

第6回統計教育シンポジウム

「仮説検定の考え方」のロジック

～ 仮説検定の本当の考え方とは～の授業実践を振り返って

お茶の水女子大学附属高等学校 数学科

三橋一行

「仮説検定の考え方」

学習指導要領解説では

(4) データの分析

ア(ウ) 具体的な事象において、**仮説検定の考え方を理解する。**

イ(ウ) 不確実な事象の起こりやすさに着目し、主張の妥当性に着目し、主張の妥当性について**実験などを通して判断したり**、**批判的に考察**できたりすること。

➡ 仮説検定の本質的考え方を教えるものと期待

実際の教科書を見てみると……

仮説検定を習って(教えて) いないのに考え方を教える。

→ 仮説検定の考え方というより 仮説検定の仕方を説明

→ 未習事項のカバーのため

検定のためのデータ収集のほかに 別の実験をプラスして、

実験 (標本調査) を実験で保証する

という 「仮説検定もどき」 のやり方

高校1年生にとって、 仮説検定（フィッシャー；ネイマン・ピアソン流） を教えるにあたって難所と思われる内容

- ①母集団と標本の関係（母集団分布，標本分布）の問題
- ②確率変数，確率分布，確率を求める（ Σ ，積分の）問題
これらをカバーするために → シミュレーションやPCの使用の問題，高1では，二項分布さえ知らない状態。
- ③統計量の問題，正規化，確率変数から統計量の導出
- ④帰無仮説，対立仮説の問題（設定法，意味，分布の属性）
- ⑤有意水準（危険率）問題（意味や過誤の問題）
- ⑥仮説の棄却に関する問題（ H_0 が棄却なら H_1 は採択？） など・・・

統計学の本の最後の方にある
「仮説検定」を最初に持ってくるなんて・・

という状況ではあるのですが・・・

- ① 単純化して、**仮説検定の本質的な考え方を指導できないか**
- ② 仮説検定の**危険性や発展的な方法も扱えないか**

3. 指導法の提案

<問題>

「A, B二つの壺がある。

Aの壺には, 白玉9個, 黒玉1個,

Bの壺には, 白玉2個, 黒玉8個

が入っている。今, 目の前にどちらか1つの壺

が置いてある。そこから, 無作為に1つの玉を取り出

したら, 黒玉であった。Aの壺か, Bの壺か判断せよ。」

[考え方 α]

確率的に考えてみると

判断をしてみる！

Aである確率 0.5	Bである確率 0.5
白玉が出る $0.5 \times 0.9 = 0.45$	白玉がでる $0.5 \times 0.2 = 0.1$
黒玉が出る $0.5 \times 0.1 = 0.05$	黒玉が出る $0.5 \times 0.8 = 0.4$

黒玉がでたので、左図の色の塗ってある部分だけを考えればよい。

(黒玉がAの壺から出ている確率) :
(黒玉がBの壺から出ている確率)
 $= 0.05 : 0.4 = 1 : 8$

黒玉がBの壺から出てきた確率は、Aの壺ら出てきた確率よりも8倍も高い。
⇒ よって、Bの壺である可能性が高い。

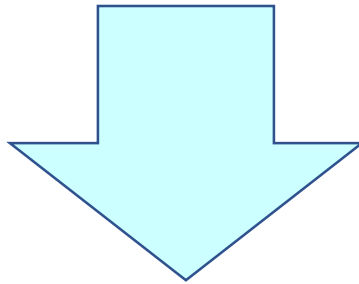
Aの壺から出る確率は $0.05 / (0.05 + 0.4) = 1/9$
Bの壺から出る確率は $0.4 / (0.05 + 0.4) = 8/9$

→ したがって、Bの壺であると考えられる。

この方法にどこか、不自然なところはないか…

最初の設定で、

「Aの壺かBの壺であるかが、確率0.5 (1/2) ずつ」としてよいのだろうか



別の方法を考えよう！！

[考え方 β]

