

統計と確率の関連付け ～推測ゲームを事例として～

おおたに ひろき
大谷 洋貴

1. はじめに

新しい中学校学習指導要領が全面実施となりました。充実した統計分野の指導をどうしようかと今から思案されている先生方も少なくないのではないのでしょうか。過去の『チャート.Info』を参照すると、高校数学Ⅰからの移行内容である四分位範囲と箱ひげ図 (No. 14, 18), テクノロジーとPPDACサイクル (No. 16), データの読解 (No. 19) がすでに取り上げられており、教材開発や統計指導の視点が提供されています。

本稿では、統計と確率を関連付けた指導について、それを可能にする教材の1つである推測ゲーム (Brousseau et al., 2001) を紹介しながら考察します。統計と確率は同じD領域に属しながらも、バラバラに取り扱われてしまっているのではないのでしょうか。両者を適切に関連付けられることは、高校数学Ⅰに新たに加わった仮説検定の考え方にもつながるでしょう。

2. 統計と確率の関連付け

周知のように、新学習指導要領では統計的探究が強調されています。現実問題の統計的解決を指向する統計的探究の重要性は明らかです。しかし, Manor Braham & Ben-Zvi (2017) は, そのような統計的探究が, 標本と母集団の関係を理解する上で必要である確率的な考察を欠いていることを指摘します。たとえ統計的探究を十分に強調しても, それだけで確率との関連付けがなされるわけではないのです。

Manor Braham & Ben-Zvi (2017) は, 統計的探究と確率的考察を組み合わせ、図1に示される統合的モデル化アプローチを提案しています。図1左側は統計的探究です。学習者は関心のある現実の問いを調査するためにデータを収集して分析し, 母集団についての結論を導きます。この結論の妥当性をチェックするのが図1右側のサイクルです。得られた結論を仮説とみなし, それを仮定するとどのような標本が得られるかをシミュレートして分析することで, 標本と母集団の関係についての情報を得ることを目指します。そこでは例えば, 標本抽出の方法の妥当性, 標本のサイズや代表性, 標本の変動性や標本分布についての考察がなされ得るでしょう。そして, この考察結果をもとに再び統計的探究がなされます。確率的考察は, 統計的探究でより良い結論を得るために不可欠なのです (e.g., 大谷, 2019)。

3. 教材としての推測ゲーム

Brousseau et al. (2001) の推測ゲームは, 小学生の確率の認識を調査するために開発されたものですが, 統計と確率の関連付けの可能な教材として注目されています (大滝ら, 2019; 大谷, 2019)。推測ゲームの最初の問題状況は次の囲みの通りです。不透明なボトルとビー玉を利用して, これと同様の状況をつくるのが可能です (Nilsson, 2020)。

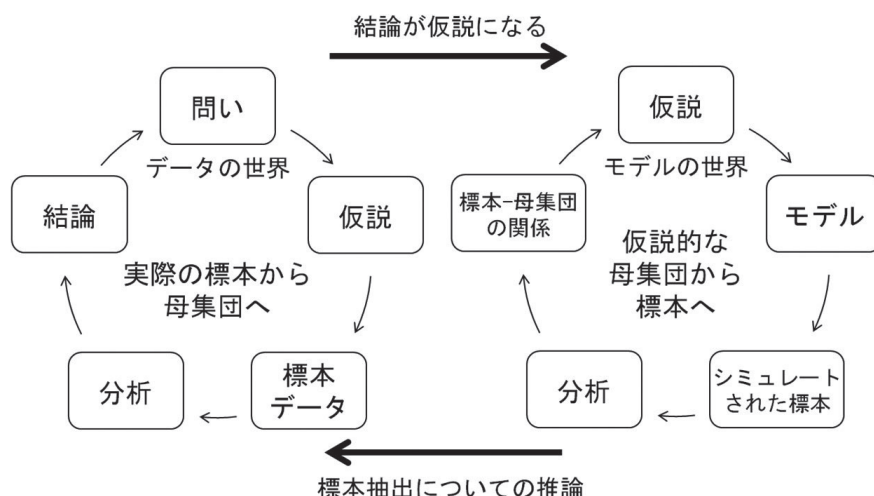


図 1. 統合的モデル化アプローチの枠組み (Manor Braham & Ben-Zvi, 2017, p.120)

