過程が医学研究のあるべき段階、過程に貫かれることを念願とした。

最近の傾向は、かつて 20~30 年前増山先生らが医学研究で模索してこられた道程を改めて、今度は医学研究に必ずしも“首までつっこんでいる” (Mainland) とは思えない人々の主導の下に繰返されているのが目に着くし、その

ことよりもいわゆる公害裁判や、環境基準の設定に統計学を悪用しているという現象も顕著である。このことに関して、日本における統計学の在り方については、『統計学の未来』（東大出版会，1976）が，そのような意味で参考になるう。

終りに，あらゆる機会にご教示を頂いた東京大学工学部 奥野忠一教授に深甚な謝意を表すると同時に，ともすれば鈍り勝ちの私のペンの運行を督励して下さった朝倉書店編集部各位に感謝したい。

1979 年 1 月

柏 木 力

目 次

1. 医学における統計学	1
2. 統計学的推論	4
2.1 データの性質	4
2.2 母集団と標本と	5
2.3 確率化の方法	6
2.4 2種類のデータ	7
2.5 分布の型	8
2.5.1 分布の性質としての分布型	8
2.5.2 分布のとり方	9
2.6 正規分布のあてはめ	12
2.6.1 正規分布のあてはめ	12
2.6.2 対数正規分布のあてはめ	13
2.7 統計学的推論	16
3. 連続量の取扱い	20
3.1 基本的な統計量	20
3.2 平均値の問題の所在	22
3.3 計 算	23
3.4 誤差の法則, 2種類の誤差	26
3.5 平均値の差の検定における分散	28
3.6 分散の比較	29
3.7 二つの母平均の差の検定	31
3.7.1 母分散共通の場合	31
3.7.2 母分散が異なる場合	32

3.8	母平均と規準値との比較	35
3.9	片側検定と両側検定と	36
3.10	処理前後の平均値の比較	39
3.11	平均値に関する推定	41
3.11.1	母平均の信頼区間	42
3.11.2	母平均の差の信頼区間	44
3.12	データを棄却する方法	44
3.13	若干の注意事項	47
4.	離散量の取扱い	50
4.1	問題の所在	50
4.2	母出現率と規準値との比較	51
4.2.1	n 個のうち r 個以上に出現する場合	52
4.2.2	n 個のうち r 個以下に出現する場合	53
4.3	二項分布の正規分布による近似	58
4.4	二つの母出現率の比較	60
4.4.1	χ^2 -検定	61
4.4.2	χ^2_0 計算の別法	62
4.4.3	正規近似による検定	63
4.4.4	連続修正	64
4.4.5	直接確率計算法	65
4.4.6	一致度の検定 (McNemar の検定)	69
4.5	いくつかの母出現率の比較	71
4.5.1	期待度数による χ^2_0 の計算	72
4.5.2	χ^2_0 計算の別法	73
4.5.3	尤度比法による χ^2_0 の計算	74
4.5.4	標本の性質による差異	76
4.6	母出現率の推定	77
4.6.1	方 法	77

4.6.2	数表の利用	79
5.	2変数間の相互関係	81
5.1	問題の性質	81
5.2	相関係数	81
5.2.1	相関係数の性質	81
5.2.2	相関係数の計算	83
5.2.3	相関表を用いたときの相関係数の計算	85
5.3	相関係数に関する検定	89
5.3.1	母相関係数 $\rho=0$ の検定	89
5.3.2	その他の検定	91
5.4	相関係数に関する推定	92
5.5	回帰関係	93
5.6	回帰直線の成立ち	95
5.6.1	母回帰直線と標本回帰直線と	95
5.6.2	標本回帰直線の構成	96
5.6.3	標本回帰係数の求め方	97
5.6.4	最小2乗法の原理	97
5.7	回帰係数に関する検定	98
5.7.1	回帰係数の有意性の検定	98
5.7.2	二つの母回帰係数の比較	100
5.8	回帰係数に関する推定	104
5.8.1	母回帰係数 β の信頼区間	104
5.8.2	回帰直線の信頼区間	105
5.9	回帰直線に対する直交多項式のあてはめ	107
5.9.1	直交多項式	107
5.9.2	n 点に2次式をあてはめる場合	111
5.9.3	n 点に3次以上の式をあてはめる場合	112
5.10	指数曲線のあてはめ	117

