「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―

数学科　　三　橋　一　行

2022 年度から，数学Iの授業で「仮説検定の考え方」が導入された。この単元の教科書の内容および学習指導要領解説に載っている指導事例は「仮説検定の劣化版の使い方指導」というべきものである。仮説検定の真の考え方とは何か。仮説検定そのものは，第 2 学年での数学 B で学ぶことになっている。したがって，数学Iでの仮説検定の学習には数学的準備が足りなさすぎる。また，数学 B においても仮説検定の考え方は十分に説明されていない。これらのことを踏まえて，数学Iの単元「仮説検定の考え方」で学ぶべき内容を検討し，授業実践を行ってきた。そして，本校の第 26 回公開研究会における研究授業でその実践を発表した。その内容と合わせて報告する。

〈キーワード〉新学習指導要領　仮説検定　仮説検定の考え方　統計的判断　データサイエンス

1.　はじめに

2022 年度から，第 1 学年の数学Iの授業で「仮説検定の考え方」の単元が導入された。「仮説検定」そのものは、第 2 学年の数学Ｂ「確率分布と統計的判断」の単元で扱われることになっている。しかも，数学Ｂを選択した生徒にとって必須となる。仮説検定の指導をする前に「仮説検定の考え方」を指導するとはどういうことなのか。さらに，仮説検定を学ぶために必要な数学的準備は，数学Iの段階では皆無に等しい。教科書の内容や学習指導要領解説に紹介されている例は良く工夫されているが，率直に表現すれば「仮説検定の劣化版の使い方指導」であって「仮説検定の考え方の本質」には達していない。このような「仕組み」に触れない授業内容は数学教育ひいては科学教育の目指す指導と逆行し，今後の日本における科学力・技術力に少なからず悪影響を及ぼすものと考える。そこで，この問題を少しでも回避するための授業を考えてみた。背理法的な考え方から確率の要素を踏まえて「仮説検定の考え方」の本質に迫り，さらにベイズ統計学の考えと対比させることで，今後注目されるベイズ統計学の基本的考え方を導入，その結果，統計学史上の大論争も疑似体験させることとなった。考案した授業が，今後，この単元指導における改善策の参考例となれば幸いである。

2.　仮説検定に関する問題点

2.1.　仮説検定の学問上の問題点

『実験や調査に関する学術雑誌は，これまで統計学における有意性検定の結果が有意であることを根拠に論文を刊行してきました。そこでは p 値と呼ばれる確率が 5%を下回ったか否かが，論文採否の主たる判断基準でした。この判定法は簡便で，客観的で，論文を書く手段として便利でした。しかし，重視されすぎたために，5%を切ること自体が目的となる傾向が生まれました。論文執筆の競争が激しいために，手段と目的の逆転現象がおきたのです。そして，不幸なことにあれこれ工夫すると学術的に無意味なデータから $p<0.05$ を示すことは，実に簡単でした。　－　（中略）－　その状況を憂い，統計学の総本山ともいえるアメリカ統計学会（American Statistical Association；ASA）は，統計的に有意でも科学的に無意味な論文をなくすために，「統計学的有意性と p 値に関する声明」を 2016 年に発表しました。そこでは，「科学的な結論や決定は p 値が有意水準を超えてたかどうかにのみ基づくべきではない。」と宣言されています。』また，『権威ある科学

誌 Nature は「統計的有意性を引退させよう」というタイトルの論文を 2019 年 3 月に刊行。800 人に及ぶ科学者がこれに署名，ASA の 2016 年の声明に対し，「我々はこの声明に同意し，統計的有意性の概念全体を放棄するよう求める」と述べた。』〔1〕

以上は参考文献からの引用で，ごく最近のアメリカの統計学に関する話題である。「それでもあなたは p 値を使い続けるのでしょうか？まだ有意性検定を教え続けるのでしょうか？」〔1〕，「だれか偉い統計学者たちが，価値ある研究とそうでない研究を峻別できる決定的な指標を早く決めて提案してほしい」〔1〕など，統計的仮説検定の使用及び教育の放棄と新しい指標を求める声が上がっている。悲しむべきは，アメリカが「仮説検定」を捨てようとしているこの時期に，日本は，「仮説検定」の高等学校生徒全員の必修化を開始することである。数学 B を履修することは選択制だが，選択すれば必ず仮説検定を学ぶカリキュラムである。また，数学 B を選択しない生徒に対しても数学 I で「仮説検定の考え方」という単元で多少なりとも学ぶことになるのである。せめて，仮説検定の危険性を教えかつ批判的に捉えていくような内容であれば，統計学の危険性や使用上の注意喚起などが行えて，統計的判断を見極める力を育てられるかもしれない。そしてそれは，学習指導要領にある批判的思考に十分当てはまるものであると考えられる。しかし，新カリキュラムの教科書では，ただ仮説検定を「便利なもの」として教えていく内容である。この結果，予想されるのは，アメリカの報告にあるように P 値を操作した（“p-hacking” と呼ばれている）論文等が多数発表されて混乱を招くという事態である。日本は，残念ながら統計教育が遅れてしまった。それを反省し統計教育に力を入れるのは良い事である。しかし，どの方向にどう教育していくかを考えられないほどに日本の統計教育が遅れ過ぎてしまっているのだと考える。文部科学省指定の SSH（スーパー・サイエンス・ハイスクール）の教員およびその指導者たちでさえ，統計学，特に仮説検定に対する認識に著しい誤りがある場合が多い。それでも，その学校が高い評価を受けていたりする。仮説検定を使用することがサイエンスなのではない。仮説検定の結論がその研究の結論にはならない。統計学（仮説検定も含めて）は「最強の学問」ではない。それにもかかわらず，統計学や仮説検定を神のごとく信奉する人々がいる。ぜひ，思い出してほしい，科学は神の存在を疑う，もしくは否定することから始まっていうことを。統計学を疑う。ここでは，仮説検定に限定して，仮説検定を疑う。仮説検定のどこを疑うべきか，その糸口が見つからないようでは，真に科学しているとはいえない。仮説検定が持っている客観的でどの分野の研究にも使えるというような利便性，その正体は『論理の飛躍』である。この点に関しては，大人たちより生徒たちの方が敏感に気付くのがせめてもの救いである。だからこそ仮説検定の危険性を含めて，正しく導く必要がある。ここでの「正しさ」とは論理的に正しいという意味である。

