



## 『有意差』の意味 —統計的検定を用いた医療系論文を理解するために—

足利 学

### I. はじめに

我々は2つの異なる事象に関連性があるかのごとく誤って認識してしまうことがあり、実際の調査の結果と異なる場合がある。医療・保健・福祉・教育などの実践を伴う領域では、法則性 (ideographic) に基づいた知識による裏付けがしばしば求められ、得られた知識を活用した介入方法を確立し、最終的には現場へ還元することが重要である。法則性に基づいた知識を得るためには、客観性を満たす(皆が納得する)データを提示する統計的手法を用いた実証的な研究が有用である。他方、人間に関する現象を対象とする学問では、対象者の個別性 (nomothetic) を無視するわけにはいかないことはいうまでもない。専門家による直感や論理的思考に基づいた示唆に富んだ研究も重要である。

近年、医療系の論文、特に看護、理学療法、作業療法の分野では統計的な処理を行った実証的な形式をとる論文が多くみられる。特に因子分析や共分散分析などの多変量解析(3つ以上の変数を同時に取り扱う統計的な解析の総称)は、多くの論文で適用されており、統計学的な知識なしでは、論文の内容を適切に理解することは困難である。

そこで本論では、統計学の専門知識なしで科学論文を適切に理解することを目的に、できるだけ数式を使用せずに、統計学的手法の初歩である統計的検定 (statistical test) について述

べる。したがって、確率分布などの統計学の理論については、本論の最後に比較的容易に書かれた文献を示したので、各入門書をご覧ください。

### II. 統計的検定

統計的検定とは、2つ以上の条件の間に、何らかの値の差があるか否かを確率的に判断するための分析であり、調査や実験の結果を一般化するための手続きである。また検定を行う意義は、研究者間で共通の判断基準を設けることによって、得られたデータを解釈する際に、主観性をできるだけ排除することである。すなわち、偶然に生じた可能性の高い結果に対して、研究者の過大な意味づけを防止し、合理的に判断することである。統計的に有意な差があるか否かを検定する方法は、分析の目的や得られたデータの特徴によって数多くあり、 $\chi^2$ 検定(カイ二乗検定)、t検定、F検定、メディアン検定、U検定、分散分析などの手法がしばしば用いられる。

### III. 危険率 (risk rate)、有意水準 (level of significance)

研究の目的によって異なるが、医療系の研究では、一般に95%以上の確率であれば仮説を十分に証明したと考えられている。一方、十分でない確率(仮説が間違っていたかもしれない)を危険率といい、危険率5%と表現する。統計的に有意な差がみられた時には、論文中の記述では、「危険率5%水準で有意差が認められた」

『危険率5%水準で(統計的に)有意である』、図表では『 $p < 0.05$ 』などと表記する。ここでいう「危険率5%」とは、100回に5回は偶然に起こるという意味であり、危険率を有意水準ともよぶ。

例えば、表1は糖尿病患者100名と非糖尿病患者100名の空腹時血糖値の平均値である。両群の平均値に統計的な差があるか否かを明らかにするために、検定(t検定)を行った結果、危険率5%水準で有意差がみられ、このデータからは、糖尿病患者は非糖尿病患者に比べて空腹時の血糖値が高いことが示唆されたと解釈する。この結果をグラフに示したものが、図1である。

表1. 糖尿病患者の空腹時血糖値

空腹時血糖値 (mg/dl)	
糖尿病	184
非糖尿病	88

\*  $p < 0.05$

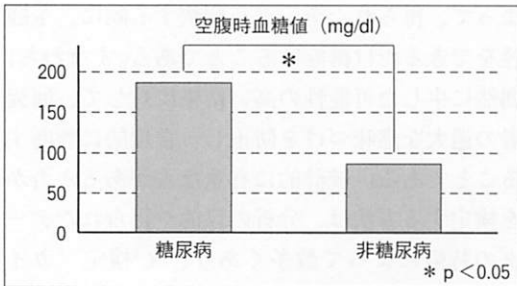


図1. 糖尿病患者の空腹時血糖値

研究によっては、仮説を証明する基準として99%、99.9%を採用する場合もあり、それぞれ『危険率1%水準で有意差が認められた』『危険率0.1%水準で有意差が認められた』と表現し、図表では『 $p < 0.01$ 』『 $p < 0.001$ 』と表記する。なお図表には、\*印(アスタリスク)の付いた数値がしばしばみられるが、この印は統計的検定の結果、有意であったことを表している。ちなみに、アスタリスクが1つ(\*)では危険率5%、2つ(\*\*)では危険率1%、3つ(\*\*\*)では危険率0.1%として慣例的に用い