教師は A, B, C のラベルのついた大きくて不透明な 3 つの布製の袋と、多数の白と黒の玉が入っている箱を用意した。袋を裏返して空であることを見せたのち、玉の入った箱に手を入れ、決して見ることなく 5 つの玉を掴む。玉の色を見ることなく手を袋 A に入れ、手を開く。袋には 5 つの玉が入っているが、白と黒の玉がいくつつつ入っているかはわからない。他の 2 つの袋にも同じことを行う。袋の中を見ることはできない。誰もが納得のいくように、袋に入っている玉の色の組み合わせを推測せよ。その際、袋から玉を一度に 1 つだけ取り出して色を確認することができるが、取り出した玉はすぐに袋の中に戻さなくてはならない。

推測ゲームの問題状況はとてもシンプルですが、統計的推測の重要な要素が豊富に含まれています。Brousseau et al. (2001) では抽出回数に制限を設けた問題状況も扱われていますし、Nilsson (2020) は玉を三色にして仮説検定を扱っています。玉の個数を変えることもできるでしょう。ぜひ実際にこの問題に取り組んでいただければと思います。シンプルな問題状況であるにもかかわらず、色の組み合わせを推測しようとする一筋縄ではいかず、誰もが納得のいく推測をするには様々な統計的知識を必要とすることを実感できるでしょう。大滝ら(2019)では、区間推定や仮説検定、さらにはベイズ推定など、起こり得る探究が検討されています。

4. 多数回の抽出による推測

中学生が推測ゲームに取り組んだときに、どのような統計的探究と確率的考察が起こり得るのかを具体的に考えてみましょう。素朴な推測方法は、袋から玉を一度に 1 つだけ取り出して色を確認すること、すなわち抽出を 5 回繰り返し、それを母数の推定値とみなすことです。例えば、黒黒白白黒の順で抽出したとき、袋の中の玉は黒 3 白 2 の組み合わせだ、と結論付けます。この方法による推測が説得的でないことは明らかです。なぜでしょうか。統計教育研究の術語を用いれば、抽出された標本 ($n=5$) の変動性 sampling variability が考慮されておらず、その標本が母集団と類似した特徴を持つか (代表性 sampling representativeness) が疑わしいから、

と説明できます。 $n=5$ の標本を何度も抽出すれば、黒4白1や黒1白4のパターンも確認できるでしょう。この素朴な方法では、誰もが納得のいく、という条件を満たさないことがわかります。

では、どうすればより説得的な結論を得られるのでしょうか。そのための方法の1つは標本サイズを大きくすることです。100回抽出して、標本($n=100$)内の黒玉と白玉の比率をもとに推測することが考えられます。母比率の推定に標本サイズが大きいときの標本比率を利用することは、大数の法則を背景にしています。大数の法則は高校数学Bの学習内容です。しかし、標本サイズを大きくすることで標本変動性を制御できること、標本サイズが小さければ偏った標本が得られやすくなることは、中学校3年生で学習され得ます。例えば、 $n=5$ の標本20個と $n=50$ の標本20個から、それぞれ標本比率を求めて図示することで、どちらの場合により説得的な結論を導くことができるかを実感することができるよう(図2参照)。この対比は標本分布の学習ともいえます。標本サイズを大きくするほど統計量の変動する範囲が狭まり統計量の平均値付近に密集してくることがわかります。

さらに言えば、中学校1年生の統計的確率でも、標本や標本サイズという言葉は出てきませんが、標本サイズの影響は学習され得ます。ペットボトルキャップや画鋲をなぜ多数回投げなければならないのか。それはサイズの小さい標本は偏る可能性があるためです。多数回の実験活動では相対度数の収束する値に関心が向けられますが、統計と確率の関連付けの観点からは標本サイズの影響にも目を向けたいところです。なお、ペットボトルキャップや画鋲だと無限回試行ができないことから真の確率を知ることができませんが、推測ゲームでは袋を開けて真の確率を確認することができます。この点は推測ゲームの良さの1つといえるでしょう。

5. 仮説の検証

推測の結果を母集団についての仮説として解釈すれば、それを検証することでより納得のいくものへと洗練することができます。検証方法は、シミュレーションを用いる場合と、確率計算を用いる場合が考えられます。いま、玉の色の組み合わせを白2黒3だと推測しているとしましょう。この組み合わせの袋やボトルを新たに作成し、それをモデルとして実験することで、新たに得られた結果と不透明袋を用いた先の実験結果とを比較することができます。物理的なシミュレーションはスプレッドシートなどのテクノロジーで代替可能です。当初の推測結果とモデルによる実験結果が似通っていればより納得のいく推測と言えますし、明らかに異なるのであれば仮説の尤もらしさは低いと判断できます。

