

## 医学研究における統計学の役割

川崎医科大学 数学教室

仮谷 太一

(昭和60年9月10日受理)

The Role of Statistics  
in medical science

**Taichi Kariya**

*Department of Mathematics,  
Kawasaki Medical School,  
Kurashiki 701-01, Japan*

(Received on Sept. 10, 1985)

### 概 要

現象の理論的解明を目指す精密実験科学の実験と、医学研究の最終段階である治療医学の実験とは、同じく実験研究とは云つても、事情が根本的に異なることを簡単な例を用いて示し、個体差に正面から取り組まねばならない治療医学では、統計学は必須であることを述べる。

次に、統計理論における母集団と標本の概念は、そのままでは実験母集団と標本データにはあてはまらないことから、実験研究における統計的推測の問題点をあげ、特に医学実験研究において留意すべき点に触れる。最後に、統計学と医学との協力によって成功した実験研究の例をあげて、医学研究における統計学の役割を明らかにした。

### Abstract

We illustrate that experiments in the exact experimental science which aims at the theoretical explanation of phenomena are quite different in characteristics from ones in the therapeutic medicine which is the final stage of medical studies. Statistics is essential in the medical studies which must face difference among individuals.

The concept of population and sample in general statistical theory does not apply to the case of medical data because the experimental population corresponding to the experimental data is originally hypothetical. Accordingly the statistical inference based on them does not always be reliable and we should pay attention to several points which will be shown later. Lastly we explained the role of statistics in medical science by citing the examples which were completed under the cooperation of statistics and medicine.

#### 1. 個体差が問題になるところでは、統計学は必須である

まず簡単な設問から始めよう。

[問題] 次の各集団のメンバーが、それぞれみんな流行性感冒にかかり、発熱・筋肉痛などの

同じ症状を示している。このとき、その治療のため、集団の各メンバーに対し、みな同量の抗生物質を解熱剤といっしょに服用させればよいだろうか。

〔集団1〕赤ん坊から、幼年・青少年・壮年・熟年・寝たきり老人にいたる種種さまざまな人達の集団の場合。

評判のよい大きな個人開業医の待合室を想定してみればよい。年齢差・体重差・男女・虚弱頑健・先天性心疾患の有無・手術後または半身不随で寝たきりなど、個体差の大きな患者達に対して、みな同じ処方ではよいだろうか。

〔集団2〕25歳から29歳までの男子の集団の場合。

ほぼ同年齢の男子である農林漁業者・工場労働者・商業従事者・会社員・公務員・医師・研究者など。

性差はなくなり、年齢差も小さくなったが、職業・基礎体力・自然環境・社会環境・生活環境・病気に対する抵抗力などについては、やはりかなりの個体差があるであろう。同一処方ではよいだろうか。

〔集団3〕25歳・身長165~175cm・体重65~70kgの男子工具で、全員寮生活をしている人達の集団の場合。

体格差・環境差は、かなり少なくなったが、なお、基礎体力・遺伝的性質・免疫の有無・楽天的な性格か悲観的な性格かなどの違いはあるであろう。処方にさじ加減が必要であろうか。

〔集団4〕同一の自然的・社会的環境におかれているコピー人間（クローン人間）の集団の場合。

全く仮想の人間集団であるが、この場合には全員均質で個体差はなく、治療条件も全員同一である。従って、物理化学の実験におけると同様、実験材料は均質、条件も全部同一だから、治療に最適のものが存在するはずである。すなわち、最適処方があるはずである。

個人差が大きく、治療に関係する要因に大きな差のある集団1から、すべての条件が同一の集団4へと移るに従って、同一処方では絶対不可から、だんだん同一処方でもよいようになり、集団4では完全に最適な唯一つの処方になるであろう。

医師は、それぞれの患者に対し最善の処置をほどこさねばならないが、その患者と全く同一条件の患者は、それまでに絶対になかったはずであるから（同一の患者であっても、前回とは関連条件は違っている）、過去の経験がそのまま役に立つことはないであろう。個人差があり、色いろ条件が異なっているが、同じような病気の患者についての過去の治療経験を、大筋でまとめて体系化し、それにもとづいて、新しい目前の患者の治療に立向うことになる。この際どうしても個体差・条件差を問題としないわけにはいかない。

近代統計学は、英国におけるコムギの品種改良実験をその発端としているが、実際のコムギの収穫量は、品種によるばかりでなく、年度による差が大きく、その外土壌・水はけ・日射量・風通し・耕作管理の良否などに影響され、簡単には決まらない。統計学は、このような条件差