2.2. 仮説検定を高校 1 年生で教えることの問題点

学術的な問題点は上に述べた。ここでは，高校 1 年生が「仮説検定の考え方」を学ぶために必要とされる内容を考えてみる。単純に高校 1 年生が仮説検定を完全に理解するのに必要な内容を以下にあげてみた。

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・・・

以上である。細かく挙げればきりがないが，ここでは①～⑥とする。これは，仮説検定を学ぶのに必要なもの殆ど全てである。「仮説検定の考え方」を学ぶ前に仮説検定に関係する内容はほとんど学んでいないので仕方のない事である。これを見ると，仮説検定どころか仮説検定に必要な準備さえも殆どない時期に「仮説検定の考え方」を教えるのがいかに無理難題を示しているかお分かりいただけると思う。ところで教科書や傍用問題集ではどうなっているかというと，仮説検定は直接教えられないので，それを「ある仮説についての調査に対して，別のサイコロ投げやコイン投げの実験調査を行って比較する」という「仮説検定の劣化版」を行って判断するという作業を行えと指示される。このような方法に数学的価値があるかどうかはもはや問題外である。調査・実験ばかりで数学は登場しないからである。こんな無理をせずに，準備が整ってから仮説検定を学べば良いということに尽きる。少なくとも，数学 B で仮説検定を学ぶのであれば，かなりの部分は既習事項となっているので問題は激減する。それでも，数学 I で「仮説検定の考え」を指導するようになっているのは，数学 I が必修科目だからである。数学 B は選択必修科目なので，数学 B にのみ仮説検定を置いていれば，学ばずに高等学校を卒業する生徒が出てきてしまうからである。それを避けるために，「仮説検定の考え方」として，数学 I に入れられているのだそうである。聞いたところによると，数 I のここの単元で数学 B の仮説検定を教えてしまっても良いそうである。むしろその方がありがたいそうである。しかし，先に示した通り，ここで仮説検定そのものを学ぶには，準備が多すぎて時間が足りない上，1 年生にとっては内容が難しく不可能である。だからといって，「仮説検定の劣化版」を教えたところでそれが社会で何の役に立つのか。社会では，正規の「仮説検定」が使用されているし，下手をすると最初に述べた通り，社会に出たら「仮説検定」が放棄されているかもしれないのだ。統計教育を急ぐあまり，非常におかしな事態が発生している。

2.3.　問題点の精選と利用

さて，上述の通り多くの準備が必要なのだが，あえて精選して「仮説検定の考え方」の本質に迫ることを考えてみよう。

①　は中学校で少し触れるが，分布の話までには及んでいない。その他は，第 2 学年の数学 B で学んでから仮説検定に入る。（ちなみに，数学 B では仮説検定の単元に入る前に教科書で 30 ページほど事前準備をしてから学ぶことになっている。この 30 ページ間をどう乗り越えるかが数学 B で仮説検定を教える際のポイントとなる。本題に入る前に，統計学を嫌いさせないようにしなければならない。）②はとても時間がかかるものであるので，高校 2 年生の数学 B で学習せざるを得ないだろう。③はデータの分析の単元で多少触れるが，②の学習の後に効果を示すものなので，②を数学 B へ回す以上，③も数学 B で学ぶこととしよう。すると，①を中学校数学程度にとどめ，④，⑤，⑥の内容を簡単にして教える。しかも，この 4 点は仮説検定の本質的な部分である。少なくともこの 4 点を簡略化して具体例で分かりやすい指導法を考えれば，仮説検定の考え方の本質に迫れるはずである。

3.　仮説検定の考え方

3.1.　ベイズ統計学

2.1.で「だれか偉い統計学者たちが，価値ある研究とそうでない研究を峻別できる決定的な指標を早く決めて提案してほしい」〔1〕とアメリカで声が上がっていることを紹介した。このような指標は，簡単には見つからない。それに，万人に使いやすく魔法のように効果があるようなもの，そういったものこそ信用できない，信用してはいけないということも記した。先ほどの声に対する答えを求めるのは同じ過ちを求める可能性が大きい。ところが，現在の仮説検定法を完成させたネイマン(ｲｪｯｼﾞ・ﾈｲﾏﾝ 1894～1981)とピアソン(ｶｰﾙ・ﾋﾟ

ｱｿﾝ 1895～1980)という 2 人の統計学者たちを中心に否定され，長い間，忘れられていた数理統計学者ベイズ（トーマス・ベイズ（1701～1761)。そのベイズの考えたベイズ統計学が再び注目を集めている。教科書で扱っている仮説検定が含まれているのは，頻度主義の立場の統計学である。これに対して，ベイズ統計学は本質的に逆の手法をとり，確率がより活躍する。そのため数学的といえるが，数学が抽象的学問であるがゆえに最初に設定する確率が客観的でないといわれてしまっている。しかし，その点を除けば，数学的には自然な流れの思考によるもので親しみやすい人は多いはずである。以下に，この 2 者の比較表をあげる。

表 1　頻度主義統計学とベイズ統計学の比較表

項目	頻度主義統計学	ベイズ統計学
理論体系の中心者	フィッシャー，ネイマンとピアソン	トーマス・ベイズが創始，ラプラスが再発見
呼称	頻度主義者	ベイジアン
母数のとらえ方	定数	確率変数（不確実）
データのとらえ方	確率変数（不確実）	定数
特徴	母数は未知だが 1 つと仮定して得たデータが偶然の産物か否かを判断する手法。客観性が高く，これまでの主流。	母数は 1 つではなく，データの変化によって更新していく統計手法。少ないデータでも確率が求められ，自分で立てた仮説の正しい確率がわかる。
批判	仮説の判断には繰り返しが必要なため，実際にはそぐわないという意見がある。	最初に主観的な確率をおくため，「信念の確率」とも呼ばれ，客観性が低いといわれる。

頻度主義統計学では，何度も標本調査を繰り返すことで真の値を突き止めるという方法であるが，現在は 1 回で済ませてしまう場合が多いようである。これは間違いである。これはベイズ統計学の手法に属する。このようにこれらの 2 者混合も問題となっていると考えられる。しかし，今回は深入りしない。

次に，ベイズ統計学の考え方をわかりやすくするために次の問題をベイズ統計学の考えに基づいて解いてみる。

＜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。2 つの壺は外見上全く同じで，外から中は見えないものとする。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。」

