

NICU

第2巻 第3号 1989年3月10日発行

大阪 **メヂカ出版** 東京

新生児

ナース

トレーニング

統計学とナーストレーニング

— 統計学の考え方と
応用について —

高木廣文 * Hirofumi Takagi

I はじめに

「統計学は難しい」「統計学なんて分からなくても臨床には関係ない」というのが、統計学を授業や講習会などで学んだ人たちが抱く感想であろう。このような意見は、ある意味ではもっともな気がしないでもない。

多くの統計学の教科書などを開くと、そこには際限なく数式が並べたてられている。かなり平易に書かれている場合でも、最少限の数式は不可欠である。したがって、数式が苦手の人には耐えられないことだろう。

この点から「統計学は難しい」という意見に異論はない。ただし、難しいのだから、単に統計学の講義を受けた、研修会に参加した、著名な統計学の成書を読んだといった程度で、統計学の各手法を理解できると考えるべきではないことも事実

である。

また、統計学の専門家に意見を聞けば、自分の抱えている問題が一挙に解決するという訳でもない。他の学問と同様に、地道に学習を積み重ねていく以外には、統計学を自分のものにする方法はない。

「統計学を勉強する時間はない。それに臨床に関係ないし」という意見はどうであろうか。確かに、日常の定型的な業務のみを行っていればよいという考えを持っている人には不必要かもしれない。ただし、科学的な看護研究を行いたい、最新の論文を読んで、正確な知識を身につけたいと考えた場合、統計学の基礎がある程度なければ、正しい判断はできないといっても過言ではないだろう。

ここでは、統計学の基礎的な考え方を中心に説明し、各手法の詳細については省略する。わずかなではあるが、実際の応用上の問題点などについても言及したい。

*聖路加看護大学助教授
(〒104 東京都中央区明石町10-1)

II 統計学とは何をするものなのか

統計学がその対象とするものは、個人ではなく「集団」である。特定の集団について、調査などにより共通の項目についてデータを収集し、その集団がもつ各項目や特性などの傾向について、検討するための方法を提供するものであるということもできるだろう。

統計学を使う目的は、大きく2分類して考えることができる。1つは、特定の事象の「記述」であり、もう1つは調査などから得たデータを基にして、調査結果や研究結果を「一般化」できるかを検討することである。前者は「記述統計学」と呼ばれ、後者は「推測統計学」と呼ばれることもある。

記述統計学の方法は、特別難しいものではない。例えば、50人の女兒の出生時の在胎週数と体重のデータがあるものとしよう。平均体重が3100g、平均週数が40週などと、平均値を計算することがよく行われる。また、体重が2500g未満が6%であったなどと、ある特性をもつものの割合を示すこともある。

これらのことは、データをまとめてある特性を示すために行うものである。すなわち、記述統計学とは、個人についてのデータが多数ある場合、それを分かりやすくするためにデータを図示したり、データを要約するための数値を求めたりする方法であるといえる。

データを図示する手法としては、「円グラフ」、「帯グラフ」、「棒グラフ（もしくはヒストグラム）」、「折れ線グラフ」などがあり、特殊な方法として「幹葉表示」なども最近はよく用いられてい

る。また、2つの変数間の関係を図示するためには「相関図（散布図）」が使用される。

データを要約するための数値としてよく使われるものに、平均値や中央値などの「代表値」がある。これらは体重や血圧などのようにデータが数値により量として把握することができる「量的データ」の場合に使用される。また、データのばらつき具合をみるためには、「標準偏差」や「分散」が用いられ、2つの変数間の関係の程度を示すためには、「相関係数」と呼ばれる指標も用いられる。

血液型のような場合には、データは数値ではなく、ある特性の有無によって、A型、B型などと表現されている。そのようなデータは「質的データ」と呼ばれ、“A型が40%”というように「比率（割合）」を用いてデータの要約を行うのが普通である。また、2つの質的な変数間の関係を見るためには「クロス表」と呼ばれる表を作成する。

上記のような方法を用いることにより、自分のもつデータの特性の把握が容易になり、また客観的に各特性を説明することが可能となる。

ここで述べたような記述統計学で通常用いられる手法や指標がよく分からない、もしくはすっかり忘れてしまったという読者でも、適当な統計学の成書を読めば、これらの内容を理解することは比較的容易なものと思う。

問題は、「推測統計学」と呼ばれる範ちゅうに属する一連の手法である。記述統計学では、データをまとめて、分かりやすく「記述」ということにその目的があった。しかし、推測統計学ではデータに基づいて「推定」と「検定」を行うことにその目的がある。それでは、推定や検定で何が

分かるのだろうか。

III 母集団と統計学的推定

推測統計学の方法を理解するためには、まずその考え方の基礎となる「母集団」と「標本」の関係を、明確に把握しておく必要がある。

いま、新生児科医100人について、これまで治療した新生児に関するいくつかの項目について質問紙による調査を行ったとしよう。その結果、これまでに治療経験のある児の最低出生体重の平均値が450gであったとしよう。それでは、一般に新生児科医は、出生体重450gの児の治療経験をもっているといえるのであろうか。すでに述べたように、平均値が450gというのは、手元にあるデータの1つの要約である。しかし、これを一般化することとは、調査していない他の医師の状況についても、その結果が成り立つということを主張することである。この関係をもう少し、統計学の用語を用いて説明しよう。

