Google Colaboratory を利用した統計学習 –仮説検定を例として–

神戸大学附属中等教育学校
数学科 中田雅之

1 前提知識

- 連続型確率分布の基本事項については一通りの学習を済ませており，正規分布の学習を終えている．
- χ^2 分布や自由度の概念について，すでに触れている．
- 区間推定については，正規分布を仮定した母平均の推定や，二項分布を利用した母比率の推定をすでに学んでいる．
- 本校では Google Workspace を利用しており，Google Spreadsheet を利用した表計算や，Google Colaboratory による統一した計算機環境を用いて情報や統計の授業を行っている[*1]．

2 仮説検定とその考え方について

例題 1

コインを 20 枚投げたところ，そのうち 15 回で表が出た．この結果から，コインは（例えば歪みや細工により）裏よりも表が出やすいと判断してよいだろうか．

この例題に対する解答を与える過程を通して，仮説検定の考え方についての解説を行う．

- 帰無仮説と対立仮説の立て方については，その背後に「背理法」の考え方があることに触れておくと，対立仮説に対する帰無仮説の意味づけが理解しやすい．
- 実験結果等の観測事実を，帰無仮説の下では「十分稀な事象である」と認定するための閾値を設定するために，有意水準を定める．
- 両側検定と片側検定の使い分けについては，最初は難しく感じるかもしれない．本例題でもそうであるように，基本的には差を検出するために使うのは両側検定であるということ，操作による効果が意図したように現れるか否かを検出するために使うのが片側検定であるということを，それぞれの例を交えて説明するとよい．ただし，片側検定を不用意に用いることは有意差を検出し易くする危険性を孕むことについて，十分注意しておくべきである．医療統計等でも，両側検定と片側検定で異なる有意水準を設けていることなどに言及するとよい．

例題 1-bis

例題 1 とは別のコインを 20 枚投げたところ，そのうち 12 回で表が出た．この結果から，コインは（例えば歪みや細工により）裏よりも表が出やすいと判断してよいだろうか．有意水準 5% で両側検定せよ．

例題 1 と類似した点が多く，この例題については帰無仮説 H_0 を棄却し損ねていることが強調できれば十分である．

この例題を通じて，帰無仮説を棄却することの意味，採択することの意味について触れられると，実際の調査における仮説検定の考え方への理解が深まると思われる．特に，帰無仮説を採択することは，「観測された標本からは帰無仮説を棄却するほどの十分な証拠が得られなかった」という程度の消極的な意味合いしかないことについては押さえておきたい．

[*1] ただし，2021 年 9 月の Google による u18 利用規約の改定により，3 か月ほど Google Colaboratory を利用できない時期があり，その間は本校の所有する情報教育用端末を利用し，ローカルにインストールされた Jupyter Notebook を用いた代替措置をとった．このあたりの一連の騒動については，`https://github.com/googlecolab/colabtools/issues/2264` を参照されたい．

例題 1-tris

表が出やすくなるように，コインの片面に細工を施した．このコインを 20 枚投げたところ，そのうち 14 回で表が出た．この結果から，このコインは裏よりも表が出やすくなったと判断してよいだろうか．

本問は，片側検定を利用する例である．この例を通して，片側検定を行うべき例について，理解が深まるであろう．

3　t 分布を用いた検定

正規分布に従う母集団から得られた標本を用いて，その母集団の情報を推測する場合，ほとんどの場合その母平均 μ，母分散 σ^2 の双方が未知である．この状況においては，母分散 σ^2 の代わりに（不偏）標本分散 s^2 を代用して母平均の推測を行うのであるが，このとき，標本の大きさ n から計算される自由度 $n-1$ によってその標本統計量 $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ の扱いが異なる．

生徒はこの統計量の従う分布である t 分布には馴染みが薄い．他方，この t 分布を用いた推定や検定は多くの研究で利用されており，生徒が探究を行う上でも利用の機会が多いと思われる．

また，χ^2 分布について学習済みである場合は，統計量 T と正規分布，χ^2 分布との関係についても触れておくと，それぞれの分布の関係や自由度についての理解が深まると思われる．

問題 2

標準正規分布 $N(0,1)$ の確率密度関数のグラフと，次の t 分布の確率密度関数のグラフを，区間 $[-3,3]$ においてそれぞれ重ねて表示せよ．また，自由度 n を変化させることで，$N(0,1)$ と t_n の確率密度関数のグラフの間に見られる関係について考察せよ．

本問では，正規分布と t 分布の差異を視覚的に捉え，自由度による t 分布の変化の様子を理解することを目指す．

Python を用いたグラフの描画については既に学習済みであるが，正規分布や t 分布を扱うために，ワークシートでは `scipy.stats` を利用しており，やや丁寧にヒントを与えつつ設問を設定している．

問題 3

ある機械で生産された 10 個の製品の重量を測定したところ，次のような結果を得た．母平均は 100g と考えてよいだろうか [a]．

101.1, 103.2, 102.1, 99.2, 100.5, 101.3, 99.7, 100.5, 98.9, 101.4 [g]

[a] 東京大学教養学部統計学教室編，『統計学入門』，東京大学出版会，1991 より．

この問題を通して，t 検定の流れを一通り体験することができる．

まずは，この問題に解答するために，適切に帰無仮説と対立仮説を設定することを求めている．

また，統計計算の方法を 2 種類学べるように，ワークシートを構成している．一つは，計算のためのコードを 1 ステップずつ組み立てるものである．Python による統計学習の利点である，計算過程を共有できることを生かすためには，この設問を通じて統計量を計算するコードを自分で記述し，それが正しく計算できていることを確認する作業を経験するべきである．

検定の結果を，不等式の真偽から Bool 値によって自動的に判定できることも，python を利用することで簡単に実装できる点として，ワークシートの活動内容に含めている．（情報科との兼ね合いや）学習の度合や必要に応じて，活動に取り入れるとよいと思われる．

もう一つは，`ttest` 関数により p 値を計算する方法である．実用上は，およそこの関数を用いて p 値を計算すると思う．

また，本問では，既に学習している区間推定を利用した推測の結果と比較することが求められている．実際の探究においても，計算によって得られた p 値のもつ意味や得られた標本がどの程度帰無仮説を棄却する能力があるのかについて考えておくことは有益である．仮説検定の結果をどのように解釈しまとめるべきかについては非常に繊細な内容を含む[*2] ため，このワークシートの学習にそのすべてを反映させることは不可能であるが，この活動を端緒にして，徐々に学びを深めていくとよいだろう．

[*2] アレックス・ラインハート，西原史暁，『ダメな統計学』，勁草書房，2017.