＜解答＞

右の図 1 の面積図を用いると理解がらくになる。

①A の壺か B の壺かであるので，それらの確率（事前確率）を　0.5 ずつ割り振る。

②A,B それぞれの壺で，白玉，黒玉が出る確率を求めると図 1 のようになる。

③黒玉が出たので，左図の色の塗ってある部分だけを考えればよい。

以上から，

（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 の壺か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。

さらに，

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。

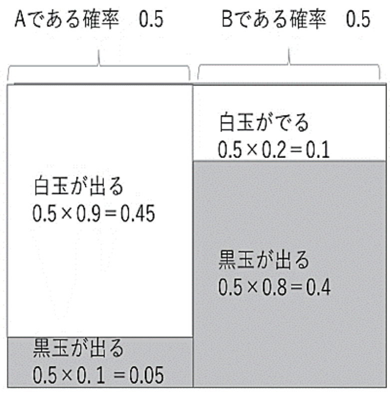


図 1　ベイズ統計学による確率の面積図

3.2.　ベイズ統計の排除と反証主義の哲学者カール・ポパー

上に示したように，最初に，A の壺か B の壺かの確率を等しく 0.5 ずつおくという事前確率の設定を除いて，ベイズ統計学の方法で何ら問題はないように思える。ベイズ統計が批判されていしまったのは事前確率の設定がどちらも同じだと勝手に決めてしまっているという点だけではない。その批判の背後にあるのは当時の哲学的思想である。カール・ポパー（1902～1994）の**反証主義**が絶大なる力を持っていたからである。「（反証主義によれば，）科学は仮説の提案，検証，棄却，より良い仮説の提案，というサイクルを通じて進んでいく。まず科学者は，ある現象についての仮説をたてる。次に，その仮説から導き出される予測を，実際に観測されたデータと突き合わせることで仮説を検証する。もし予測とデータが一致しない場合，科学者は当該仮説を誤りとして棄却し，より良い仮説の探究に向かうことになる。では，予測が成功した場合はどうだろうか。この場合でも，仮説は正しかったと喜んではいけない。というのも，ある仮説の論理的帰結が正しいからといって，その仮説自体が正しいと判断するのは，いわゆる後件肯定の誤謬を犯すことになるからだ。実際のところ我々が言えるのは，当該仮説は今回のテストを生き延びた，ということだけに過ぎない。今回は大丈夫だったが，次のテストではだめかもしれない。よって科学者がすべきことは，仮説からさらなる帰結を引き出し，それを新たなるデータと突き合わせることで再び仮説を検証にかけることである。もちろんこのテストを通ったとしても，仮説の正しさが認められるわけではない。」〔2〕というのが反証主義である。反証されればそれで終わり，反証されなくても正しいとは限らない。反証されなければ，検証を続けるしかない。しかし，その過程においてどのくらい真理に近づいたかを人間は知ることはできない。後件肯定は誤謬であるが，後件否定は論理的に妥当である。反証主義は後件否定の論理である。そこまでしなければ，真理らしきものでさえ得られないというわけである。ベイズ統計の事前確率の設定が等配置されるには根拠がないので，真っ先に批判される。そして，お気づきだと思うが，この反証主義の方法と頻度主義統計学の仮説検定の手法は酷似している。これは，ネイマンとピアソンが反証主義によって自分たちの考案した手法が批判を受けないように，反証主義の方法を真似たということである。話が逸れるが，仮説検定において仮説が棄却されなかった場合，「仮説を採択する」と表現している書籍が少なからず存在している。しかし，これは厳密には誤りである。反証主義にのっとった仮説検定においては，棄却はできても，採択はできないのである。その理由は，ここに述べたことから明らかであろう。

3.3　仮説検定は「確率を用いた背理法」～崩れた背理法の隙間を確率で補う

仮説検定が反証主義をもとにしていることから考えて，論理的にしっかりできているに違いない。しかし，検証して仮説の真偽を判定していかねばならないという事実かして，論理のみで完全ではないということは

明らかである。おそらく，その辺の白黒つけがたい部分に対して，「反証すれば仮説の間違いは明らか」，「反証できなくても仮説が正しいとは言えない」と，どちらにしても否定で統一することで，曖昧さの排除に努めたに違いない。だが，ここに確率的判断が許されるとするともう少し穏やかになるのではないだろうか。また，反証主義の手法は背理法に通じるものがある。先ほどの問題で，これらのことを実験的に試してみる。

＜問題＞(再掲)

「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。2 つの壺は外見上全く同じで，外から中は見えないものとする。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。」

この問題に対する思考過程を論理的表現にするのを妨げているのは，黒玉と白玉の混在であり，たった一つのデータからでは，A，B のどちらの壺であるかを判断するには情報が不足しすぎている。そこで，「壺 A がすべて白玉，壺 B がすべて黒玉とする」という特殊な条件をいったん付けて思考過程を強引に論理的にする，全体の流れは背理法の論法によって表現してみる。

＜背理法的表現①＞（背理法そのもの）

壺 A がすべて白玉，壺 B がすべて黒玉とする

1. 壺 A だと仮定する。（壺 B であることを主張したいため）
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。

次に，厳密な論理としては崩れてしまうが，設定を戻し，それに合わせて文章表現を変化させてみる。

＜背理法的表現②＞（論理的に崩れた背理法）

壺 A が白玉 9 個，黒玉 1 個，壺 B が白玉 2 個，黒玉 8 個にもどすと

1. 壺 A だと仮定する。（壺 B であることを主張したいため）
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**。

となって論理的には，非常に歯切れの悪いことになる。この曖昧な表現に対して，確率を数値で与えることで補えば，かろうじて背理法の論理性を保てると考える。もちろん不完全な論理なので，ある程度の反例が生じることも覚悟しなければならない。しかし，それの出現率も確率という数値で表せていれば，全体としてその推論結果の妥当性を数値で評価できるであろう。曖昧な表現をその起こりやすさと，そうではなことが起こる起こりやすさを数値で比較すれば，論理的に破綻していても，確率を用いた推論としては上手くいっていると考えて良いはずだ。ただ，上の例のように曖昧な箇所が複数になってしまうのはやりにくい。そこで，論理的な強さを回復するために，「確率がある数値以下の起こりにくいことは『起こらない』と決めてしまう。」という論理的な近似を行って背理法の論理力を復活させ，そして，その近似の代償の妥当性を評価するという「仮説検定の考え方」が浮かび上がってくるのである。この少ない確率で起こるかもしれない