5.10.1	$y = L\{1 - e^{-r(X-a)}\}$	117
5.10.2	$y = Ae^{rx}$	121
5.10.3	$y = L/\{1 + e^{-r(X-a)}\}$	123
5.11	算定曲線による S 字状曲線の線形変換	128
6.	実験計画法	130
6.1	実験計画法の意義	130
6.2	実験の意義	131
6.3	実験における二つの条件	133
6.4	基本的な解析の方法	135
6.4.1	誤差分散の計算	135
6.4.2	処理の効果の計算	136
6.4.3	全体の測定値の変動	137
6.4.4	平方和の簡単な計算法	137
6.4.5	実測値のモデル	138
6.4.6	仮説検定	139
6.4.7	F-検定	140
6.4.8	分散分析表	141
6.4.9	級間の差	142
6.5	個体差の検出	143
6.6	交互作用	145
6.7	母数模型と変量模型と	147
6.8	実験計画法の型	149
7.	多変量解析	150
7.1	多変量解析の意義	151
7.2	多変量解析の応用	153
7.3	医学研究に用いられる多変量解析	154
7.4	主成分分析	155

7.4.1 幾何学的意味	155
7.4.2 主成分分析の解法	156
7.4.3 例 解	159
7.5 判別関数法	164
7.5.1 判別関数の概念	164
7.5.2 判別関数の導き方	165
7.5.3 判別関数の検定	167
7.5.4 新しい標本の判別	168
7.5.5 誤って判別する確率	168
7.5.6 検査法の追加, または削減するための基準	170
7.5.7 多群の判別	171
7.5.8 例 解	173
8. 分布の型によらない方法	177
8.1 方法の特徴	177
8.2 方法の選択	178
8.3 二つの標本の比較 (Wilcoxon 検定)	179
8.4 いくつかの標本の比較 (Kruskal-Wallis の H -検定)	184
8.5 対応のある場合の 2 標本の比較	188
8.6 関連性の問題	192
8.6.1 Spearman の順位相関係数	192
8.6.2 Kendall の順位相関係数	193
8.6.3 数値例	195
9. より高度の解析に関する指針	198
例題一覧	201
数値表	205
索 引	233

Ⅰ。 医学における統計学

医学研究の特殊性のひとつは、いうまでもなく対象あるいはその場が有機的な(organize された)統一体にあるということであろう。その統一体をどのような水準でとらえるかによって、それぞれその研究が明らかにしようとしている真実のあり方が異なる。したがって何が単なる現象の記載であって、そこをいかように昇華させれば本質に到達することができるかということは、その研究の水準により本来多様であるべきであって、ある病気の原因が一つなりいくつかの物質、分子の運動形態の変化に帰せられなければその病因が解明されたとは言い難いということは当を得ていない。

他方、今日 Pavlov 流の考え方が必ずしもわれわれを首肯させえない原因のひとつは、かれらの実験における誤差成分をすべてあるいは一部分“大脳皮質(関与)”要因と帰結してしまうところにある(もちろん、ソヴィエト学者の、“パヴロフ、ないしセチェーノフの名において”という事大主義の故もあろう)。

Organize された統一体という条件を保持する最低の条件は、血液の供給と、神経の支配とであると考えられるが、このことは当然神経系の反応性、血液供給の状態の個々の対象における差異をもたらす。いわば個々の対象は、個性(個体差)を内包している。

このような条件にある対象に対して、あるひとつの処理を施したとき、それに対する反応が正の方向に出現したというだけでは、それはその限りのことであって、別の時期に同じ個体に同じ処理を与えたときも、同質の反応が得られるとは限らない。これではこの処理はこの対象に対し正負いずれの反応を惹起すべきものか、確定的なことはわからない。

ここに統計的思考の必要性が認められるのであって、それなくしてはまさし

く“Einmal keinmal”といわれても仕方がないのである。

統計的な思考は 個体差に代表されるようなもろもろの誤差を管理 (control) する方策を与える。それは誤差を排除して、0 にしてしまう——神や Hitler でなければ到底不可能であるが——のではなくて、ある目的とする要因の発現の仕方に対して、全体として偏りや傾向性を無にするという意味の管理である。

この誤差成分といわれるものも、その内容は相対的なものである。ひとつは目的とする要因の効果に対しての誤差という意味で相対的であると同時に、見方を変えれば、誤差成分のある部分はまた他の部分に対して、ひとつの要因に変わるという意味でもまた相対的である。固定的にものを考えるのではなく、何よりも柔軟な態度で、flexible にものに接することこそ要求される。

Organize された統一体のもつもう一つの重要な特性は、自己調節性、自己修復性ということであろう。それゆえ、ある処理、たとえば薬剤の治療効果は少なくともこの自己調節性による効果を上回るものでなければならない。当然ここに対照の意味の重要性のひとつがあるのであって、対照の設置の仕方も、統計学は与えられた条件と目的との下で、適切な処方を下す。