・個体差のある生物の集団に対し、その集団における現象の法則性または集団の構造（A品種の収穫量の方が、B品種のそれより平均して多いとか、病虫害に対する耐性の優劣など）を発見しようとするものである。しかも避けられない年度差（天候差）の下で、それぞれ差のある試験圃場からの無作為標本（ランダム・サンプル）を手がかりにして、集団全体としての法則性を発見し、われわれ人間の行動の指針を得ることを目指している。従って、個人差・個体差を避けて通ることのできない治療医学、いやむしろ個体差が中心的課題である治療医学研究と甚だ類似した研究条件下にあり、最終目標が人間行動の決定にある点でも酷似していると云えよう。

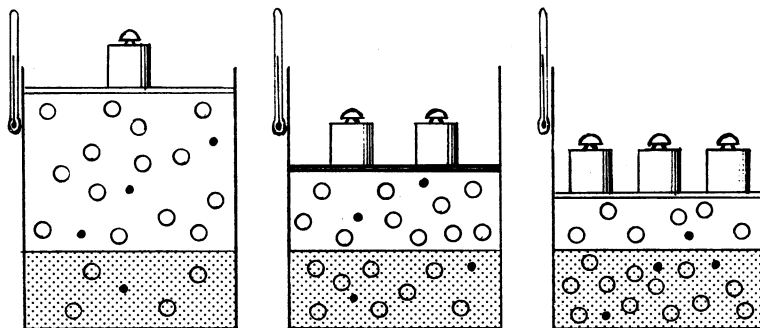
上記の集団1では、流行性感冒に特有の治療方針に従い、当然個人差による「さじ加減」を、統計的推測の下で行うべきである。また、集団4では、個人差も環境差も全然ないから、物理化学の場合と同様、最適な処方が存在するはずであり、その意味で統計学の必要は殆んどない。

## 2. 同じく「実験研究」と云っても、物理化学の場合と治療医学の場合とは大違い

### — 実験材料と環境条件の制御・管理について —

均質な材料で、完全に条件制御ができる場合の現象の法則性と、個体差・条件差を避けて通れない場合の法則性とは、その探究の方法・結果の要約に大きな違いがあるが、この辺の事情を簡単な例を通して考えてみよう。

物理化学における実験としては、「溶媒にとけこむ気体の質量と圧力との関係」についてのヘンリーの法則をめぐる実験を、医学生物学における実験としては、「運動能力（垂直とびの高さ）と年齢との関係」を調べる実験をとりあげ、それぞれの実験材料の均質・異同性、環境条件（圧力・温度、年齢・実施条件など）の制御・管理、および実験結果の法則性に関する確実性などを比較してみよう。

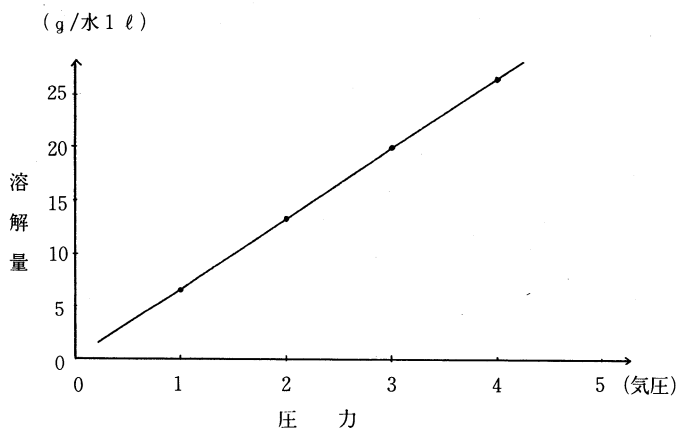


第1図 圧力と溶解量（説明図）

1) 溶媒にとけこむ気体の質量と圧力との関係を研究したいのであるから、溶媒にとけこむ気体の質量と圧力との間に、ある種の関係が成り立つことは、すでに知っているわけで、その関係を精密に確立したいのである。

ところで、圧力が一定のときは、温度が高くなると液体内の分子の運動がはげしくなり、液体分子の引力にうちかって液体の外に飛び出す気体分子の数が増えるので、気体の溶解度は小さくなることが知られている。また、温度・圧力が一定ならば、溶媒の量が多いほど気体は多量に溶解するから、実験中温度を一定に保ち、一定量の溶媒を用いなければならない。

