

2023年3月19日(日) 15:00～

理数系教員統計・データサイエンス 授業力向上研修集会 宮崎

仮説検定の考え方」の指導について

～実践を通して考える仮説検定のロジックと指導のポイント～

お茶の水女子大学附属高等学校 数学科

三橋一行

数学 I「仮説検定の考え方」

＜例題＞ （自作）

シャープペンシルを製造している会社が、すでに販売している製品Aを改良して、製品Bをつくった。そして、消費者にBがAよりも書きやすいと評価されるかどうかを調査したいと考えた。そこで、無作為に選んだ20人にこれらのシャープペンシルを使ってもらい、A、Bのどちらが書きやすいか回答してもらった。その結果、15人がBと答えたことがわかった。この結果から、製品Bが書きやすいと消費者に評価されていると考えて良いか。基準となる確率を0.05とし、次のコイン投げの実験結果を利用して考察せよ。

実験 公正な1枚のコインを投げる。そして、コイン投げを20回行うことを1セットとし、1セットでの表の出た枚数を記録する。この実験を200セット繰り返し、次の表のような結果となった。

表の枚数	5	6	7	8	9	10	11	12	13	14	15	16	計
度数	4	10	15	20	28	33	29	25	21	12	3	1	200

主な現場の声①……

- ①「20人中15人なら75%はBを選んでいる、面倒なことしなくてもBが良いのは予想がつく。……」
- ②「表から15回表が出るのは、 $3/200$ ($\doteq 0.015$) の確率でしか起こらないから棄却する……」⇒ 16回表が出る確率をたさない。また、「なぜか」と質問がでる。
- ③ 仮説Ⅰ（教科書では主張） $P=1/2$ と仮説Ⅱ $P>1/2$ が背反（補集合の関係）ではない。 $P<1/2$ は考えないのか。 $P\neq 1/2$ なら理解できる。
- ④ なぜ、コイン投げとかサイコロ投げの実験をしなければならないのか、そんなことをして良いのか、調査（実験）を実験で比較（理論なし）。実際に実験させてみたら、なかなか（二項分布）に収束しない。
- ⑤ 基準となる確率は何か、なぜ0.05なのか。

主な現場の声②……公開研協議会から

- 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）

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・・・

これらに深入りしないためには，（準）仮説検定をするマニュアルにそって分析するのがよいだろう。……

「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

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

壺Aである確率 0.5

壺Bである確率 0.5



--	--

壺Aである確率 0.5

壺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 $0.5 \times 0.8 = 0.4$
黒玉が出る $0.5 \times 0.1 = 0.05$	

壺Aである確率 0.5

壺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 $0.5 \times 0.8 = 0.4$

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。

[考え方 β]

背理法的に考えてみると

「困った時の背理法！？」①

壺Aが**すべて白玉**， 壺Bが**すべて黒玉**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，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であることに矛盾する。よって壺Bである。