このような統計的思考の過程を経て到達する研究の個々の段階における判断は、“正である方向の反応が得られる可能性がきわめて高い”という、いわば蓋然的 (plausible) なものである。それは“命題 A が存在すれば、必ず B を帰結する”という断定的な、あるいは形式的な判断のあり方とは大いに相違する。既に 25 年前、Polya がその著、“Mathematics and Plausible Reasoning” (邦訳、『数学における発見はいかになされるか』) の中で鋭く指摘しているように、両者の判断のあり方を混同することは許されない。

以上述べたような観点から、医学研究、特に臨床医学研究において、比較的繁用される手法を以下の各章にわたって紹介するが、Mainland もいうように、統計的思考を身につけていれば、簡単な手法で驚くべきほど豊富な成果をあげることができる。それゆえ、生体の機能が多面的であり、相互に依存しあっているからといって、たとえば多変量解析の方法を安易に援用することは好ましくない。それだけ満たすべき前提条件が多数になり、それだけ実施の困難さが

増す。

このことは一般に統計学的手法を適用する上で常に意識しておかなければならないことであって、機械的な適用は最も忌むところである。明確な研究の目的に適った実験なり試験なり調査なりの計画の下に、実施があって、初めて得られたデータの解析が意味をもち、この過程を経てのみ前に述べたような性格の結論に到達することが保証される。

2. 統計学的推論

2.1 データの性質

医学研究，特に臨床医学におけるそのあり様を，一つの例に即して考えよう．いまある種の周期性神経筋疾患の病理を明らかにしたいと考えて，当面の仮説を，おかされた筋肉の解糖系の酵素群の活性は，正常のそれに比較して低下しているのではないか，ということにおいたとしよう．実験的に動物でこの疾患のモデルを作製することは困難なので，この仮説を検証するためには，勢い臨床的に患者から材料を採集して諸酵素活性を測定しなければならない．当然，健康者からの材料も対照として採集しなければならないこともいうまでもない．

対照としては，この場合健康者の他に，神経疾患，筋疾患，代謝性疾患，内分泌疾患等を除外した他の患者からの材料をあてることもありえよう．

表 2.1 データの形式

患者群	酵素 A	酵素 B	...
被検番号			
1	x_{A1}	x_{B1}	...
2	x_{A2}	x_{B2}	...
3	x_{A3}	x_{B3}	...
$\vdots$	$\vdots$	$\vdots$	...
対照群			
101	y_{A1}	y_{B1}	...
102	y_{A2}	y_{B2}	...
103	y_{A3}	y_{B3}	...
$\vdots$	$\vdots$	$\vdots$	...

そこでできるだけ精力的にこの疾患の患者を集めることになるが，多くの場合，患者は受動的にしか得られず，患者が向うからやってくるのを待つことになる．たとえその疾患が家系的に，あるいは地域的に集積が認められていて，その場所に調査に出かけ材料を得たとしても，基本的な状況は変わりようがな

い、

このようにして得られた材料から適当に筋肉を処理してそれぞれの酵素活性を測定することになる。そうすると初めて表 2.1 のような形式のデータが得られる。

いまこのような形のデータの解析法を示すのが目的ではなく、ここではこのようなデータは、そもそもどのような性質、背景をもったものであるかを問題にする。

このデータは、ある種の神経筋疾患の病理を明らかにするために集められたものではあるけれども、明らかにさるべき病理は、このデータを提供した個々の患者に限って適用されることで足りるとしているのではないことはいうまでもない。これこれのデータからしかじかと結論することができると、原則的にこの疾患の患者すべてに適用されることを本来意図していたはずである。

例外はもちろんないわけではない。症例報告という観察記録が、学会でも、雑誌にも発表されることが多い。これは一般に珍しい病気、いままで知られていなかった症状を呈した既知の病気、あるいはごくまれなことであるが全く新しい独立の疾患単位等の観察報告であるが、これは統計調査でいう典型調査に類するものといえる。

われわれが集めたデータは本来このような意味のものである。

2.2 母集団と標本と

上に述べたような意味をもつデータが生れ出てくる構造からいえば、この神経筋疾患の無数の、過去から未来にわたる患者の集団があって、実際にデータを提供した現実の患者はその集団を代表する一部分である。この関係を統計学ではもとの集団を母集団 (population) といい、われわれの眼前の患者、あるいはそのさまざまな測定値——この例では諸酵素活性値——自体は標本 (sample) として取扱い、標本から得られた結論を母集団全体に適用する、あるいは標本について得られた知識から母集団の特性を推論するという過程をとる。