背理法的に考えてみると

「困った時の背理法！？」

壺Aがすべて白玉，壺Bがすべて黒玉とすると・・・はっきり言いきれぬが，壺Aが白玉9個，黒玉1個，壺Bが白玉2個，黒玉8個とすると・・・

1. 壺Aだと仮定する。（壺Bであることを否定する）
2. 壺Aならば，**ほぼ**白玉が出る。
3. 無作為に壺から1つ取り出したら黒玉であった。
4. 壺Aでない**とは言い切れないが，可能性が低いので壺Bだろう。**
何か弱い・・・

そこで、

確率 $1/10$ 以下で起こることは、起こりにくい事なので「起こらない」としてしまおう。

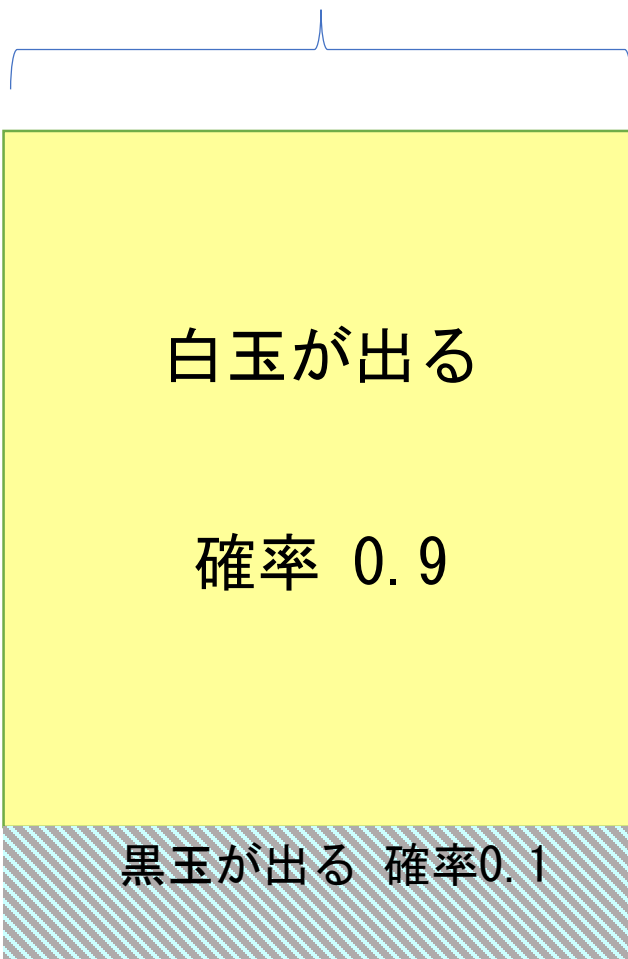
壺Aが白玉9個，黒玉1個，壺Bが白玉2個，黒玉8個

1. 壺Aだと仮定する。（壺Bであることを否定する）
2. 壺Aならば，白玉が出る。←確率 $1/10$ で起こる黒玉は出ないと考えた。
3. 無作為に壺から1つ取り出したら黒玉であった。
4. 壺Aであることに矛盾する。よって壺Bである。

判断はできたが，壺Aから確率 $1/10$ で黒が出ている可能性を無視したので，この判断は $1/10$ の確率で間違っている危険性をもっている。

めったに起こらないことなので目をつぶったこの確率を**有意水準（危険率）**という。

壺Aである



確率 $1/10$ 以下の事象は
起こらないと考えると

壺Aから黒玉が出るとは
考えられない。

(正しくは考えにくい)

→ 壺Bである

有意水準（危険率）とは？

「考え方β」で、判断はできたが、壺Aから確率1/10で黒が出ている可能性を無視したので、この判断は1/10の確率で間違っている危険性をもっている。

めったに起こらないことなので目をつぶったこの確率を有意水準（危険率）という。

有意水準（危険率）は0.05を使うことが多い。

100回に5回起こる，すなわち20回に1回程度しか起こらない事象は「起こらない」と判断するということ。

諸説あるが，数学的，科学的根拠はない。

フィッシャーの「毎年推定している場合，20年に1度のミスは許される」との発言によるものではないかとされている。

大きくすれば，判断はつき易いが，間違った判断も多くなる。

小さくすれば，間違った判断は減るが，判断が付きにくくなる。

問 1.