ことを「起こらないこと」と判断してしまう基準の確率が『有意水準』にあたる。これについては次節で述べる。

3.4. 有意水準（危険率）の設定とその意味

前節の＜背理法的表現②＞の表現をここでもう一段階変化させる。「確率 1 /10 以下で起こることは，起こりにくい事なので『起こらないこと』，つまり確率 0 としまおう。」という考えのものとに書き換えてみる。なぜ，1/10 かといえば，それが説明に都合がよいからである。また，この数値が普通の仮説検定では，主に 0.05 などで表される「有意水準」といわれる数値である。統計学の歴史上，はじめのうちは，0.05 となる深い数学的根拠はなく，頻度主義統計学のフィッシャー（1890〜1962）が，その程度の誤差は許されるといったからだという説が有力である。あとになって検出力などの研究で妥当性は示されているようであるがここでは触れない。つまり，比較的恣意的に設定できる数値である。そして，有意水準とは，「起こる確率を 0 とみなしてしまう確率の上限」なのである。言い換えるなら「『見て見ぬふりをする』ことが許される確率の上限」なのである。「有意差 0.05 で，仮説が棄却された」など，有意差（有意水準以下で仮説が棄却できるほどの差）があったことを神の承認を得たかの如く結論を誇らしげに語る人がいる場面によく出くわすが，「確率 0.05 以下で起こることは無視したよ」と自分の罪を自慢しているに過ぎない。少し前の統計学の教科書には，**「有意水準を危険率ともいう」**と書かれていた。危険率の方が妥当と思われる。なぜなら，確率 0.05 以下とはいえ，起こるかもしれないことを起こらないことにして，つまり無視して強引に結論を出すからである。無視した 0.05 の確率で起こることが起こっていたら，その結論は間違っているのである。仮説検定は常にこの危険を犯してまで白黒つける，という方法なのである。だからこそ，その結果を安易に用いるべきではない。どうしても白黒つけたいときに，危険を承知で用いるべきものである。その危険性が現実のものとなったときに発生するのが，第 1 種の過誤（正しい仮説を棄却してしまう），第 2 種の過誤（棄却されるべき仮説を残してしまう）なのである。前置きが長くなったが，＜背理法的表現③＞は次の通りである。

＜背理法的表現③＞

<u>壺 A が白玉 9 個，黒玉 1 個，壺 B が白玉 2 個，黒玉 8 個</u>

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。

つまり背理法的論理を保証するために，ある基準以下の確率でしか起こらない起こりにくい事象を「起こりらないこと」決めつけてしまうこと，そのようにして結論を，仮にではあるが，明確にしてしまう方法，それが「仮説検定」なのである。そして，有意水準はその基準の確率を表すものなのである。＜背理法的表現③＞における有意水準は 1/10，つまり 0.1 となる。

3.4. 仮説検定のロジック

仮説検定のロジックとは，上で述べてきたように論理的に崩れた背理法とそれを補う有意水準による補強による論理である。前出の白玉黒玉問題で用いた論理を，一般化して記述すると次のようになる。

仮説［1］が正しいと判断したいとき，

① 仮説［1］に反する仮説［2］(厳密には，仮説［1］の属する集合の補集合に属する仮説）を立て，仮説［2］が正しいと仮定する。

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。次の⑤か⑥が成立している。

⑤ ③で求めた確率が②の有意水準より小さかった。→ 仮説［2］が正しくなかった（仮説［2］は起こらないとみなされたので)。よって，仮定［1］は正しい（だろう）と判断して良い。

⑥ ③で求めた確率が②の有意水準より大きかった。→ 仮説［2］は捨てられない（仮説［2］が起こらないとはみなせないので)。よって，仮説［2］か仮説［1］かの判断は保留する。

このように，仮説検定のロジックを表すとほとんど「仮説検定のやり方」になる。このことから，教科書が「仮説検定のやり方」を説明することが，まったく「仮説検定の考え方」と無関係にはならない。この点は私の意見を少し修正しなければならない点である。しかし，上で述べてきたように，背理法と確率を基に具体例を交えて「仮説検定のやり方」を示した場合と，直接的に「仮説検定のやり方」を示した場合では，学習者のとらえ方は全く違うものと考えられる。私が示したように，既習事項の背理法及び確率を用いて思考過程を説明することに重点を置かなければ「仮説検定の考え方」の単元とは言えないと考える。さらに，教科書では，仮説の立て方や有意水準の扱いが天下り的で本来の意味の理解ができず「考え方」の指導にはならないと思われる。本格的ではないにせよ，上述のような説明をすることで「考え方」の方に重点が移り，なおかつ第 2 学年の数学 B で取り扱われる本格的「仮説検定」の興味･理解を促進させるものと考える。

前出の白玉黒玉問題をこの仮説検定のロジックを用いて，解いてみると次ようになる。

＜問題＞(再々掲)

「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。2 つの壺は外見上全く同じで，外から中は見えないものとする。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。」

＜解答＞

壺 B であることを主張するために，壺 A であることを仮定する。したがって，ひとまず，壺 A のみについて考えればよい。壺 A から白玉が出る確率は 9/10，黒玉が出る確率は，1/10 である。
ところで，1/10 以下で起こる事象は，「起こらないもの」と見なすので図 2 のように，黒玉の部分は考えない。すると壺 A からは白玉しか出ないと考えられる。したがって，これは壺 A から黒玉が出たという事実に反する（壺 A から出た黒玉だとは考えにくい)。

よって，壺 B である（可能性が高い，かつ 1/10 で黒玉が出る可能性を見ないことにしているという危険が伴う)。

図 2 頻度主義統計学の仮説検定による面積図

4.　仮説検定のロジックを生かした授業

4.1.　指導案

以上までの内容をもとに授業案とワークシートを以下の通り作成した。公開研究会当日のものに若干ではあるが，手を加えている。

１年　数学科（数Ｉ）学習指導案

「仮説検定のロジック　～　仮説検定の本当の考え方とは～」

授業者：お茶の水女子大学附属高等学校
教　諭：　三　橋　一　行

１．日　時　2022 年　11 月　19 日（土）　9：50～10：40

２．場　所　お茶の水女子大学附属高等学校　書道室

３．対　象　１年梅組(女子 42 名)

４．単　元　データの分析　（仮説検定の考え方)

５．授業設定の理由

2022 年度から数学Ｉで「仮説検定の考え方」という単元が導入された。「仮説検定」自体は、2 年生の数学Ｂで扱われる。しかも、新課程では、数学Ｂの「確率分布と統計的判断」は（数学Ｂを選択した生徒にとって）必須となる。数学Ｉの段階で仮説検定のための準備にあたる内容を教えずに、仮説検定を知らない生徒にどのように「仮説検定の考え方」を指導導したらよいのか。このことは、高等学校の現場で数学教師が頭を痛めているところである。