られている。さらに危険率10%で有意な差を示す際には、†(ダガー)で表記し、『危険率10%水準で有意差が認められた』ではなく、『有意な傾向が認められた』と記述することが多い。危険率が低い(5%→1%→0.1%)検定ほど、信頼性の高い研究であり、研究者はできるだけ低い危険率を選択することが求められる。

また、調査の対象者(患者、被験者)の数が少ない場合には、統計的な分散(ばらつき)の影響のために、有意差がみられない場合がある。論文の中では、『有意な差は認められなかった』、『検定の結果は(統計的に)有意ではない』などと記述する。この場合には、図や表にはn. s. (no significance、有意でない)と略記することがある。

#### IV. 統計的検定の限界

統計的検定は得られたデータに判断を下す際の主観性をできるだけ排除するための一つの手法であるが、決して絶対的なものではなく、以下の3点を考慮しながら論文を読み進めていくことが必要であろう。

##### 1. データの数

統計的検定の結果はデータの数によって異なる場合があり、データ数が多ければ統計的には有意な差になりやすい。例えば、2つの事象の関連性について検定を行った場合、データ数が50ぐらいの少ない場合には有意差がないにもかかわらず、データ数が1,000ぐらいになると、有意な差になることがある。表2は、5%水準で有意な相関関係が認められるピアソンの相関係数(2変数の関係の強さを表す指標)とデータ数を示したものである。データ数が1,000になると、相関係数が0.062と低い値でも統計的には有意になるが、実質的な意味を保証するものではない。逆に、データ数が少ない場合には、実際の母集団においては差が認められる場合でも、有意差を検出することができない。したがって、有意な差が認められなかった場合でも、

積極的に差がないことを肯定することはできない。

表2. 5%で有意になるピアソンの相関係数

データ数	相関係数
5	0.878
10	0.632
50	0.279
100	0.195
500	0.088
1,000	0.062

### 2. 有意差の過大評価

得られたデータが同じで、同じ統計的手法(解析方法)を適用すれば、誰が行っても同じ結果が導かれるために、統計的な手法を用いた研究では、客観性が十分に保たれているような錯覚に陥りやすい。しかしながら、基本的な前提条件(例えば、無作為に抽出されたデータであることなど)が完全に満たされた研究のみばかりとはいえ、たとえ有意差が認められたとしても、確率論から導かれた結果であり、個別のデータ(症例)全てにあてはまることではない。すなわち、数学的に処理された論文であっても、有意差を過大評価してはならないのである。

### 3. 危険率5%と1%

一般にデータ数が同じであれば、危険率5%よりも1%の方が、差が大きいことを示している。しかしながら、データ数が多くなると、有意差がでやすいという事実を考えると、危険率5%よりも1%の方が、大きな差があると断定することはできない。

## V. おわりに

医療系の論文を理解するために、統計的検定の有意差について述べてきた。平易に説明することを心がけ、できるだけ統計学の専門用語や数式を使用しなかったために、逆に曖昧になったり説明不足になったりしているかもしれない。統計学を基礎から勉強し、科学論文を読むようになる方法もあるが、論文を読み進めながら、少しずつ統計学の知識を高めていく方法も一つの手法であろう。

### 【統計学の入門書】

- 岩淵千明編著. あなたもできるデータの処理と解析. 東京: 福村出版: 1997.
- 古谷野亘. 数学が苦手な人のための多変量解析ガイド. 東京: 川島書店: 1988.
- 古谷野亘, 長田久雄. 実証研究の手引き—調査と実験の進め方・まとめ方—. 東京: ワールドプランニング: 1992.
- 東京大学医学部保健社会学教室編. 保健・医療・看護調査ハンドブック. 東京: 東京大学出版会: 1992.
- 進研アカデミーグラデュエート大学部編. すっごく簡単! 0からの心理統計. 東京: オクムラ書店: 2005.
- 服部環, 海保博之. 心理データ解析. 東京: 福村出版: 1996.
- 吉田寿夫. 本当に分かりやすいすごく大切なことが書いてあるごく初歩の統計の本. 京都: 北大路書房: 2004.