大阪市立大学法学部「法学政治学計量分析 第 6 講: 仮説検定」の質疑応答

宋財泫 *

2017 年 5 月 22 日

伝統的な統計学における仮説検証の手続きはかなりややこしいです。統計学を専門としている[1)]私にとっても言葉にするのはかなり難しいです。モヤモヤするのは当たり前です。ちょっとした頭の体操みたいなもんです。実際に多くの作業はパソコンがやってくれるので、だいたいの感覚が分かればいいレベルです。

以下は、授業後アンケートの質問に対する回答です。

1. おすすめの本を教えてください。
   - 最も分かりやすい本といえば個人的には本日の講義で紹介した本を含め、以下のものをオススメします!
     (a) 小島寛之. 2006. 『完全独習 統計学入門』ダイヤモンド社
     (b) D. ロウントリー著. 加納悟訳. 2001.『新・涙なしの統計学』新世社
     (c) 鳥居泰彦. 1994.『はじめての統計学』日本経済新聞社
2. p 値が分かりづらかった。(4 名)
   - p 値 (有意確率) とは「帰無仮説が正しいと仮定した場合、今回以上の極端的な結果が得られる確率」ですね。最も代表的な例であるコイン投げを考えてみます。コイン投げを 100 回繰り返し、表が 55 回出たとします。「このコインはイカサマコインだ!!!」という仮説を検証するなら、帰無仮説は「この仮説は公平なコインだ」になります。つまり、表が出る確率と裏が出る確率の

* 神戸大学大学院法学研究科 博士課程後期課程; 日本学術振興会特別研究員 (DC2)
E-mail: tintstyle@gmail.com; Homepage: http://www.jaysong.net

1) 実は政治学が専門ですが、最近、アイデンティティー・クライシスです...

差が 0 ということです。この場合、p 値は公平なコインを 100 回投げた場合、表あるいは裏が 55 回以上出る確率です。

- なぜ「裏」も含まれているのか?は簡単です。公平なコインという帰無仮説を棄却するのは「表がよく出る」という方向だけでなく、逆に「裏もよく出る」という方向もあり得るからです。
- とにかく、この p 値が 0.184 (=18.4%) だとすれば、「このコインが公平なコインで、表 (あるいは裏) が 55 回以上出る確率は 18.4% です」ということです。一般的に有意水準が 5% であり、この値より有意確率の方が大きいので、「このコインはイカサマコインとは言えないよね」となります。
- もし、1000 回投げて 550 回表が出たならどうでしょう。先ほどの例と同様に表が出る確率は 55% ですね。しかし、この場合の p 値は 0.0009 (=0.09%) になり、「このコインはイカサマだ!」と主張できますね。このように試行回数 (データならサンプルサイズ) が増えると、p 値は小さくなりやすいです。

3. 有意水準 (α) が分かりづらかった。(2 名)
   - 有意水準は簡単に言えば、「有意確率が何 % 未満なら帰無仮説を棄却するか」に関する基準です。つまり、何かの分析で $p = 0.079$ で、有意水準が $\alpha = 0.05$ なら帰無仮説の棄却ができず、対立仮説 (検証したい仮説) が主張できません。しかし、有意水準を事前に $\alpha = 0.10$ とすれば、帰無仮説が棄却でき、対立仮説が支持されます。
   - このように有意水準の設定によって帰無仮説の棄却有無が変わります。この意味で帰無仮説の検定には恣意性が含まれるわけです。しかし、社会科学において $\alpha = 0.05$ は一般的な慣行です。これに従いましょう。
4. なぜ p 値が α より小さいと帰無仮説が否定されるのでしょうか
   - 有意水準 (α) を設定することはは「第一種の過誤 (帰無仮説が真なのに、棄却する確率) を犯す確率が何 % 未満なら帰無仮説を棄却するか」の基準 (何 %) を決めることです。p 値は「帰無仮説が正しいと仮定した場合、今回のような結果以上が得られる確率」であり、「今回のような結果以上」は帰無仮説を棄却する方向にいくことなので、結局、p 値というのは「第一種の過誤を犯す確率」となります。先ほどの文章を置き換えたら有意水準の設定は「p 値が何 % 未満なら帰無仮説を棄却するかの基準を決めること」となりますね。
5. 実際に仮説検定をしてみたら理解できそうです。
   - そうですね。来週からは実際に色んな分析手法 (t 検定, ANOVA, χ^2 検定, 回帰分析などなど) をやりながら、出てきた結果の解釈なども一緒に解説してい