教科書の内容や学習指導要領解説に紹介されている例は，良く工夫された内容であると思うが，ひどい言い方をご容赦いただけるなら「仮説検定の劣化版」で，しかも原理や仕組みを教えるのではなく，「使い方」指導になってしまっていると私は感じる。これでは，現在の数学教育の課題を解決する指導とは逆行する状況になりかねない。今回の授業はこの問題に答えるべくして考えられている。今後，この指導における一つの提案となる授業となれば幸いである。新課程の教科書にはない、オリジナルの授業である。次年度以降への本単元に対する課題と効果を見るため，また，数学Ｂでの学習状況への効果を考えて設定した。

６．授業内容についての考察

新課程のカリキュラムでは、「仮説検定の考え方」に配当された時数は 1 時間である。その短時間で仮説検定を考えるのはかなり難しい。新しい教科書においても 1 ページ程度の取り扱いである。主張したい仮説１（対立仮説）を立て、それを否定する仮説２（帰無仮説）を立てたのち、実験などを行って仮説２が確率的に誤っているかどうかを判断する。という手順と簡単な例が１ページほどに詰まって載っている状態である。教科書にもよるが、これでは手順を説明しただけで「仮説検定の考え方」を学んだとは言い難い。また、数学Ｂでの本格的な仮説検定の学習時においても、手順が優先され「なぜ、そうなるのか」というところにはかみ砕いた説明がなされない。この授業は、仮説検定そのものを教えずに仮説検定の本質に迫りながら、その考え方を学ぶことを目指している。仮説検定の本質は、「有意水準（危険率）という判断基準で、帰無仮説が捨てられるかどうかを判断すること」であると私は考えている。その判断の際に「起こりにくいこと」を「起こらない」と見なしてしまうことにより，灰色の部分に白黒どちらかの判断を与えるのである。このとき，わずかな確率ではあるが，得られた判断とは反対の可能性をつぶしてしまう。その功罪についても仮説検定を行う際には忘れてはならないことである。そのことについても本授業の中でふれるつもりである。

７．生徒の実態

課題やグループ活動に積極的に取り組むまじめなクラスである。発表などで間違ってしまっても，周りの生徒が批判せず受け入れてくれるので，発言がしやすいクラスと考えられる。また数学に興味のある生徒が比較的多く、「虹の数学」による光線経路の作図問題では，活発な意見や発表が見られた。全体的に好奇心旺盛で，学習への取り組みが良いクラスである。

８．本時の学習

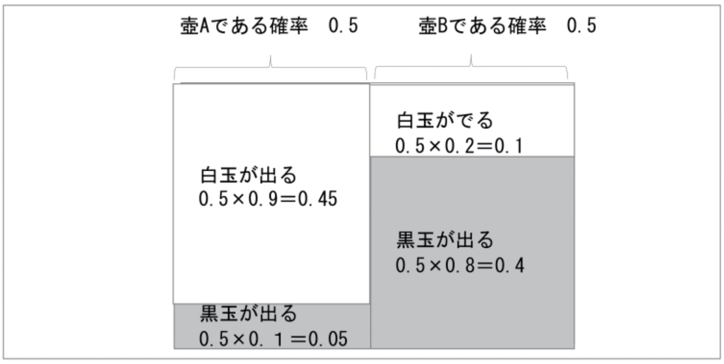
（１）本時の目標

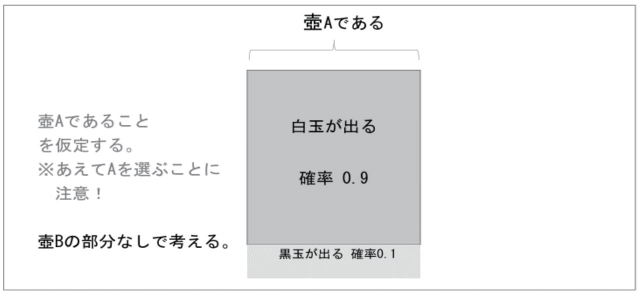
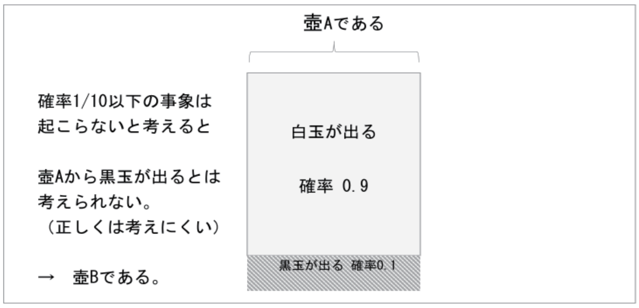
① 仮説検定の考え方を論理的に理解する。

② 有意水準の意味が，過誤の原因であることを知る。

（２）学習活動

学習内容	学習活動・指導過程 主な発問（Ｔ）と予想される生徒の反応 (S)	留意点（☆） 学習支援（○）
導入 （5 分）	**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。** (T)「この判断は正しいか？ (S)「正しいのでは」,「正しくない」「そうとは限らない」など (T)「後件肯定の誤りである。インフルエンザとは限らない。」 論理の学習の復習も兼ねる 夏休みの宿題にワークで「仮説検定の考え」は予習していると思う。その考え方を深めるための授業をする。非常に簡単な例を取り上げてるので，遠回りに感じるかもしれないが，慎重によく考えてほしい。」	（○）パワーポイントのスライドを利用して授業を行う。 （☆）正式には、「統計的仮説検定」という。 （☆○）後件肯定の誤りと論理の復習も兼ねて導入問題を設定している。 （○）データ分析のワークを夏休みの宿題として出している。 （○）これからの流れを軽く説明する。
展開① （20 分）	＜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。」 (T)「これから２通りの方法でこの問題について，説明します。あとで，この説明について話し合いをしてもらいますので，論理的に，数学的にどうかという点についてよく聞いておくこと。また，どちらの考え方が良いか，その理由も含めて答えください。」	（○）問題を紹介する。 （☆）少し考えさせたら，この問題の解答の例として２つ紹介する。 １つはベイズ統計学によるもの。もう１つは，頻度主義統計学による仮説検定によるものである。ここで，統計学史上の頻度主義者とベイジアンとの論争を追体験してもらう。

