

統計的仮説検定において学習者の理解を難しくする 幾つかの要因の検討

綾 皓二郎

Email: aya@isenshu-u.ac.jp

石巻専修大学 理工学部基礎理学科

©Key Words 仮説検定, 帰無仮説, 対立仮説, 統計的エラーとリスク

1. はじめに

統計学が昨今ブームとなっている。専攻分野によらず、統計教育の重要性が共通認識となることは喜ばしいことである。しかし、学生にとって統計学を学ぶことは必ずしも容易ではない。統計教育の中で学生がどこで躓くかがわかってきた。この報告では、「統計的検定」に限定して、学生の躓く箇所と原因の幾つかを明らかにし、よりよい教え方／学び方を検討する。さらに、従来の多くの教科書では触れられていない重要な点について言及する。

2. 統計用語の問題

第一の躓きの原因は統計用語の分かりやすい説明が欠けていることである。日本語の訳語よりも原語の意味を理解させた方が分かりやすい場合が多い。たとえば、以下にみるように、英語の方が意味が明快である。

- **有意** significant 意味がある, 重大な, 重要な。
large or important enough to have an effect or to be noticed
having a particular meaning (OALD)
- **受容する** accept 採択する
to receive something as suitable or good enough
to agree to or approve of something
accept は、論理的に評価して、適切であるとして、受容することである。
- **棄却する** reject 否定する
to refuse to accept
- **帰無の** null 価値がない。ゼロの, 無効の
without value, effect, consequence, or significance.
have the value zero
「帰無仮説」は、本来は無に帰する仮説とされるが、⁽¹⁾ 価値のない仮説と考えてもよい。そこで、「帰無仮説」は、価値がゼロであるから、棄却(否定)してはじめて、仮説検定が意味を持つことが理解できる。

帰無仮説の用語の問題は、後述する (3.2)。

- **対立の** alternative 代替の, 代わりに選ばれるべき, 二者択一の
that can be used instead of something else
「対立仮説」は、帰無仮説に代わって選ばれるべき仮説のことである。統計学では、これを「対立」と訳しているが、これは誤訳に近く、適切な訳ではない。「代替仮説」の方がよい。

3. 仮説検定の論理と手続き

第二の躓きの原因は、仮説検定の論理では「背理法(反証法)」が使われていることである。すなわち、仮説検定は、帰無仮説が正しいとして論を進めて、その仮説のもとできわめて起こりにくい現象が起きていることを導くことによって帰無仮説を棄却し、対立仮説を採択するという「背理法」の考え方をとっている。⁽¹⁾⁽²⁾⁽³⁾ しかも、この検定の論理では、本当に証明したいことを帰無仮説にしても、帰無仮説の正しさはそれ自体で証明されることはないことを理解することは、きわめて重要である。⁽⁴⁾⁽⁵⁾

3.1. 帰無仮説の立て方

第三の躓きの原因は、帰無仮説の設定で、なぜ「差がない, 差が null (ゼロ) である」を立てるかということである。「差がある, not zero」の仮説を立てようとする、大きな差・中位の差・小さな差があるなど無限に仮説が立てられ、一つに定まらないので、検定の手続きを進めることができないことをきちんと説明する必要がある。

3.2. 帰無仮説の性質

第四の躓きの原因は、帰無仮説として「差がない」「ゼロである」ことを、どのように理解するかである。分析者の立場により、「差がない」ことを次のように分類できる。

- (1) 「差がない」ことが望ましくない場合

例：新製品の開発（新薬や新品種の開発など）

（２）「差がない」ことが望ましい場合

例：製品の生産（品質管理など）、擬似乱数の生成、放射線被曝など

ただし、「差がない」ことが望ましくないこともあれば、望ましいこともある場合もある。例：男女差、無罪・有罪など。後者の場合、被疑者と検事では望ましいことは、正反対であることが普通である。

第五の躓きの原因は、仮説検定が本来の上記

（１）だけでなく、（２）の場合にも使われることである。この場合にも通常、帰無仮説という用語が使われるが、無に帰することになる（１）のように否定の目的で設定されていると解釈することは適切ではない。そこで、帰無仮説という用語を避け、「ゼロ仮説」、あるいはもっと一般的に検定の直接の対象となる仮説をすべて「検定仮説」と呼んでいる場合もある。⁽⁶⁾