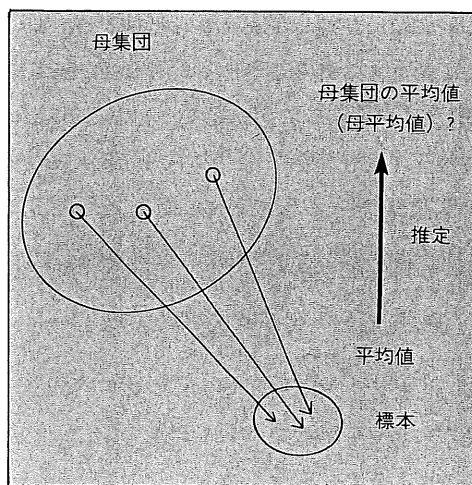
実際の調査対象は「標本」と呼ばれ、標本に対してその背後にある、より大きな集団は「母集団」と呼ばれる。上記の例では、新生児科医という集団が「母集団」であり、実際の調査対象となった100人が「標本」ということになる。データは標本から得られるものであり、ある特性の平均値は標本についてのものとなる。分析結果を一般化することとは、母集団について言及することに他ならない。少し考えれば分かるように、母集団から標本を選ぶ方法は無数にあり、たまたま選んだ標本の平均値などが、母集団がもつ平均値（「母平均値」と呼ぶ）に一致するとは限らない。標本のデータから母集団の特性値（「母数」と呼ばれる）

について言及するための手続きが「推定」と呼ばれるものである。この関係を図1に示した。

実際に推定を行う場合、統計学では標本が母集団を「代表」していることを必要とする。この要求を満たすためには、標本の取り方が問題になる。データが集めやすいので知り合いに頼んだ、ボランティアに頼んだなどというのは問題外である。「無作為抽出法」と呼ばれる標本選択の方法のみに、この要求を満たすものである。無作為抽出法の基本的な考え方は、母集団の構成員について、各人の選ばれる可能性が全て等しくなるような方法で標本を選ぶという点にある。具体的には、母集団の全員に一連番号をつけ、コンピューターなどで人間の思惑が入らないように、「乱数」と呼ばれる全く作為のない数値を求めさせ、その数値に一致したものを標本とするという方法がよく用いられている。

標本が正しく選ばれた場合、推測統計学の手法により推定を行うことができる。ここで、推定の

図1 推定の考え方



方法は「点推定」と「区間推定」にさらに分類される。

点推定とは、母平均値などを1つの数値を用いて推定するものであり、例えば、「新生児科医が治療経験した児の平均最低体重は450gである」というような場合である。標本が無作為に選ばれている場合、標本から得た平均値と比率は、母集団の母平均値と母比率の点推定値となることが知られている。

ところで、点推定は1つの数値しか用いないため、必ずしも母集団の母数と一致するものではない。そこで、ある統計学的な確からしさで区間を設定し、その範囲内に母平均値などの母数が存在することを示す方法も考えられる。そのような方法が区間推定と呼ばれるものである。そして、母数の存在すると推定される区間は「信頼区間」と呼ばれる。さらに、その信頼性の程度をつけて「母平均値の95%信頼区間」などと呼ばれる。この場合、100回の同一の調査を行い、同一の計算方法を用いれば、95回はこの信頼区間の中に母平均値が入るということを「95%」により表現している。逆にいえば、5%は間違える可能性もあるということであり、これを小さくしたければ「99%信頼区間」などを求めることになる。先の例では、「新生児科医が治療経験した児の平均最低体重の95%信頼区間は430~470gである」といった表現になる。

IV 検定とは何をするものなのか

検定も推定と同様に、標本についてではなく、母集団の各種の特性について言及するための方法である。しかし、推定と異なり「帰無仮説」と呼

ばれる母集団についての仮説を設定し、それが成り立つかどうかをデータを用いて調べるという手順をとる。ただし、検定のための仮説は、研究上の作業仮説とは異なるものである。

例として、妊婦の喫煙に関して、「喫煙者は非喫煙者に比べて、未熟児出産率が高率である」という作業仮説を立てて調査を行ったとしよう。

このような場合、妊婦を喫煙群と非喫煙群に分け、各グループでの未熟児の割合を調べることになる。調査データでは確かに、喫煙群で未熟児の割合が高率であったとしよう。しかし、2つしかグループがなければ、完全に2つの割合が一致する場合よりも、偶然の影響により、どちらかが高率となる可能性はかなりあるだろう。