このことは一般に科学認識のあり方というべきものと対応しよう。もっとも、現象と本質との関係に対する見方は人によってさまざまであろうが、われわれが直接知ることができるのはあくまで現象についてであって、現象の背後にあって、それを貫いて存在する本質、あるいは法則性ともいうべきものは、直接には知るよすがはなく、ただ現象

を手がかりにしてのみ把握できるとする。それはさまざまな形で本質が現象に反映しているからであると考える。このように母集団と標本との関係はまた、本質と現象との関係として対応づけられるといえる。

さて標本から母集団を推論できるためには、標本が正しく母集団を代表していなければならない。正しくという意味は、標本が母集団全体の正確な縮図であるということである。そのように標本を構成するために、母集団から確率的に、または、作為なしに標本の成員を抽出する。この操作を確率(無作為)抽出(random sampling)と呼び、この場合の標本を確率標本(random sample)という。

確率標本によって母集団を正しく代表させることができると同時に、標本から母集団について推論を行う際に、誤った推論を行う危険を確率的に制御することが可能になる。具体的にいえば、誤って推論する場合は、全体のたとえば高々5%であるというように。

2.3 確率化の方法

確率化は、確率標本を得るためだけではなく、動物実験において、ある複数の種類の処理をいずれの群の動物に割りつけるか、また実験ないし測定の手順をどの群のどの動物から始めるべきか、あるいは薬効試験において試験薬と対照薬とを患者にどのように割りつけるか等、さまざまな局面で重要な意味をもっている。

このような確率化を行うためには、一般に乱数表が利用される。簡単には、乱数サイによるのがよい。

1° 乱数表： 乱数表とは0から9までの数字を確率的に配列したものである。乱数表の任意のページをあけ、目をつぶって鉛筆を落とし、最も近い2桁の数字を読む。 n が2桁の大きさならば、読みとった表の数字を n で割ってその余りでおきかえる。

たとえば $n=25$ 人の被検者(1~25の通し番号がつけてある)の順序を確率化したいとき、上の方法で、28, 83, 07, 41, 86, ..., の数列が得られたならば、これを25で割った余り、3, 8, 7, 16, 11, ..., の数列におきかえ、これを被検者番号として、順序を確率化する。重複はとばす。

2° 乱数サイ： 正20面体サイコロ(図2.1)に、0～9の数を2面ずつに記入したものである。赤、黄、青の3個1組になって販売されており、たとえば赤を100位、青を10位、黄を1位と定めると、3桁の乱数が得られる。これを1°の場合と同じように用いる。

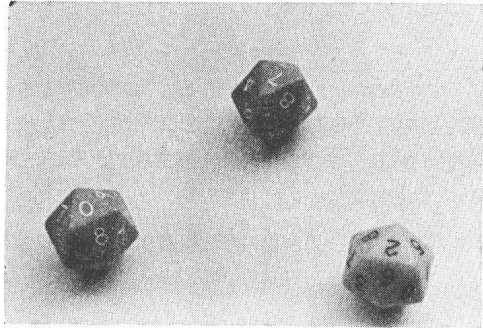


図 2.1 乱数サイ

2.4 2種類のデータ

§2.1で述べたようにデータそのものは統計学でいう標本値である。そしてデータは通常2種類に大別することができる。

一つは、1, 2, ..., n と、被検者なり動物なりの頭数を数えるような場合のデータである。たとえば、ある病気の一群の患者50人のうち、症状Sを示した患者は35人であったとか、あるいは満6歳の学童100人のうち、身長が130cmをこえる者は30人であったという場合である。またサイコロを100回投げたとき、1の目が出た回数は15回であったというような場合である。このようなデータを計数データといい、これを一つの量と見たとき、計数值、より一般には離散量という。

この型のデータを扱う場合には、上例でわかるように、50人のうち35人が症状Sを呈していた、すなわちこの症状の陽性率は70%であったとか、100回のうち16回、すなわち16%に1のサイコロの目が出たとかのように、出現率の問題としてとらえるのが一般である。

しかし、病理学上の組織学的所見を表現する場合、ある病変、たとえば白血球の浸潤の強度を(卅)、(卅)、(+), (±), (−)等と記載する場合がある。これに4, 3, 2, 1, 0という評点を与えて、部分的に計量化する場合がある。また、ある症状Sの有無を、存在する場合は1とし、欠く場合は0とするという取扱い、2進法的な取扱いをしているのであって、ある種の機械診断