	（T）「では，1つ目の説明を始めます・・・・考え方Aです。」	（○）説明後の活動を知らせ，注意深く説明を聞くことを伝える。
		（☆）ベイズ統計の考え方でまず説明する。 （○）面積図を用いて，説明する。実はこれが，確率分布の表や面積に当たる部分である。
	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の壺か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	（☆）ベイズ統計は数学液であり，生徒にとってわかりやすいものではないかと考えられる。 　歴史的には，事前確率が 「科学的でない」として弾圧をうけてしまった。
	（T）「では，2つ目の説明を始めます・・・・考え方Bです。」	
	仮に　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。 　背理法の論理をなぞっていけないか・・・・	（☆）頻度主義のネイマン=ピアソンの仮説検定方の考え方に沿って説明する。 　仮説をつくり，背理法など既習事項の論理にそって，思考の流れを説明する。 （☆）有意水準は危険率であり，ある種の危険を伴った判断であることが自然と明確になる。

	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。  ↓  　・・・・と考えれば，判断ができる。」 (S) 説明をきく。 　質問があれば，受け付ける。	(☆) 質問は最低限の解答に止める。生徒の話し合い の話題を奪わないようにする。
展開② (5 分) 発表 (10 分)	(T)「以上，2 つの説明について，それぞれ自分のワークシートをもとに話し合ってみよう。どちらの考え方が良いか。その理由も含めて考えよう。」 (T)「では，答えてもらいましょう。」 　個人または，グループの代表として，意見を発表させる。 (S) 最初の考えが良いか，あとの考えが良いか。それともその他の考えか，それぞれ理由とともに，意見を述べる。	(○) ワークシートを配布する。 (○) どちらが良いか，あるいは，両方良いか，両方ダメか，第 3 の方法はあるかなど，幅広く意見を聞いてみる。 (○) 意見が振るわないときは，指導者がヒントを与える。

まとめ① （7 分）	（T）「前半が，ベイズ統計の考え方，後半が，ネイマン＝ピアソンの仮説検定の考え方で，・・・」それぞれの，利点，欠点を簡単な歴史の流れとともに伝える。その際，有意水準（＝危険率）の意味も伝える。 ネイマン＝ピアソンの仮説検定の考え方がややこしくなっている背景には，**この授業の冒頭で行った後件肯定の誤りを防ぐためであったことも伝える。** その他，まとめの解説を行う。 ●ここまでの復習と思考力を使うためワークシートの残りの問い（問２と問 3）を考えさせる。時間があれば，発表させる。	（☆）それぞれの考え方には，一長一短がある。歴史的背景として，ベイズが先に使用され，カール・ポパーの科学的認識論の話などから，仮説検定も科学的にしようということから，仮説検定の方法がつくられた，しかし，コンピュータの発展から再びベイズの時代がきている。
まとめ② （3 分）	（T）「（プリントで）仮説検定の問題を宿題とします。今日の授業のことを思い出しながら，問題を解くこと。また，解いた後に，夏休みに仮説検定の問題を解いたときと，今回の問題を解いたときの感想も書くこと。」	（○）夏休みの問題と同じでも良い。感想を見て，仮説検定の考えが深まったか確認する。

4.2.　**指導案の説明**

ワークシート（仮説検定の考え方）

問１．　［考え方 A］と［考え方 B］を聞いて，以下の問いに答えなさい。

①　どっちらの考え方が，良いと思いますか？理由も書きましょう。

（スペース省略）

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。

（スペース省略）

「仮説検定の考え方」（教科書 P. 203）

主張 → 仮説と置き換える。

仮説［１］が正しいと判断できるかを考えるとき，

①　仮説［２］（仮説［Ｉ］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］のもどで，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。

そこで問題です。

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。

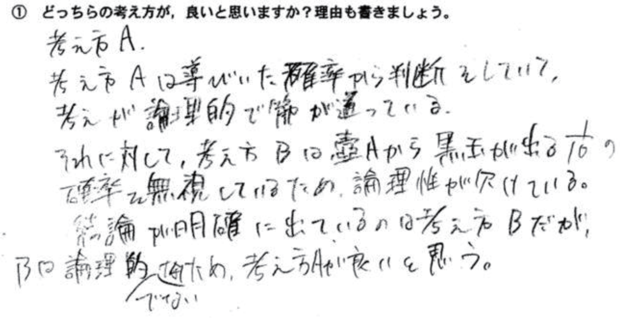
（スペース省略）

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。

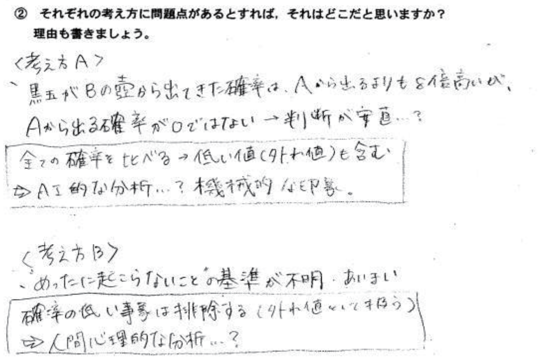
（スペース省略）

4.3.　授業の結果から

ワークシートの問 1①，②から主だった記述を以下に載せる。授業中の考え方 A か考え方 B かの議論は大変活発に行われた。この学年の生徒の性格による要因もあると思われるが，合わせて統計学史上の大問題でもあるので，その要点がつかめれば，議論は絶えないはずである。また，各グループの発表も予想外に長くかなり多くの意見が出たことがうかがわれるものであった。これらのことから，2 つの考え方の説明は上手くいったといえるだろう。このクラスにおける頻度主義対ベイジアンの論争は，およそ，1 対 7 でベイジアンの勝利となった。しかし，授業者の反省もあって，この論考の冒頭にも述べたように頻度主義の仮説検定に疑念をもっているので，ベイジアンの考え方に少々力が入っていたようだ（授業後の反省アンケートに記載されていた）。また，ベイジアンの説明の方がすっきりしており，これまでの確率の考え方がストレートに使えるので生徒たちには馴染みやすかったと思われる。対しての頻度主義仮説検定の考えは，長くてひねりがあるので，わかりにくい印象がありそれが妨げになったと思われる。