統計学的な仮説検定法では、「喫煙者と非喫煙者の未熟児出産の母比率（割合）は等しい（差はない）」という仮説を立てる。この仮説が正しいとした場合に、調査データで認められるような2グループ間の比率の差が偶然生じる可能性（確率）を計算する。もしもその確率が大きければ、2つのグループでの差は偶然によるものとも言えることになり、それ以上のことは何も言えなくなる。逆に、確率が小さければ、そのように稀なことがたまたま自分の研究で起こったと考えるのは無理があり、そもそもの前提となる仮説が誤っていたと考えねばならない。したがって仮説は否定され（「棄却する」という）、2つのグループでの母比率には差があると言えることになる。上記の検定の手順を図2に示しておいた。

このように統計学的仮説検定における仮説は、棄却されるために存在して、そのために帰無仮説と呼ばれるのである。かりに棄却されなければ、

検定は無意味となる。なぜならば、仮説が棄却できないということは、仮説を積極的に支持する根拠を与えた訳ではなく、そのような状況もありえるということを単に示すに過ぎないからである。

ところで、検定のための目安となる確率の大きさは「有意水準」と呼ばれ、慣習的に5%か1%がよく用いられているが絶対的なものではなく、状況によっては10%程度の甘い基準でよいかもしれないし、重要な意思決定に関わる場合には0.1%を採用するなど、状況によって適宜判断を自分自身で下す必要がある。

V 実際の応用はどうするのか

これまでの説明で分かるように、推測統計学の手法を用いようとした場合、分析のためのデータは無作為抽出法により選ばれた標本から得たものでなくてはならない。臨床で担当した患者から1例1例集めたものが、ある程度の数になったから集計してみようというのでは全く駄目なのである。せいぜい記述統計学の手法を用いて、データの要

約や図・表などを作成して、実態を報告できるだけである。

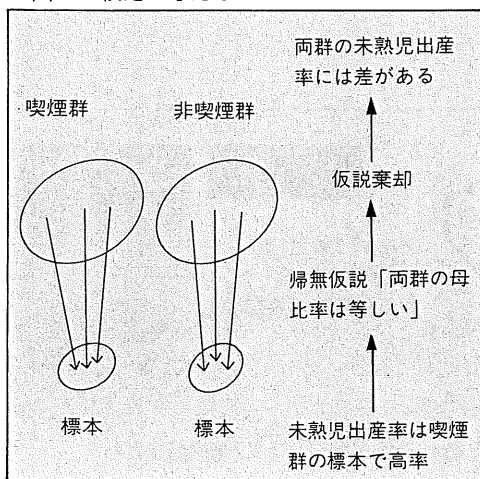
それでは、臨床で得たデータが何の役にも立たないかというと、必ずしもそうではない。標本が母集団を代表するものであるとある程度言えれば、推測統計学の方法を用いても、それほど大きな誤りを犯すことはないだろう。そのためには、性別・年齢別などで患者をみた場合、それが典型例であるという保証が必要だろう。したがって、当該疾患の専門家と共に、それに該当する患者を選択すればよいだろう。このような標本抽出法は「任意抽出法」と呼ばれ、必ずしも良いものではないが、臨床場面ではそれ以外にあまり良い方法は思い浮かばない。

分析可能なデータを手に入れたとしても、それではどうやって実際の分析を行えばよいのであろうか。数十例程度ならば手計算により分析も可能であるが、項目数が多かったり、例数が多い場合には、少なくともパソコンの助けが必要である。しかし、パソコンがあっても良い統計学のソフトがなければ何もできない。ここでは、その1つとしてわれわれが開発した「統計学パッケージ HALBAU (現代数学社から販売)」を紹介しておく。数百例、数百項目の集計・分析ならば、HALBAUを用いて比較的簡単に行うことができるだろう。ただし、ある程度のパソコンの知識は必要であるし、統計学の知識がなければ結果の解釈が不可能な場合も少なくない。どの分野でも同様であるが、ある程度の「慣れ」と独習は必要である。

VI おわりに

今回は、統計学の基本的な考え方を中心に説明

図2 検定の考え方



した。そのため、具体的な分析方法の解説がなく、物足りない読者も多いものと思われる。しかし、統計学の講義を行っていて常に感じる事は、多くの学生・受講者が、なかなか推定や検定の考え方を理解できないということなのである。

検定などの数式は、どの書物にも書かれているし、かなりの範囲まで独学が可能である。しかし、その基本となる思考方法を理解する事は、かなり難しいものと考えられる。実際の応用上の問題な

ど、ほとんど触れることができなかったが、参考文献を示しておいたので、それを参考にして欲しい。人間のもつ多種多様な側面を扱う看護研究の分野で、積極的に統計学の手法が応用されることを期待している。

参考文献

- 1) 高木廣文：ナースのための統計学、医学書院、1984.
- 2) 高木廣文、柳井晴夫：統合型統計パッケージHAL-BAUの開発、聖路加看護大学紀要、14；27-34、1988.
- 3) 林知己夫監修：調査とサンプリング、同文書院、1985.

専門医からの健康アドバイス

ワンポイントクリニック

毎日新聞長寿連載—医療・健康コラム「待合室」
読者の声に応えて第2集刊行!!

好評

- A5判 ●216頁
- 定価1,200円
- 毎日新聞学芸部編



メヂカ出版

〒564 大阪府吹田市広芝町18-24 ☎(06)385-6911 FAX (06)385-6721