きます。その時に分からないことがあったらいつでも質問してください。

6. p 値と α の比較が極小だった場合、結局結果の解釈が大切ということで、数量データといえども絶対じゃないなと思った。
   - そうですね。$\alpha = 0.05$ に設定するのがこの業界の一般的な作法でありますが、それでも恣意的なものということは否めません。なので、p 値そのものだけでなく、実質的な効果 (投票率の差, 体重の差など) も一緒に見て解釈すべきですね。$p < 0.05$ なら万々歳ということはありません。数量データは客観的な値を得るには効果的な手段ですが、その値がどのような意味を持つのかは分析者の主観に依ります。
7. p 値についてネットで適当に検索すると、「その仮説が正しいと仮定したら、今回みたいな結果が起きる確率はこんなにも低い」という主張の「こんなにも低い確率」のことであると書いてあり、ある程度納得できたのですが、そういう理解で大丈夫ですか?
   - ほぼ合ってます。ただし「今回みたいな結果」を「今回みたいな結果以上」に置き換える必要はありますね。もし、トマト摂取の有無による体重差が今回 3kg でしたら、「3kg 以上」の結果が得られる確率になります。
8. スライド 8 の後半) ぴったり 5% の差がある可能性は極めて 0 に近い ⇒ (実は) 特定のポイントである確率は 0← 分からない!何言ってるの!ソンさん、しっかりして!
   - 確率というのは全ての可能性を足したら必ず 1 (=100%) になります。コイン投げなら「表」と「裏」、二つの可能性があり、「表が出る確率」+「裏が出る確率」は 1 (=100%) になりますね[2]。表が出る確率と裏が出る確率が等しいなら各可能性に対する確率は $\frac{1(=100\%)}{2(=\text{可能性の数})} = 0.5$ になります。
   - ならば、無作為に選ばれたある人の身長を考えましょう。その人の身長が 150cm である確率は?151cm である確率は?150.1cm である確率は? 150.0000000000002cm である確率は? このような無限大の可能性があります。もし、各可能性に 0.00001% の確率を付与しても可能性の数が無限なので足したら 1 (=100%) を超えてしまいます。上の例のように 1 を無限大に割ったらどうなるでしょうか。極限の話になるので詳細は省きますが、$\frac{1}{\infty} \simeq 0$ となります[3]。

---

2) コインが立つというミラクルは除きましょう。

3) コイン投げの例は厳密に言えば、特定の値をとる確率は 0 にならないですが、それにしても非常に小さいことには変わりません。また、伝統的統計学の基本的な思想は「無限回」の試行を前提としているので、結局は 0 になります。まぁ、ある意味、深く考えたら負けです。専門的に勉強したい場合は大標本理論

9. スライド 8 の後半) 「◯◯ より大きい/小さい」確率は計算可能 ← 分からない!助けて!ソンさんかっこいい!
    - 確率の定義によると出現しうる値が得られる確率の総和は 1 となり、これを図で表現すると図 1 のようになります[4)]。

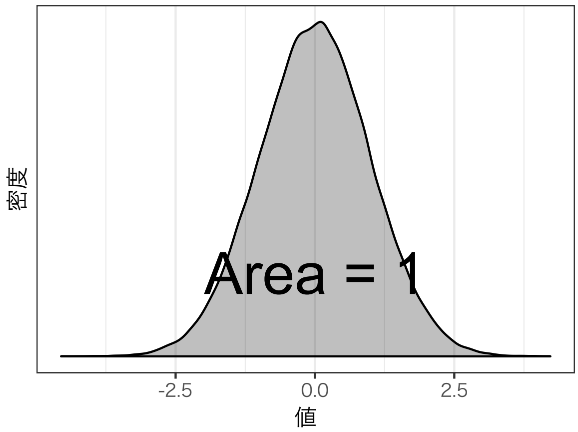

図 1　確率の総和は必ず 1

色塗りの領域の面積は 1 です。つまり、ある変数が約-3 から 3 の間の値をとる確率は 1 (=100%) です。面積=確率ということを考えれば、ある変数が 1 以上、2 未満の値をとる確率は図 2 の色塗り部分の面積となります。

しかし、上で説明したように、特定の値を取る確率は 0 です (図 3)。なぜなら、面積の基本は「横 $\times$ 縦」ですが、横の長さが 0 だからです。

しかし、図 2 と同様、特定の値より大きい確率は以下のように面積を計算すれば算出可能になります (図 4)。

10. α 値の根拠は何なのでしょうか? 結局棄却するかどうかは個人の裁量に委ねられるということなのでしょうか?
    - $\alpha = 0.05$ に設定する一般的慣行に根拠はありません。しかし、この慣行があ