「困った時の背理法！？」②

元に戻して、

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。**

「困った時の背理法！？」③

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。

壺Aである

壺Bである

先程は、並列
的に扱ったが、

状況から
壺Bであると
思われるので、

壺Aである

壺Aであること
を仮定する。

※あえてAを選ぶことに
注意！

壺Bの部分なしで考える。

白玉が出る

確率 0.9

黒玉が出る 確率0.1

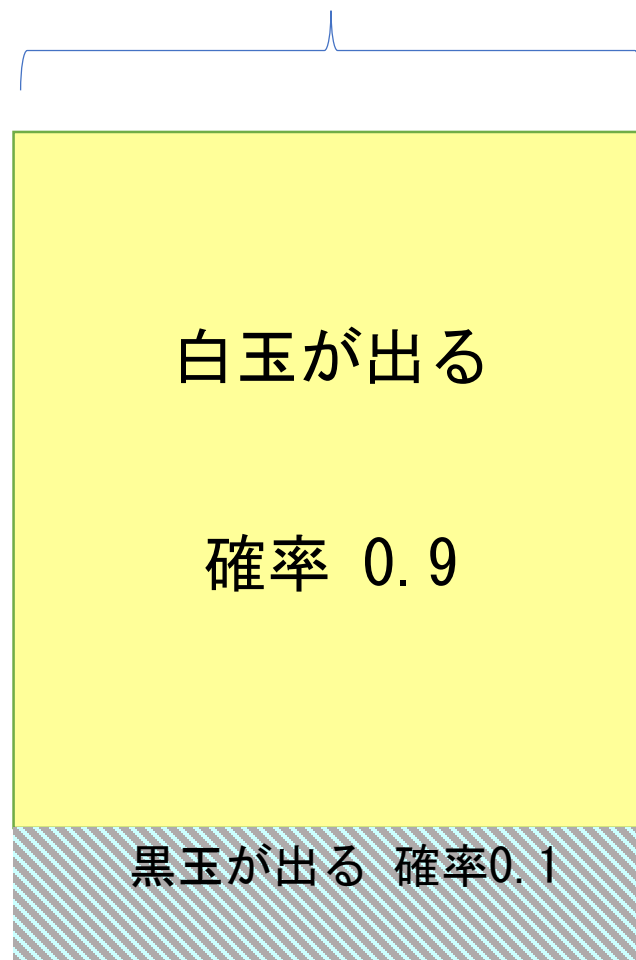
壺Aである

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



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れば、それはどこだ
と思いますか？ 理由も書きましょう。

実は、

[考え方 α]

ベイズ統計学による仮説検定の基本的考え方。

- ・ 事前分布 → 理由不十分の原理
- ・ 確率が主観的 → 科学的でない。

[考え方 β]

フィッシャー・ネイマン・ピアソン流の仮説検定の基本的考え方。

- ・ 反証に堪えうる仮説を残す → 科学的にした。
- しかし、 . . . 。

反証主義 哲学者カールポパーの多大な影響を恐れた！！

歴史的に見ると、

最初

ベイズ統計学 主観的要素が多い。

批判



フィッシャー・ネイマン・ピアソン流
の仮説検定

(科学的に。反証主義 ポパーの発言)



再び

ベイズ統計学 (AI, コンピュータの発達)

背理法（反証） のようにするのは， 誤りを防ぐため。

後件肯定の誤り。論理的に正しくない推論

もし P ならば、Q である。

Q である。

したがって P である。

（例）

インフルエンザにかかっているなら、高熱がでる。

今、私は高熱を発している。

したがって、私はインフルエンザにかかっている。

→ インフルエンザ以外にも高熱の出る病気は
あるので明確な誤り。

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。

「仮説検定の考え方」

※ 主張 → 仮説と置き換える。

教科書の内容を補足して下のようになりました。

仮説〔1〕が正しいと判断できるか。

- ① 仮説 [2]（仮説 [1] と反する仮定）を立てる。
- ② 有意水準＜危険率＞（これ以下の確率は起こらないと考える）を決める。
- ③ 仮説 [2] のもとで、実際に起こった出来事が起こる確率を調べる。
- ④ ②と③で求めた確率を比べる。→ 小さかった
＜大きかったら、仮説 [2] を棄却できない＞
- ⑤ そもそも、仮説 [2] の仮定が正しくなかった。
- ⑥ 仮定 [1] は正しいと判断しても良いと考えられる。

教科書の内容を補足して上のようにしました。

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 どちらの考え方が、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れば、それはどこだ
と思いますか？ 理由も書きましょう。

「仮説検定の考え方」（教科書をもとに再掲）

※ 主張 → 仮説と置き換える。
教科書の内容を補足して下のようになりました。

仮説 [1] が正しいと判断できるか。

- ① 仮説 [2] （仮説 [1] と反する仮説（仮定））を立てる。
- ② 有意水準 < 危険率 > （これ以下の確率は起こらないと考える）を決める。
- ③ 仮説 [2] のもとで、実際に起こった出来事が起こる確率を調べる。
- ④ ②と③で求めた確率を比べる。→ 小さかった。
＜大きかったら、仮説 [2] を棄却できない＞
- ⑤ そもそも、仮説 [2] （の仮定）が正しくなかった。
- ⑥ 仮説 [1] は正しいと判断しても良いと考えられる。

教科書の内容を補足して上のようになりました。

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！

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）

＜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