他方で、「帰無」は、‘無に帰することも予定して’くらいの意味で深い意味はないので、「帰無」の語を無視しても差し支えないという著者もいるが、⁽³⁾ こういう曖昧な説明では、学生の頭を混乱させるだけであるので、好ましいとはいえない。

3.3 検定における分析者の異なる二つの立場

第六の躓きの原因は、仮説検定において、上述のように、「差」をどのように理解するかにより、分析者には異なる二つの立場があることを、統計学の教科書がきちんとした説明していないことである。すなわち、帰無仮説：差がない、対立仮説：差がある は同じであっても、分析者がものごとの「非同一性」（差があること）を主張したいのか、「同一性」（差がないこと）を主張したいのかで、仮説検定の考え方に次のような明確な違いがある。

分析者の立場（a）

「非同一性」を主張したい場合

たとえば、新製品の開発の場合には、旧製品／他製品の特性との非同一性、差別化、効果の違いに価値がある。この場合には、帰無仮説は、分析者にとっては、これが文字通り、無に帰すべき、価値のない仮説である。差がないという帰無仮説を棄却することに、積極的な意味がある。帰無仮説の棄却により、対立仮説を採択し、新製品の特性の非同一性を判定する。例として、新薬の開発をあげる。

■ 帰無仮説：新薬と偽薬で差がない。

→ これは価値がなく、分析者が主張したいことの否定となっている。

■ 対立仮説：新薬と偽薬で差がある。

→ これは、分析者が主張したいことである。この場合、帰無仮説が棄却されれば、対立仮説が採択され、新薬と偽薬に違い（非同一性）があり、新薬に効果ありという結論が得られる。

分析者の立場（b）

「同一性」を主張したい場合

たとえば、製品の継続的安定生産（品質管理など）では、製品の特性の同一性、不変性に価値がある。この場合には、帰無仮説は、分析者にとっては、本当は無に帰しては困る仮説であり、分析者が真であることを証明したい、価値のある仮説である。差がないという帰無仮説を棄却することは、自らの主張を否定することになる。他方で、帰無仮説が棄却されなかったからといって、帰無仮説が積極的に支持されるわけではない。ここで、仮説検定では帰無仮説が真であること自体を証明できないという論理となっていることに注意させることがきわめて重要である。例として、品質管理の問題をあげる。

■ 帰無仮説：品質に差がない。

→ これは価値があり、分析者が主張したいことである。これは本来、帰無仮説としては適切でないが、帰無仮説としては「差がない」という仮説しか立てられないので、帰無仮説にせざるをえない。

■ 対立仮説：品質に差がある。

→ これは価値がなく、分析者が否定したいことである。

この場合、帰無仮説が棄却されれば、対立仮説が採択され、非同一な不良品の排除ができる。

学生には、分析者には（a）、（b）の二つの立場があることをきちんと理解させ、それぞれの立場から、統計的検定の演習を行わせる必要がある。

3.4 帰無仮説を棄却できないときの解釈

第七の躓きの原因は、帰無仮説を棄却できないときに、どのように判断すればよいか、学生の多くに迷いが生じることである。統計学の教科書は、帰無仮説が棄却できないときの判断を、次のように説明している。

（１）何もいえない。何もわからない。

（２）判断を保留する。

（３）帰無仮説とは矛盾しない。

(4) 帰無仮説を採択（受容）する。

帰無仮説が棄却できないときの判断で、いちばん厳しく慎重なのは、上記の(1)である。⁽⁷⁾⁽⁸⁾ 帰無仮説が棄却できないときには何もいえないという結論では、決定を下さなければならない分析者はデータを前にして困る場合がある(後述 3.5)。次に厳しいのは(2)の保留である。帰無仮説を棄却できないのは、帰無仮説を棄却するための十分な証拠が得られなかったのであって、帰無仮説を採択して差し支えないと結論するのは、推論の間違いであり(2)とするのが正しいとする。⁽²⁾⁽⁹⁾ 帰無仮説とは矛盾しない(3)は、帰無仮説が棄却されなかったからといって、帰無仮説が積極的に支持されたわけではないという判断である。⁽³⁾ 帰無仮説を採択する(4)の判断がいちばん甘い。単に「採択する」と記述しているだけの教科書も少なくない。⁽¹⁰⁾ しかし、帰無仮説を採択したからといって、帰無仮説が真であることが積極的に「証明」されたわけではない。⁽³⁾ 上記(3)(4)の場合には、多くの著者は、検定統計量の分布において、その全領域を「棄却域」と「採択域」の二つの区域に分割している。ここで、「採択域」の代わりに、「保留域」とすることはないようである。