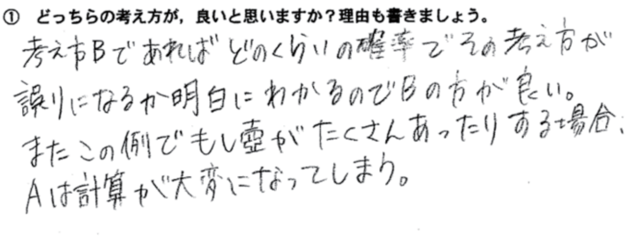

① どっちらの考え方が，良いと思いますか？理由も書きましょう。

考え方A.
考え方Aは導びいた確率から判断をしていて，
考えが論理的で筋が通っている。
それに対して，考え方Bは壺Aから黒玉が出る 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は論理的でないため，考え方Aが良いと思う。

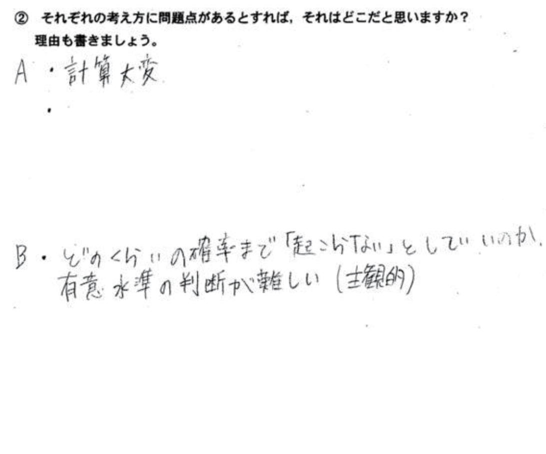

② それぞれの考え方に問題点があるとすれば，それはどこだと思いますか？
理由も書きましょう。

〈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→判断が安直…？
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・あいまい
確率の低い事象は排除する（外れ値として扱う）
⇒人間心理的な分析…？


① どっちらの考え方が，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。

A ・計算大変
・

B ・どのくらいの確率まで「起こらない」としていいのか，
有意水準の判断が難しい（主観的）

① どっちらの考え方が，良いと思いますか？理由も書きましょう。

B. Bは $\frac{1}{10}$ で間違えるが、

Aは $\frac{1}{9}$ で間違える。

Bの方がよいと思う

方 Aの方がよい。

(Bはつぼ Bの都合がよいように考えている？

Aの方が数値に基づいて。)

① どっちらの考え方が，良いと思いますか？　理由も書きましょう。

考え方α

考え方βでは $\frac{1}{10}$ 以下の確率を無視してしまっているので、誤差がでやすい。対し、考え方αはどちらの方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表せるため、より正確だと考えるから。ただし、考え方αでは確率の大小でしか結論づけられないため、どちらが出るかという結論が確定しているβの方が、主張をするのにはむいていると思った。

次に，問 2 と問 3 の解答例も挙げておく。問 2 は，結論がはっきりいえているわけではない仮説検定の危険性がわかっているかどうかを試す問いである。教科書では，はっきりと結論を述べられると誤解されるような記述があるので，それがどこかを考えさせることによって仮説検定の危険性を知ったうえで使用するこという態度を身に着けさせようとする問題である。2 つほど例を載せるが，殆どの生徒（約 93%）が「言い切れない」とか「結論に危険性がある」ことが指摘出来ていた。問 3 は，有意水準設定のタイミングを問う問題である。あらかじめ，決めておくのが公平な態度である。分析後に決めるとすると有意水準は恣意的であるので，都合に合わせた結果をコントロールして導き出すことも可能になるためよろしくない。これも殆どの生徒（約 95%）が理解できていた。なお，実社会の統計学では，有意水準で棄却域を考えるのではなく，p 値という数値を計算しそれが，小さいほどその仮説は起こりえない，つまり棄却すると判断する。0.05 以下かどうかと考えずに，「小さいほど良い」という考えになり，“p-hacking”を誘発したのではないかと考えられる。

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「正しくなかった」と言い切ることはできない。

仮定をしてからだと良いけれど、急にはできないよ。

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 絶対にだめだと言い切ることができない。

((仮説検定はとてもリスキー))

めをつぶった確率

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から。

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ってしまうという危険性があるから。

4.4. 授業の反省会から

今回の授業の反省としては，ベイジアンの方に肩入れしすぎている印象をもった。話し合いや発表などせっかく活発に行えたので，ICT を活用して全員によりわかりやすく共有すると良かった。なかなかチャレンジングな授業をしている。などのご意見やご感想をいただいた。先ほど述べた通り，頻度主義統計学の仮説検定に疑念を抱いているので，無意識のうちに肩入れしてしまったのだと思う。次回からは公平になるよう努力してみたい。ICT を利用して，発表を有効に行うということは非常に同感する。しかし，その設備がないのとあったとしてもそれを活用する技術が私にはないので今後の課題としてさらに勉強していきたいと考えている。教科書や指導要領の内容にいわばケチをつけて自分なりものを考えているので，チャレンジングな内容であることは確かである。しかし，この論考のはじめの方で述べた通り，正解の動きや仮説検定というものをよく知らずに授業内容が決められている印象が強い。統計学になじみの薄かった日本人が，統計学に力を入れている外国に抜かれた。だから，日本も統計学をやろうというのは良いとして，統計学が最強だとか，神様のごとく信奉するのはいかがなものか。統計学はなけなしの非常手段なのである。そして，コンピュータでビッグデータが簡単に扱えるようになった現在では，標本調査による統計的判断は必要ない。全数調査を行えばよいのである。そうすれば，確率の出番はなくなる。そういった統計学はもはや数学の範疇を出て情報科の学習になるのではないだろうか。そのような状況が教科書や学習指導要領に生かされていないのは非常に残念なことである。したがって今回のようなチャレンジングな授業を行う必要があるのである。

また，公開授業後の反省会では各学校で指導に困っているというお話がでた。いくつか主だったものを，質問に私が答える形式であげておく。なお，実際に出た質問をまとめたり，答えにその後の知見を加えるなど加工して整理していることをご容赦いただきたい。

［質問］ 仮説検定を行うために別実験を行う（コイン投げやサイコロふり）ことについてどう思うか。

（私） 良くないと考えている。実験（調査）を実験で評価するなんてことはやってはいけないと思う。しっかり理論値で評価したい。そうでなければもはや数学ではない。ここは無理せず，数学 B の確率変数や確率分布の授業が終わるまで待った方は良いだろう。数学 I では，「考え方」を簡単な例で説明することに重点を置いた方が良いと考える。