精密実験科学では、実験中温度を一定に保つことができ、また正確に一定量の溶媒を実験に供することができること、及び実験に用いる気体として酸素・水素・窒素の混合気体を用いるとすれば、常に組成一定の混合気体を用いることができることを、実験の大前提としている。実験中温度を一定に保つことができなとか、溶媒の水1ℓというのを分量でしか測ることができないとか、混合気体の組成を一定比率に管理できないといった場合に、「溶媒にとけこむ気体の質量と圧力との関係はどうなるか」といったことは考慮の外である。もしそのような考慮の必要があるとするれば、もはや精密科学ではなくなってしまうのである。



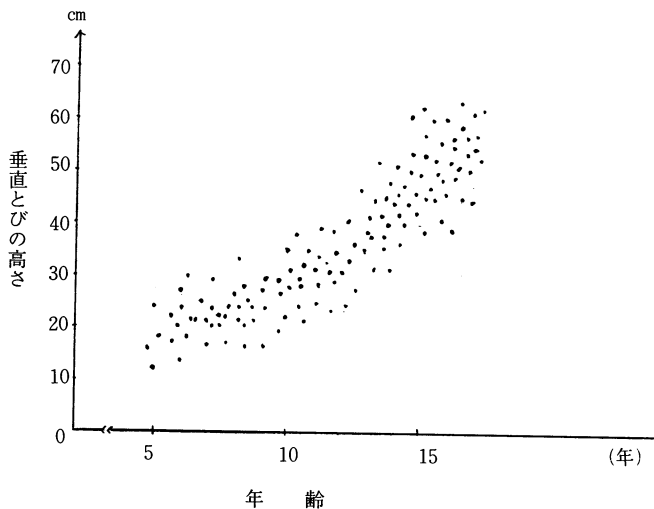
第2図 ヘンリーの法則

上の実験結果については、ヘンリーの法則が成り立つことが、ひろく知られている。すなわち、

温度が一定のとき、一定量の溶媒にとけこむ気体の質量は圧力に比例する。

2) 一方、運動能力（垂直とびの高さ）と年齢との関係を調べる実験では、「5歳から17歳の男子の間では、垂直とびの高さは年齢とともに、平均的には増大傾向にあるだろう」という予想の下に、できるだけ正確に両者の関係を調べようというのである。ところで、年齢が同じであっても、身長が異なる上に肥満児とやせ型の者とは、垂直とびの高さはかなり違うであろうし、跳躍実験を行う時間とか、実施者の寛厳などによっても影響を受けるであろう。されば

といて、混合気体の組成を一定にするように、5～17歳の男子の体型をそろえることは非現実的であり、温度を実験中一定に保ったように、跳躍実験の時間とか実施者の寛厳を同一にすることはむつかしい。できるだけそのような状態に近いように努力するのが精一杯である。さらにまた、あまりに厳密な条件下で実験が仮りに実施できたとしても、その結果は現実とはかなり離れたものとなり、実際場面での適用を考えると、無駄な骨折りのそしりをまぬがれ得ないであろう。個人差がある上に、垂直とびの高さに関係のある要因を十分管理して各人を同一条件下におくことは不可能に近く、現実的でもない。従って得られた実験結果は、気体の溶解量と圧力とのように比例関係にはなく、(年齢、垂直とびの高さ)のデータは、第3図に示すように、右上りの帯状域に散らばることになる。



第3図 年齢と垂直とびの高さ (男子)

従って、ヘンリーの法則のような確定的法則は得られない。新しく無作為に選んだ15歳の男子の垂直とびの高さは、第3図における15歳の人達の平均と標準偏差とから、ほとんど確実にこれこれの範囲内にあるという情報が得られるだけである。このような確率的な法則しか得られないのは、個体差をもつ集団を研究対象にし、実験結果に影響を与える条件の制御・管理が十分できない場合の宿命とでも云うべきであろう。実験の結果に関する知見を不確実にしている要因として、上記の実験材料の個体差及び関連要因の制御・管理が不十分であることの外に、現実の実験対象がもとの集団から選ばれた小数の標本であることが重要である。個体差があるのだから、標本を再抽出すれば、またまた実験結果が異なってくるのは必定である。精密科学における実験の場合には、実験材料は均質であり条件制御も完全であるから、標本による違いは全く考えられない。あるとすれば、計器の精度によるもの、実験者の測定誤差などにもとづく小さな変動だけである。