仮説から期待される確率分布を計算する場合も同様に、それと不透明袋を用いた先の実験結果とを比較することによって、仮説の尤もらしさを判断できます。樹形図を使えば(やや大変ですが)中学校2年生でも確率分布の構成に取り組めるでしょう。例えば袋の中の玉を、白₁、白₂、黒₁、黒₂、黒₃として、復元抽出を3回繰り返す状況を考えます。起こり得る場合は同様に確からしく、全部で $5^3=125$ 通りあり、3回とも白玉は8通り、白2黒1は36通り、白1黒2は54通り、黒3は27通りになります。この期待される確率分布と、不透明ボトルで $n=3$ の標本抽出を多数回繰り返した結果とが一致しているかどうかを検討することで、白2黒3という仮説の尤もらしさを判断できます。図3は両者が明らかにずれている状態の分布です。

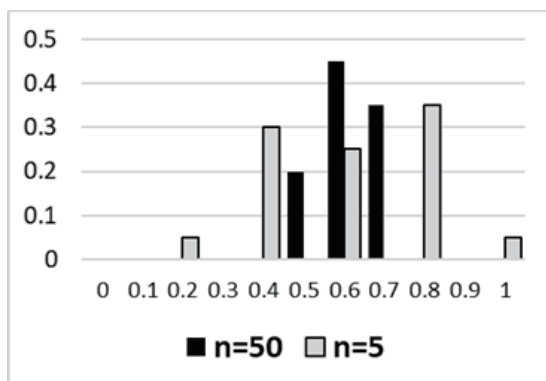


図 2. $n=50$ と $n=5$ の標本各 20 個の対比例
(標本分布の平均はどちらも約 0.6)

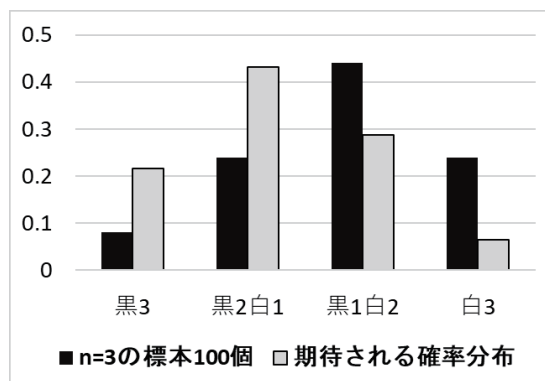


図 3. $n=3$ の標本 100 個 (左) と期待される確率分布 (右) との対比例

実験回数が少なければ、競合する複数の仮説が提起されることがあり得ます。母集団についての仮説を検証することは、どの仮説がより妥当なのかを明らかにするために必要です。これは高校数学 I で学習する仮説検定の考え方に直結するものです。その素地となる経験は中学生でも可能ですし、シミュレーションを用いるなら中学校 1 年生でも十分取り組めるでしょう。

6. おわりに

推測ゲームは統計と確率の関連付けの可能な教材であり、学年を超えて取り扱うことができます。中学校 1 年生では多数回抽出と統計的確率を使うことで、2 年生では仮説から期待される確率分布を求めることによって、3 年生では標本サイズの影響を踏まえることで、それぞれ納得のいく推測をすることができるでしょう。確率的考察は高校数学 I で扱われる仮説検定の考え方に直結するため、その素地を養う意味でも統計と確率の関連付けは重要です。まずはこの推測ゲームに実際にチャレンジして、統計と確率の関わりを実感してみてください。

文献

- Brousseau, G., Brousseau, N., & Warfield, V. (2002). An experiment on the teaching of statistics and probability. *Journal of Mathematical Behavior*, 20(3), 363–411.
- Manor Braham, H., & Ben-Zvi, D. (2017). Students' emergent articulations of statistical models and modeling in making informal statistical inferences. *Statistics Education Research Journal*, 16(2), 116–143.
- Nilsson, P. (2020). Students' informal hypothesis testing in a probability context with concrete random generators. *Statistics Education Research Journal*, 19(3), 53–73.
- 大滝孝治, 袴田綾斗, 大谷洋貴, 福田博人 (2019). ブルソーの推測ゲーム：統計的探究のための教材研究．全国数学教育学会第 51 回研究発表会発表資料．
- 大谷洋貴 (2019). 探索的データ解析アプローチから確率に基づくアプローチへの移行に関する一考察．日本科学教育学会第 43 回年会論文集, 129–132.

(日本女子大学 助教)