［質問］ 2 つの仮説が論理的に排反でないことがあるが，大丈夫なのか。

（私） 数理統計学上は補集合の関係（排反）である，ところが実際は曖昧になっている。$\mu A=\mu B$ 1 点での（単純）仮説と $\mu A \neq \mu B$ その他の（複合）仮説 とで考えるならば背反であるが，帰無仮説 $\mu A=\mu B$ と対立仮説 $\mu A<\mu B$ の設定は不自然である。論理的には，$\mu A>\mu B$ の可能性が除かれている。ここは起こらないと勝手に踏んでいるもしくは，そこには，興味がないから調べようとしないだけである。こういう使用の仕方が統計学の信用を失うきっかけになる。分析者が都合よく作り変えている可能性がある。もちろん $\mu A>\mu B$ を排しても良いだけの理由があれば良いということになるが。

なお，仮説が立てられるのは，パラメータ（母数）のみについてのみである。これは，データを作り出す確率変数の背後の密度関数が持っているパラメータ（母数）でなければ，数学的に無意味だからである。

［質問］ 問題となる確率のみでなく，なぜそれ以上（それ以下）の確率をもとめるのか。

（私） 大学の測度論で学んだ通り，確率は面積である。問題にある，ある一点の事象が発生する確率は，

面積を持たない線分なので測度 0 である。だから横幅を持たせて面積にしないと確率にならない。現実に数多くある事象の中で，そのたった一つの事象が起こる確率は殆ど 0 である。例えば，40 人のクラスの生徒をランダムに一列に並べた時，出席番号順になっているということは絶対起こらないわけではないが，まず期待できない。事実上確率 0 となる。逆に確率が 0 であってもそれは全く起こらないということにはならない。これらのことは大変難しいので，無理をしないで，数学 B の確率分布の話が出てきてから，考えたほうが良いだろう。現在は，ヒストグラムを作成するなど横幅を持ったグラフで考える方が良い。（質問されたが学校ではすでにヒストグラムを使用していた。すばらしい。）

［質問］　**有意水準の設定について 0.1 とか 0.05 とかどう決めるのか？**

（私）　有意水準を今回は「見ないことにする確率」と決めた。これを大きく設定すれば，どんどん棄却できるので，歯切れよくなるが同時に棄却してはいけない仮説も棄却する可能性が増加する。逆に小さく設定すれば，歯切れが悪くなり棄却すべき仮説の棄却できなくなる。問題にあわせて，大きめにするか小さめにするかを考えると良いと思う。わからなければ，標準の 0.05 にすればよいと思う。

［質問］　**データサイエンス，統計学の結果の信頼性は？**

（私）　不確定な要素，確率を用いた要素を含めて分析する限りどんな手法であれ，リスクを背負った判断になる。そのことを忘れないようにすることと，リスクを背負ってまで，どちらが妥当か判断すべき事態になっているかどうかということを分析前に考えたほうが良い。リスクをなくすなら，全数調査をするか，もしくは，ややこしい手法を用いずに円グラフなどで各項目の割合を表示する提度にとどめておいた方が良いだろう。妥当性とリスクが両方確認出来て公平な感じがする。

5. 今後

今回の授業では，わかりやすさを優先させるために，有意水準を 1/10 に設定した。しかし，これだと確率としては大きいので，生徒によってはそんなに目をつぶるのなら考え方 B は正しくないと思う生徒もいた。玉の数を思い切って 100 個ぐらいにするパターンも実践してみたいと考える。今回，10 個にしたのには，もう一つ理由があって，実際に壷と玉を用意して代表生徒に実演してもらって当たるかどうか実験的要素を高めた授業にするという原案もあったからである。しかし，これだとこちらの都合の良い結果になるとは言えず時間を浪費する恐れがあったので，今回はやらなかった。今回の授業反省会で指摘されたように，どちらの考え方にも公平であるという態度を強めておくことも課題である。可能であれば，ICT を活用してより発表や意見交換がスムーズに効果的に行えるようにするのも次回以降の課題である。私としては，当面，数学I「仮説検定の考え方」をこのスタンスで授業できればと考えている。仮説を立てるときの論理性，背反である場合とそうでない場合など帰無仮説と対立仮説の論理的関係についてももっと研究する必要がある。大学の統計学の授業では何気なくスルーしていたが，改めてよく考えるとおかしい。また，有意水準 0.05 の妥当性を検出力という観点から高校生にわかるようにしていきたい。頻度主義統計学が衰退した後は，ベイズ統計学が主流になってくると思われる。ベイズ統計学についてももっと研究を進めたい。

以前，院生時代，指導教官から「統計学は人間と自然（神）とのゲームなのだ。」という言葉をいただいた。失礼ながら，ええかっこしいでキザなセリフを述べられたのだと思い続けていたら，50 年近く前に発行された数学書に「ゲーム理論で統計学者の最適戦略を探る」という研究が載っていることを，つい最近発見した。大変興味をそそられる。現在の統計学者の手法は果たして，最適解なのであろうか。とにかく，統計学，仮

説検定についての興味は尽きない。こちらも可能な限り研究してみたいと思っている。

そして，統計学が高等学校の数学に中心的に扱われるようになるにあたっていろいろと問題点が出てくると思う。それらの解決策を考えていきたい。研究授業などを行ったり，外部の研究会に参加するなどして，教育現場での問題点，指導上の悩みなども察知し高校数学で対処できるようにしていければと考えている。

それにしても，アメリカがやめると宣言している仮説検定を，今後日本はかなりの時間を割いて授業していくわけであるが，本当に大丈夫なのであろうか。しかも，導入を急ぐあまり不完全で誤解された内容を身に着けてしまいかねない状態になっているというのに。犠牲者が少ない事を祈るのみである。

私としては，自分が良かれと思う授業を開発し行っていくのみである。今回提案した授業の方向にすすめば，世界が頻度主義仮説検定を捨てたとしても，ベイジアンの基本の考え方に触れているので，そちらの方に進むと考えれば他の授業を受けているよりも，わずかではあってもベイズ統計学を受け入れやすいだろう。そこに，わずかながらも希望の光を見つつ，本稿を閉じることとする。

参考文献

［1］豊田　秀樹（2020）『瀕死の統計学を救え！
-有意性検定から「仮説が正しい確率へ』-　朝倉書店「はじめに」の p. ⅰ ～ ⅱ

［2］大塚　淳　（2020）『統計学を哲学する』　名古屋大学出版会　p .97～p.98

［3］文部科学省（2018）『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

［4］数研出版　（2022）『高等学校数学Ⅰ』　教科書