[考え方 α] と [考え方 β] を聞いて、以下の問いに答えなさい。

① どっちの考え方が、良いと思いますか？

理由も書きましょう。

② それぞれの考え方に問題点があるとするば、それはどこだ

と思いますか？ 理由も書きましょう。

① どちらの考え方が、良いと思いますか？理由も書きましょう。

考え方 A.

考え方 A は導出した確率から判断をしいて、
考え方が論理的で筋が通っている。

それに反して、考え方 B は壺 A から黒玉が出る $\frac{1}{10}$ の
確率を無視しているため、論理性が欠けている。

結論が明確に出ているのは考え方 B だが、
B は論理的ではないため、考え方が良いと思う。

① どちらの考え方が、良いと思いますか？理由も書きましょう。

考え方Aの方が良いと思います。

なぜなら考え方BはつぼAにも黒五が出る確率があることを無くしてしまっているため正確ではないから。

考え方AはつぼA、つぼBそれぞれの確率を調べた上で結論を出すため全ての場合が含まれている。

② それぞれの考え方に問題点があるとしたら、それはどこだと思いますか？

理由も書きましょう。

A 今回は、 $1=8$ であるから、⁺「可能性が高い」を判断
しにくいから、 $4=5$ などのときに活用できる。

B 「起りに cu 」と振動確率が $\frac{1}{10}$ で、⁺「高い」では
ないかと考えた。人に $1, 2$ 、無視確率が異なり、 2 以上
の cu とないかと考えた。

② それぞれの考え方に問題点があるとすれば、それはどこだと思えますか？

理由も書きましょう。

〈考え方A〉

黒玉がBの壺から出てきた確率は、Aから出るよりも8倍高いが、

Aから出る確率が0ではない → 判断が宇宙...?

全ての確率を比べる → 低い値(外れ値)も含む
⇒ AI的な分析...? 機械的な印象。

〈考え方B〉

「めったに起こらないこと」の基準が不明。あまり

確率の低い事象は排除する(外れ値を除外)

⇒ 人間心理的な分析...?

① どちらの考え方が、良いと思いますか？理由も書きましょう。

考えがBであれば"どのくらい"の確率で"その考え方が誤りになるか"明白にわかるのでBの方が良い。
またこの例でもし壺が"たくさん"あったりする場合、Aは計算が"大変"になってしまふ。

② それぞれの考え方に問題点があるとしたら、それはどこだと思いますか？

理由も書きましょう。

A ・ 計算大変

B ・ ϵ の α の確率まで「起るはずない」として ϵ の α の場合、
有意水準の判断が難しい (主観的)

① どちらの考え方が、良いと思いますか？理由も書きましょう。

B. Bは $\frac{1}{10}$ だけ間違えるが、

Aは $\frac{1}{9}$ だけ間違える。

Bの方がよい
とAは

② Aの方がよい。

(Bは「Bの都合がよい上」に考えられている？
Aの方が数値に基づいている。)

① どちらの考え方が、良いと思いますか？ 理由も書きましょう。

考え方A

考え方Bでは $\frac{1}{10}$ 以下の確率を無視してしまっているので、誤差が大きい。
対し、考え方Aはどちらの市が出やすいかという結論に加え、どちらが
どのくらい出やすいのかというところまで着くため、より正確だと考えるから。
ただし、考え方Aでは確率の大小でしか結論がつけられないため、
どちらが出るかという結論が確定しているBの方が、主張をするのには
むいていると思った。

問 2.

【考え方B】 と比べて、上の①～⑥の表現でふさわしくないと
思われる表現を用いているのは、どれですか？ 理由とともに
答えなさい。

問 3.