“帰無仮説は棄却することはあっても「採択」してはいけないという注意は英語の文献ではしょっちゅう見かけるが、(日本語の文献では)、著名な人でもこの表現を普通に使っている”という指摘がある。⁽¹¹⁾ アメリカ心理学会は、統計的方法の特集号の会誌で、Never use the unfortunate expression “accept the null hypothesis.” と明確に述べている。⁽¹²⁾

3.5 帰無仮説を事実上「採択」する場合

分析者の立場(b)では、帰無仮説を棄却できないときに、「何もわからない」「何もいえない」では、製品の継続的安定生産や生成に支障をきたすので、帰無仮説を事実上「採択」している場合が少なくない。たとえば、酒の品質管理⁽¹³⁾ や擬似乱数の乱数性の検定などである。擬似乱数の検定では、 p 値 \geq 有意水準のときには、帰無仮説を採択して、“良い乱数”としている。⁽¹⁴⁾⁽¹⁵⁾

3.6 帰無仮説を保留するケース

分析者の立場(b)では、帰無仮説を棄却できないときには、分析時点での判断を保留すること、さらに継続してデータを収集し分析していくことが求められる場合がある。例として、子どもの甲状腺ガンの発生率の増加が原発事故によるもの

のかどうかについて、統計的検定をおこなう場合をとりあげてみる。

帰無仮説：事故の前後で発がん率に差がない。

対立仮説：事故後に発がん率が増大した。

を立てる。この検定の手続きでは、次の結果が得られる。

(1) 帰無仮説を棄却して、対立仮説を採択する。

「事故後に発がん率は増大した」と結論する。この結論が間違っている確率が危険率であるが、この検定の結論は、ありがたくない結論ではあるとしても、検定の論理から、子どもと親は納得して受け入れざるを得ないであろう。

(2) 帰無仮説を棄却できないときに、帰無仮説を採択してよいか。

帰無仮説を棄却できないから、帰無仮説を採択して、「事故の前後で発がん率に差はない」「事故の影響は見られないので、安心してよい」という結論を出されて、子どもと親は納得して、これを受け入れるだろうか。検定の解釈としては、現在のデータでは、事故の影響については「何もわからない」、「保留」という判断が適切である。この判断では、数年後に発がん率が増大する可能性があることが否定されたわけではないので、今後、継続して検査を受ける必要があることを、子どもと親に伝えなければならないだろう。このような実際例に接すれば、学生は、統計学が現実社会の問題把握に役立つことがよく理解できるはずである。

4. 統計的エラーとリスクの問題

第八の躓きの原因は、統計的エラー(第1種と第2種の誤り)と消費者あるいは生産者のリスク(危険)が、分析者の立場(a)と(b)では逆になることが、教科書で明快に説明されていないことである。ただし、一方のリスクは、結局は他方のリスクにもなるので、消費者/生産者のリスクという用語は誤解を招くだけであり、使わないにこしたことはない。

第1種のエラーとは、帰無仮説が正しいのに、これを棄却する誤りである。第2種のエラーとは、帰無仮説が間違っているのに、これを棄却しない誤りである。

分析者の立場(a)：新製品の検定の場合

第1種のエラーは消費者のリスクであり、第2種のエラーは生産者のリスクである。

第1種のエラーは、良品と違って不良品を購入した消費者が損失を受ける誤りであるから、消費者のリスクである。⁽¹⁶⁾⁽¹⁷⁾ これは、不良品を買われた消費者が生産者を訴えて賠償を求めるこ