医学でヒト・動物の集団を扱う実験では、研究の対象である（不揃いな）個体の全部（母集団）にあてはまる知見・情報を、そこからの無作為標本（くじ引きで選んだ標本）に対する実験結果を手がかりにして、その変動を究明することが根幹であり、数理統計学における母集団と標本の理論が、その研究の基盤になっている。

念のため、物理化学における実験と、医学特に治療医学における実験の異同を表にまとめておこう。

表1 同じく実験研究というけれど

	実験結果に変動を起こさせる要因		実験から得られる知見	実験の主目標
	実験対象・材料	環境要因・実験要因以外の要因の制御・管理		
物理化学	常に均質 (空気の組成一定)	徹底的(温度一定、 溶媒一定量)	決定的(ヘンリーの法則)	現象の理論的 解明
医学 -治療医学-	個体差あり <標本のぬき方で 増幅>	できるだけ (実施条件に違い)	近似的	人間行動の指針

( ) 内は上記実験における具体例

### 3. 母集団と標本

精密実験科学においては、天然現象や、実験室内で制御された条件の下で生起する自然現象を、観察したり数量的に測定したりして、その結果から当該現象を支配する法則を究明しようとしている。そしてこれまでの輝かしい研究成果は、その方法論に殆んど疑念をさし挟む余地を与えないかのように思われる。実験材料は均質であるから、実際に実験に供された実験材料（標本）と、研究対象と考えられているもの（母集団）とは、全く同質であり区別できない。従って、それらに関して観察される現象も、実験条件の制御・管理が完べきであれば、全く同一であって問題は残らないはずである。（もちろん実験材料の均質性にしても、実験条件制御の完べき性にしても、その時点における技術レベルに関連することであり、また思考様式そのものにも関係があって、実験は目指す現象の一つの近似であり、一般に対しやはり特殊と云わざるを得ない。この意味で、特殊のあるいは具体的な事例から、一般的な法則・命題を導き出そうとする帰納法の制約を脱却することはできないわけで、このことは科学の歴史を見れば明らかである。実験材料の均質化・純粋化、ならびに実験条件の完全制御は、この特殊性からの脱出を意図していると云えよう。）

これに対し、実験材料の不揃いな集団、個体差のある個体の集団を母集団とする医学特に治療医学の実験では、まず実験の出発に当り、実際に実験に供される個体の集団（標本）の選択

が大問題である。

もともと母集団の個体は個体差があるのであるから、そこにおける現象・事象に関して確定的な法則が得られるはずはない。確率的な、ある変動幅をもった法則しか得られないことは明らかであろう。ところで、標本が母集団の完全な縮図になっておれば、そこでの実験結果から母集団におけると同じ確率的な法則が期待できるであろう。もしそうなら、一応目的は達成せられたと云うべきで、めでたしめでたしである。

しかし、母集団の完全な縮図になっているような標本は、どうしたら選べるであろうか。「それは簡単！ 無作為標本（ランダム・サンプル）を選べばよい」というのは、話がとび過ぎている。いま研究しようとしている現象・事象に関して母集団の相（実相）が完全に分っているのなら、その縮図としての標本抽出は可能であろう。しかし、母集団の相が分っているのなら、今更研究の必要はない。従って実験そのものが不必要である。分っていないからこそ、標本を抜き出して、実験しようとしているのである。これでは出発点において矛盾である。

矛盾しているのなら、その実験は止めてしまえというのなら、話はそれまでであるが、最善が駄目なら次善の道をとというのが人間行動の規範である。患者を前にした医者が、最善の方法が分からなくても、いや時としては打つ手がないことが分っていても、何らかの処置をとることを要請されるのと同断である。

母集団の、目指す現象に関する関連情報を使い尽した場合、最終的には無作為標本に頼らざるを得ない。その無作為標本が、母集団のよい縮図になっているかどうかは分らないが、母集団からはなれ具合は確率論を用いて計算可能であるから、定量的な結論を導くことができる。と云うに過ぎないのである。無作為標本は母集団の縮図という意味において、良い標本とは限らないが、自分の予想とひどく違うときは、何べんでも標本を抜き直して確かめることができるから、公正な標本とでも云うべきであろう。

1985年春ニューヨークに現われた日本大相撲の一行は、日本人という母集団からの1つの可能な無作為標本（小錦を除いて考える）であるが、母集団のよい縮図にはなっていない。この標本が無作為標本として抽出される確率は0ではなく、明らかに正の数である。ただ彼等の集団と同等以上に大きい方にずれた標本の得られる確率が非常に小さいことは云うまでもない。