有意水準（危険率）を決めるのは、③より前に決めておくことが
よいとされています。その理由を考えてみましょう。

プリントの問題を解いてみよう！

問2.

[考え方B] と比べて、上の①～⑥の表現でふさわしくないと思われる表現を用いているのは、どれですか？ 理由とともに答えなさい。

⑤ 絶対に「ダメだ」と言い切ることができない。

↓
((仮説検定はと2もリステ))

めそつぷた確率

↓

問3.

有意水準（危険率）を決めるのは、③より前に決めておくことがよいとされています。その理由を考えてみましょう。

③の値が分かってから決めてしまうと②の値を自分の都合の良い確率にしてしまう可能性があるので。

②で決めた方が公正。

問3.

有意水準（危険率）を決めるのは、③より前に決めておくことがよいとされています。その理由を考えてみましょう。

結果が出た後有意水準を意図的に操作して結論を自分の思う方に持っていくという危険性があるから。

授業を終えて

- 仮説検定の考え方として、ベイズの方が、NP仮説検定よりもしっくりにくるようだ。しかし、説明の長短、有意水準を「無視する確率」といっているなどの影響があると思われる。(数学らしいので自然に受け取れるようである。)
- 仮説検定（夏休みの宿題でやり方だけは学習していた）ので、どんなことをしているのかのイメージがつかめたという生徒が多かった。
- 危険性を孕んでいる判断であることが分かった。
- グループ活動で意見交換が活発に行われた。
- 代表生徒の発表内容が充実したものであった。(予想外)

<再掲>

高校1年生にとって，仮説検定（フィッシャー；ネイマン・ピアソン流）を教えるにあたって難所と思われる内容

- ①母集団と標本の関係（母集団分布，標本分布）の問題
- ②確率変数，確率分布，確率を求める（ Σ ，積分の）問題
これらをカバーするために → シミュレーションやPCの使用の問題，高1では，二項分布さえ知らない状態。
- ③統計量の問題，正規化，確率変数から統計量の導出
- ④帰無仮説，対立仮説の問題（設定法，意味，分布の属性）
- ⑤有意水準（危険率）問題（意味や過誤の問題）
- ⑥仮説の棄却に関する問題（ H_0 が棄却なら H_1 は採択？） など・・・

授業を終えて②<協議会での話題から>

- **別実験**を行う → **理論値で評価したい**。 → **無理せず、数Bまで待つ**
「**考え方**」を簡単な例で説明することに重点
- 2つの仮説が**論理的に排反でない**ことがある
→ **数理統計学上は補集合の関係（排反）、実際は不可能。**
 $\mu A = \mu B$ 1点での（単純）仮説と $\mu A \neq \mu B$ その他の（複合）仮説
とで考える。実は、帰無仮説 $\mu A = \mu B$ と対立仮説 $\mu A < \mu B$ の設定は不自然。
 $\mu A > \mu B$ の可能性が除かれている。仮説が立てられるのは、**パラメータ（母数）のみについてのみ** **競争に勝つ主張をすべき（現実的考えかた）**
or ベイズ統計
- 問題となる確率のみでなく、なぜ**それ以上**の確率をもとめるのか
→ **分布の形状とのそれらのズレ、ヒストグラム作成** **数Bまで待つ**
- 有意水準の設定について（**棄却域の生成**こそが要）
→ 「**見ないことにする確率**」（今回） **帰無仮説の分布前提（モデル）からの矛盾**
- データサイエンス、統計学の結果の**信頼性**は？（条件、リスクを背負った判断）

＜参考文献＞

- 「高等学校学習指導要領（平成30年告示）解説 数学編 理数編」 文部科学省
- 「データ分析に必須の知識・考え方 統計学入門」阿部真人 著 ソシム
- 「完全独習 ベイズ統計学入門」 小島 宏之 著 ダイヤモンド社
- 「統計学を哲学する」 大塚 淳 著 名古屋大学出版会