

研究・論文のための統計学

八木茂典*

【統計学は避けたい】

高校時代、数学で赤点を取ってきた筆者は統計学が苦手です。できれば避けて過ごしたいです。

かつての水銀体温計は計測に5分程かかっていましたが、電子体温計は数10秒です。これは、実際に計測しているのではなく、温度上昇曲線から統計学的に予測した値を表示しています。計算はすべてコンピューターがやってくれます。われわれに要求されているのは、統計学を適切に使うことです。

本稿では、私のように統計学が超苦手な方のために、研究・論文でどのように使うかを述べます。

【平均値と標準偏差】

データを集計するとき、最初に頭に浮かぶのは「平均値」でしょう。国民生活基礎調査によると、一世帯あたりの貯蓄金額の平均値は1139万円です。これを聞いて「多!」と思った方もいるでしょう。非常に高額な貯蓄の方もいるので、グラフは恐竜のしっぽのように右へ低く長く伸びていきます。この場合「平均値」は、全体の代表値として適切でしょうか。平均値は、正規分布(この値を中心として左右均等にばらついている)でないときは有用ではありません。この場合、中央値を使用します。「中央値」は419万円です。

平均値や中央値は、点を表していますので、次は範囲を考えてみます。平均値から左右にどの程度ばらついているか、平均値との差の平均が「標準偏差」です。この範囲にはデータの68.3%が入っています。2倍には95.4%、3倍には99.7%が入っています。ある研究発表で「対象は平均年齢±標準偏差、78.0±15.0歳」とありました。63~93歳、2倍で58~108歳、3倍で43~123歳です。123歳!。対象に世界最高齢より上の人がいることとなります。標準偏差は、正規分布でないときは使用できません。せつかくの研究も、すべてのデータが疑わしいと判断されてしまいます。

正規分布のときは「平均値」「標準偏差」を使用します。正規分布でないのときは「中央値」「四分位範囲」「最小値~最大値」を用い、図表は箱ひげ図を使います。

【検定と推定】

検定は危険率で示し「0.05未満の場合、有意な差がある」といいます。分かりやすくいうと「まぐれで勝った可能性が5%未満の場合、勝ったと

判定する」となります。いくらデータに差があっても(100対0であっても)、統計学では「差があるとはいえない(勝ったとは判定しない)」んです。「なんだそれ?」ですよ。

サイコロの各面が出る可能性は1/6ですが、6回振っても各面が1回ずつ出るとは限りません。6回では不十分と分かっているからですね。統計学的有意差を述べるには、少なくとも20試合必要ということです。

研究で統計学を使用するつもりなら、始める前に必要なサンプルサイズを把握しておきましょう。対応のあるt検定は34、ウェルチ検定は64、マンホイットニー検定は67です。

推定は信頼区間で示します。サンプルの結果から予測して95%の確率で存在する範囲を示します。危険率のように「差がある/ない」という二者択一の答えではありません。電子体温計の例でいえば、37.0°Cと表示されているのは、95%信頼区間が36.8~37.2°Cということです。

【研究・論文のための統計学】

2019年衝撃的な発表がされました。米国統計学会と、最も権威ある科学雑誌「Nature」が「有意性検定を用いるのを禁止する」と述べたのです²⁾。近年は、統計学的有意差を出すことが目的となっているような論文が大量に公開されています。事態は深刻で、ついに科学の発展に不利益なので使用禁止に至ったということです。

統計学では20試合しないと勝敗がつかえません。しかし100対0なら勝敗は明らかだと考えるのは、臨床的な感覚といえるでしょう。統計学的有意差よりも臨床的有意性の方が大切です。研究・論文に統計学は使わなければいけないわけではありません。サンプルサイズが小さければ、統計学を使うまでもなく、10例中4例だったと記述すればいいのです。

言いたいのは、統計学が研究・論文のハードルになってはならないということです。

【文献】

- 1) 清水信博：もう悩まない!論文が書ける統計。オーエムエス出版、埼玉、2010。
- 2) Amrhein V, et al.: Scientists rise up against statistical significance. Nature. 567(7748): 305-307, 2019.

*中央大学理工学部
Shigenori YAGI Ph.D, PT, JSPO-AT
Chuo University, Faculty of Science and Engineering