#### 4. 実験母集団

統計理論における母集団対標本概念と、医学統計学における実験母集団対標本概念との間には、大きなへだたりがあり、医学実験における統計的推測をあいまいにしている。一般に統計学では、母集団の定義を明確にし、その中から標本を無作為に抽出し、測定データを求めてから統計的推測にかけるのである。すなわち、母集団はすでに存在しているのであり不変不動のものと考えられている。しかもその中から無作為に標本を抽出できることを前提としている。ところが、実験母集団と標本の場合には、これと全く事情を異にしている。一定の実験計画に

もとづく標本データだけがあって、その母集団に該当するものは実在しない。強いて云えば、全く同一の実験計画に従って、何回も何回も実験を繰り返したと仮定するとき、それら何回も繰り返された実験データの全体が、実験母集団を形成するはずである。このように考えるとき、はじめて母集団と標本との関係が明確になり、推定や検定の論理を展開することが可能になる。しかし、どれほど明確に定めた実験要項の下で実験を繰り返したとしても、現実で得られた標本データは、それに続いて考えられる標本データとは、質的にみて一般には異なるであろう。実験を繰り返すごとに、実験者の技量は向上し(手術を伴う動物実験の場合には特に顕著)、実験に関連する情報も増加するから、管理すべき要因の取扱いもおのずから変化するであろう。従って、最初の標本データの実験母集団と、それに続く標本データの実験母集団とは変化していると考えべきであろう。同一の実験の中においても、最初に手術した動物と、終わりごろに手術した動物とでは、手技の向上に伴う反応の偏りを考慮することが、厳密に云えば必要であろう。このような意味で最初の実験において予期するような結果が得られなくとも、失望落胆するのはいささか早計に過ぎると云えよう。実験計画を改善しながら実験を繰り返すうちに、実験者の技量が向上し、環境要因の制御管理も手ぎわよくなり、実験施設も充実してくれば、理想的な実験母集団に近いものからの標本に近いものが得られるようになり、成功が得られるかも知れないからである。

##### 5. 統計学を必要とする研究・必要としない研究

医療の最終段階では、個人差の中で確率的法則を見出し、それを基準にさじ加減が是非とも必要であるから、これまで述べてきたように統計学は必須である。どのような薬を、どの位飲ませ、どのように養生させればよいかの大よその目安、すなわちその病気に対する大筋の治療方針—確率的法則—があって初めて(いく通りもの治療法があるときには、その中からいずれを選ぶかをも含めて)、個体差に必ずる微調整、いわゆる「さじ加減」が意味をもつのである。大筋の治療方針のないところには、個人差に必ずるさじ加減も何もあったものではなく、場当りの対症療法か、当るも八卦当らぬも八卦的な気やすめ療法しかないであろう。1人ひとり顔が違うように、治療も患者ごとに違うのが当然で、患者をのっぺらぼうにして得られる一般法則など、治療の世界には関係ないなどと云うのは、科学としての医学を否定するものに外ならない。

近時医学研究においても、ヒトまたは動物そのものを対象としない研究、細胞レベル・分子レベル・遺伝子レベルの研究が盛んになってきている。いわゆるバイオテクノロジーを用いる研究であるが、そこでは生命現象の原理的解明を志向するものが多く、方法論的にも精密実験科学そのものと云うべきであろう。試験管の中で培養された同一種の細胞に、ある種の完全に制御された操作を加えて、細胞の変化を調べる実験などにおいては、実験材料の均質性・環境条件の完全制御ともども、精密実験科学と全く同じである。生きものについての実験は、それが独創的であればあるほど、時に研究者の管理外の局面を不可避免的に含んでいる可能性が、物理



化学の実験より一層高く、他人の手による追試をより必要とするという点に、差があると云えば云えるくらいであろう。このような分子レベル・遺伝子レベルの研究では、統計学の出番は少ない。メンデルの遺伝法則では、原理的にはDNAまでさかのぼらなければならないが、遺伝子のランダムな振舞いにもとづく集団現象が主題であり、統計学の好個の対象であった。一方、物理化学的実験研究においても、原子分子レベルになるとまた、原理的には少し異なるけれども、確率論を必要とし、統計学的な考察を必要とすることは申すまでもないことである。