(漸近理論) を勉強する必要があります。

4) グラフの縦軸は相対的な確率 (密度) です。説明はややこしいので割愛します。

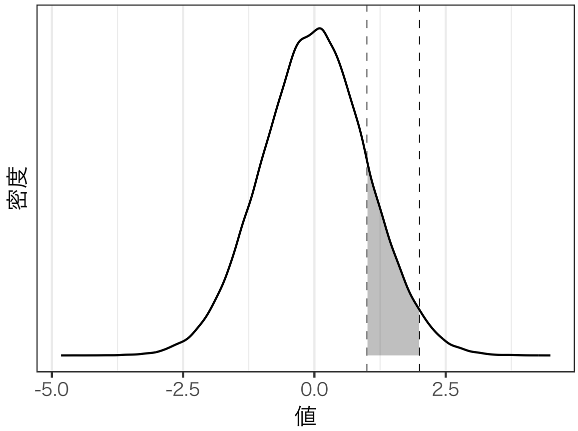

図 2　ある範囲内の値をとる確率は面積と等しい

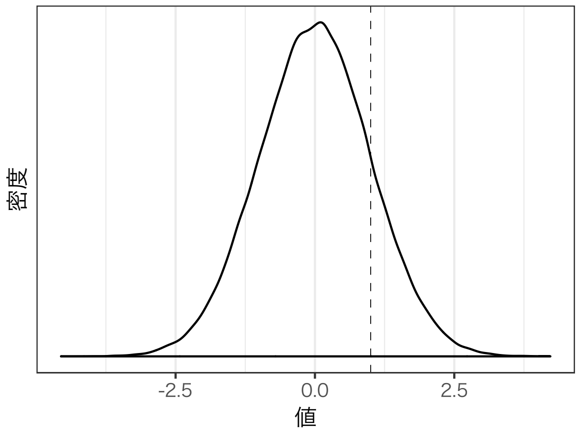

図 3　特定の値をとる確率は常に 0

まりにも一般的すぎて、$\alpha = 0.05$ にしたら誰も文句は言わないでしょう[5)]。

- 分析で得られた結果で p 値が 0.087 の場合、$\alpha = 0.05$ なら帰無仮説が棄却出来ませんが、$\alpha = 0.1$ なら棄却できますね。α は分析者が決めるものなので、棄却するかどうかは個人の裁量となります。したがって、有意水準は必ず分析に入る前に決めるべきだと言われています。分析前に α を決めておけば、恣意的に棄却有無を調整できないからです[6)]。
- しかし、$\alpha = 0.05$ が慣行なので、これを変更するのはかなり大変です。もし、$\alpha = 0.05$ ではなく、$\alpha = 0.1$ とかにしたい場合は、なぜ α を 0.1 にするのかを

5) しかし、医学統計、保健統計などでは命に関わる問題なので $\alpha = 0.001$ とかにするケースも多いです。
6) しかし、論文やレポートでは事前に設定したかどうかは分かりませんので、厳密に言えば個人の裁量になりますね。

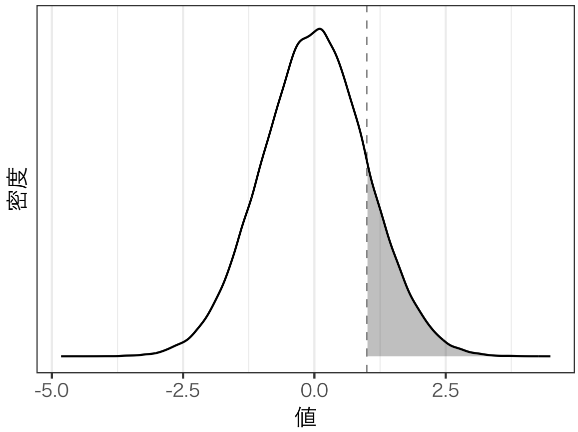

図 4　特定の値以上の確率なら面積で計算可能

説得すべきです[7]。

11. 具体的な帰無仮説、対立仮説などの定義がよくわからなかった
    - 統計的検定の多く[8]は「0 か 0 じゃないか」の検証になります。例えば、トマトを毎日 200kg 食べる人と食べない人の間に体重差があるかどうかは「差が 0 か 0 じゃないか」になります。ここで「0」が帰無仮説、「0 じゃない」が対立仮説になります。つまり、「トマトを毎日 200kg 食べる人と食べない人の間には体重の差がない」が帰無仮説となり、「差がある」が対立仮説になります。
12. 帰無仮説と対立仮設の説明の箇所がよくわからなかったです。初めのリサーチクエスチョンと仮設検定の関係がよくわかりませんでした。なぜ?という問いに対して、どのような仮説を立てると各集団間の性質を比較するようになるのかがわかりません。
    - リサーチクエスチョンから仮説を立てるのは、ある意味、訓練が必要なものです。投票率の例を見ると投票参加を規定する要因は無数にあります。性別、学歴、所得、居住地域、動員の有無、ネットワーク、政党帰属意識、政治的有効性感覚など数え切れないほどです。ここから仮説を導くのは色んな方法がありますが、「未知の要因」を発見してそれを検証するか、「既知の要因」の否定・修正するかです[9]。
    - RQ から仮説へ、仮説から集団間比較への経路を、いくつか例から見てみましょう。

[7] これが面倒くさいので $\alpha = 0.05$ にするか、ベイズ統計をやるわけですね
[8] 全てといっても過言ではありません。
[9] 講義で紹介したのは既知の要因の再確認なので、実は学術的な意味はないんですね。

RQ1　日本の野球はなぜ強いのか

仮説　日本人は毎日米を食べているから[10)]

比較　米が主食である国の野球成績と、主食ではない国の成績を比較

RQ2　なぜ法学政治学計量分析の講義は楽しいのか

仮説　それは講師のソンさんがかっこいいから

比較　講師がかっこいい授業とかっこよくない授業の履修者の授業満足度を比較

RQ3　ソンさんはなぜかっこいいのか

仮説　毎日ラーメンを食べているから

比較　毎日ラーメンを食べている人と食べない人のイケメン度を比較

10) 具体的にいえば、「国が主食である国の野球成績は良い」という形でしょうね。