とになれば、生産者のリスクでもある。

第2種のエラーでは、良品であるのに、不合格であると判断するものである。このエラーは、生産者がぼんやりして大もうけの機会を見逃してしまうものであり、生産者が損失を受ける誤りであるから、生産者のリスクである。⁽¹⁶⁾⁽¹⁷⁾ これは、良品が安定して供給されないことになれば、消費者のリスクでもある。

分析者の立場 (b) : 品質管理の検定の場合

第1種のエラーは生産者のリスクであり、第2種のエラーは消費者リスクである。

第一種のエラーは、合格すべき良品を不合格と判断する誤りであり、生産者が損失を受ける誤りであるので、生産者のリスクである。⁽³⁾⁽¹⁸⁾ これは、良品が安定して供給されないことになれば、消費者のリスクでもある。

第二種のエラーは、不合格とすべき不良品を合格と判断するという誤りであるから、消費者のリスクになる。⁽³⁾⁽¹⁷⁾ これは、不良品を購入した消費者が生産者を訴えて賠償を求めることになれば、生産者のリスクでもある。

5. 要約

いままでの結果を表にしてまとめておく。

表1 分析者の立場により異なる仮説検定

項目	分析者の立場 (a)	分析者の立場 (b)
主張	非同一性	同一性
帰無仮説	無価値	価値あり
対立仮説	価値あり	無価値
帰無仮説	否定したい	肯定したい
帰無仮説	棄却	棄却
対立仮説	採択	採択
対象	良品	不良品
第1種エラー	消費者のリスク	生産者のリスク
第2種エラー	生産者のリスク	消費者のリスク

6. おわりに

統計的検定において、学生の躓きやすい箇所を幾つか指摘し、躓きの原因を検討した。学生に統

計的検定の方法を身につけさせるためには、入門用の教科書には特に丁寧で分かりやすい、曖昧でない説明が求められていることを指摘した。

著者の統計学の理解に誤りがあることを怖れる。ご指摘とお許しを願う次第である。

参考文献

- (1) R.E.Henkel, 松原 望・野上佳子訳：“統計学検定—統計学の基礎—”, pp.31-42, 朝倉書店 (1982)
- (2) 栗原伸一：“入門統計学”, pp.92-95, オーム社 (1996)
- (3) 東大教養学部統計学教室：“統計学入門”, pp.233-239, 東京大学出版会 (1991)
- (4) 佐伯 胖・松原 望：“実践としての統計学”, pp.44-51, 東京大学出版会 (2000)
- (5) C. Taylor: Why Say Fail to Reject, <http://statistics.about.com/od/Inferential-Statistics/a/Why-Say-Fail-To-Reject.htm>
- (6) 森田優三・久次智雄：“新統計概論 改訂版”, pp.27-278, 日本評論社 (1993)
- (7) 小針暁宏：“確率・統計入門”, p199, 岩波書店 (1973)
- (8) 柴田文明：“確率・統計”, pp.121-127, 岩波書店 (1996)
- (9) 市原清志：“バイオサイエンスの統計学”, pp.14-16, 南江堂, 1990
- (10) 稲垣宣生・山根芳知・吉田光雄：“統計学入門”, pp.114-117, 裳華房 (1993)
- (11) 奥村晴彦：“帰無仮説を採択」？ <http://oku.edu.mie-u.ac.jp/~okumura/blog/node/2492>
- (12) L. Wilkinson: “Statistical Methods in Psychology Journals: Guidelines and Explanations”, American Psychologist, Vol.54, No.8, pp.594-604 (1999)
- (13) 佐藤 信：“推計学の手ほどき”, p.96, 南江堂 (1970)
- (14) 伏見正則：“乱数”, p.109, 東京大学出版会出版会 (1989)
- (15) 東京理科大学工学部金子研究室：“擬似乱数生成系の検定方法に関する調査報告書” (2004) www.cryptec.go.jp/estimation/rep_ID0211.pdf
- (16) 蓑谷千風彦：“推定と検定のはなし”, p.156, 東京図書 (1988)
- (17) 青木繁伸：“統計数字を読み解くセンス—当確はなぜすぐにわかるのか?”, p.130, 化学同人 (2009)
- (18) 村上征勝：“工業統計学”, p.76, 朝倉書店 (1985)