最後に史上最大と云われるポリオ・ワクチンに関する実験研究について触れておこう。この医学実験は、1954年アメリカ合衆国で実施された。ポリオによる麻痺や死亡を予防するのに、ソーク・ワクチンが有効かどうかを評価するために行われたものである。100万人以上の幼児がこの実験に参加し、直接経費だけでも500万ドル以上と云うから、その規模の大きさには驚かざるを得ない。ポリオ根絶のために、このような大がかりな調査研究を実施するに当っては、アメリカ合衆国第32代大統領フランクリン・D・ルーズベルト（彼は若いころポリオに襲われた）の偉大さもさることながら、アメリカ合衆国が持つ底知れぬパワーも見逃すことはできない。このことは現在、ガン制圧に向けての国としての取組み方を見ても明らかであろう。

さて、ポリオの原因はウイルスであることがわかり、「ポリオ・ウイルスを何世代にもわたって培養することにより、有害種から無害種を開発育成する研究」、「ポリオ・ウイルスの感染性を除去し、その抗原効果をできる限り破壊しないよう十分ホルムアルデヒドで処理し、不活性化ウイルスを開発する研究」等がなされた。これらの原理的な研究段階では、前述したように統計学は殆んど必要ない。ところで、1950年代の初めに小児麻痺のための国家基金によって招集された諮問委員会は、「ピッツバーグ大学のジョナス・ソークの開発した不活性化ウイルス・ワクチンが、それを試験接種された子供達にとり安全で、かつ高水準の抗体作用を惹起しうる」ことを認めた。これにより、ソーク・ワクチンは一般向け使用ワクチンの有力候補にはなったのであるが、「実際に感染の危険にさらされている人々を、ポリオから防ぐことを実証すること」すなわち有効性の確認が最終的に残されたのである。医療の最終段階では必ず、統計学を必須とする個体差前提の評価の問題があり、これが過去最大の公衆衛生実験といわれる1954年の実験研究となったのである。

実験は様ざまの紆余曲折を経て、二重盲検法により実施された。二重盲検法が治療や薬の有効性研究に最適であることは、現時点では疑問の余地はなく、わが国でも各種の薬剤の有効性再確認試験に多用されていることは周知の事実である。この二重盲検法は、(1)病気は生体のもつフィードバック機能により自然に動くこと、(2)患者には個体差があること、(3)偽薬効果（心理効果）があること、(4)医師の効果判定も投与した薬のイメージによって影響を受けることに対し、無作為化を応用して患者側にも医療側にも、どちらにも知られないように公正な対照群をつくり、適正な有効性の数量的評価を可能にした統計学・医学共同研究の一大成果である。

なお、100万人を越える大規模な実験とならざるを得なかったのは、ポリオの気まぐれな低率の臨床出現率のためで、要請された分析項目に有効な情報を提供するのには、それだけの標

本数が必要だったのである。

この外、医学研究における統計学の役割については、

- 麻酔薬の安全性評価 (標本抽出法・割合の標準化)
- 新しい良薬を求めるための絶えることのない研究 (実験計画法)
- 投薬量基準の設定 (逐次実験計画法)
- 新治療法は外科や麻酔において、どの位の頻度で成功するか
- 喫煙と健康
- 伝染病の発生と衰退のモデル構築により、都市の大きさを設計する
- 人と類人猿の骨の判別
- ヒトのリンパ球水準の測定において、2法のうちどちらの測定法がよいか
- 睾丸癌の治療における用量水準の設定
- 臨床検査室における定量法の評価
- 微小集落のラジオ・オートグラフィーにおける推定問題
- 乳児の原始歩行に関する研究
- 心臓移植の研究 (生存日数予測と予測変数の探索)

等、文献 [2] [3] [4] に親切な解説がなされている。

## 文 献

- [1] 仮谷太一：医学部進学課程における統計学について，川崎医学会誌一般教養篇，10：1-9 (1984)
- [2] Lehmann, E. L. (ed.): Statistics : A Guide to the Unknown, 2nd Ed., Holden-Day (1978)/  
安藤洋美監訳:統計学講話—未知なる事柄への道標，現代数学社 (1984)
- [3] Rupert, G.M. JR. and others : Biostatistics Casebook, John Wiley & Sons (1980) /医学  
統計研究会訳:医学統計実践事例集，マール社 (1982)
- [4] Brown, B. W. & Hollander, M. : Statistics — A Biomedical Introduction, John Wiley &  
Sons (1977) /医学統計研究会訳:医学統計解析入門，マール社 (1984)