【統計一口メモ　第9話】

＜両側検定と片側検定　使い分けていますか？＞

名古屋市立大学大学院医学研究科 非常勤講師　薬学博士　松本一彦

第8話でFisher正確検定の両側検定について、統計ソフトによって手法はまちまちであることを述べました。今回はさらに、片側検定の使い方について統計家のニュアンスが異なることと検証的な臨床試験と探索的な前臨床試験では有意水準が異なることについても触れてみようと思います。

＜帰無仮説と対立仮説＞

２種の薬剤のうちA剤は標準品でB剤は新薬。in vitroではB剤の有用性が確認されています。

仮説には「２つの薬剤に差がある」 と 「A剤よりB剤の新薬の方が優れている」が考えられます。

帰無仮説は，　“差がない”　　H_0：$\mu_1=\mu_2$

対立仮説は、　“差がある”　　H_1：$\mu_1\neq\mu_2$

“優れている”　　H_1：$\mu_1<\mu_2$　or　$\mu_1>\mu_2$

対立仮説は前者を選ぶと両側検定、後者は片側検定となります。

＜片側検定を行う際の帰無仮説の取り方＞

対立仮説をH_1：$\mu_1<\mu_2$に設定するとします。このとき、帰無仮説をH_0：$\mu_0\geqq\mu_1$としてはダメ！

帰無仮説は便宜的な“仮説”に過ぎず、成り立っているかどうかを知りたい仮説ではありません。成り立っているかどうかを知りたいのは対立仮説の方です。したがって、H_0：$\mu_1=\mu_2$としなければなりません。

＜片側検定の方で有意になったのでこちらを採用しました！＞　※影の声「オイオイ」

永田先生は「統計量が1.80でした。両側検定では1.80<1.960で“有意差なし”でした。でも、片側検定では1.645<1.80で“有意差がみられた”ので、こちらを採用しました！」という報告に

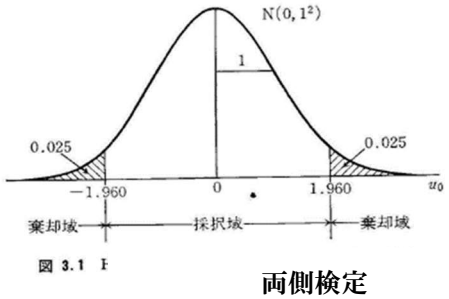
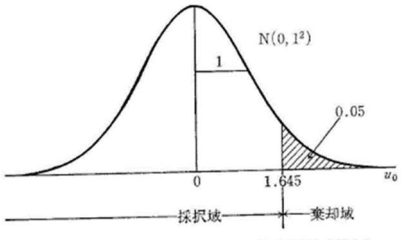


図 3.1 F

両側検定



片側検定

「これは、データを見たあとに仮説や棄却域を自分の都合のよいように設定すること、いわば、“反則行為”です。仮説や棄却域の設定は、データを取る前に行わなければならない[1)]」と述べています。

※論文執筆者の行為は恣意的と言われても仕方ありません。片側検定をするかしないかは、永田先生の言われるようにデータを取る前、プロトコールを書く段階で決めておかなければいけません。そのことについて、多くの統計家が著書の中で、いろいろな表現で述べています。それを紹介しながら片側検定の使い方について学んでいきましょう。私の考えは※で述べていきます。

＜片側検定は稀！いつも両側を＞

Dr．Altman は「片側検定が適切なことは稀である。事前に強い期待、例えば新治療は旧治療より悪いはずがないという期待をもっているときでさえも、それが正しいことを確信することはできない。確実であれば実験する必要がない！片側検定が本当に適切であると思うのであれば、その決定はデータを解析する前になされなければならない；結果がどうであったかによるべきではない。いつも両側を使うことを勧める」と述べています[2)]。

※適切であるかどうかをどうやって判断するのでしょう。解析前に対立仮説で$\mu_1 < \mu_2$を設定し、それを検定することで判断するものと思うのですが。Dr．Altman がレフリーの論文では片側検定はボツになりますね。

＜両側検定を用いる方が問題になることが少ない＞

丹後先生は「一般に、一方の群が、他方の群に比べて、大小関係が予想できる場合に片側検定を用いることができる。しかし、大小関係が予測できたとしてもその関係が、研究者間で認められていない場合には片側検定を用いるのは望ましくなく、さらに、それは、両側検定に比べて、有意でない差を有意としてしまう可能性が高くなるので、両側検定を用いる方が問題になることは少なくてすむ」[3)]。

※「研究者間で認められる」ということってどういう場合なんだろう？あたらしい事象に関しては研究者は自分だけかもしれない。「問題になる」のを避けるために両側検定にするという行為は多くの統計家が取っている。

＜両側検定としておくのが安全＞

Dr. Armitage は「両側検定よりも片側検定を用いたいという誘惑が生じるかもしれない。データをながめて、仮説からのずれの方向を見定めてから片側検定の使用を決定することは決して許されるべきものではない。仮説からのある方向への乖離は常に偶然にしか生じず、従ってどんなにずれが大きくてもそれは意味のないものであると見なせることが確実な場合に限って片側検定を用いることを決めてよい。しかも、その決定はデータを調べる前に行うのである。このような状況は現実には稀であり、有意性検定はほとんどいつも両側検定としておくのが安全であろう」[4)]

※対立仮説の H_1：$\mu_1 < \mu_2$ をよほど稀な場合にしかとってはならぬという主張。血沈のような上にあがることがないようなデータしか使えないということだろうか？ここでも両側検定に安全の地を求めている。

＜あえて両側検定にしておくのが賢明＞

佐藤、小西先生は「片側検定か両側検定か？というのも、どうもよく分かりにくいところです。あまり片側検定を使っている例は見かけません。どうしてでしょう。ある薬の効果を調べるためにプラシーボとの比較を行う際、プラシーボの方が実際の薬より優れた効果があると仮定することはないでしょう。したがってこの場合、理屈で言えば片側検定を使うのが適当だと考えるのが自然です。でも、片側検定による報告がされていることは稀です。その理由の１つには、片側検定では「有意差が出やすくなる」ということで、厳しい査読委員に好まれない傾向があると言えるでしょう。したがって、結論としては、両者の違いを理論的によく理解したうえで、あえて両側検定を使用しておくのが賢明かと思います」[6)]。

※理屈で言えば片側検定を使うのが適当と言いながら、最終的には両側検定を使うのが賢明とレフリーに忖度しているのはいただけません。実務者は再試験ができない緊迫した状態にある場合に忖度している余裕はありません。レフリーに結果を見る前から片側を設定したことをしっかり述べるべきです。

＜片側検定は事前に決めておけばよい＞

浜田先生は「両側検定では対立仮説の方向を欲張るため片側検定に比べて、ｐ値が大きく有意になり難くなります。多くの場合、両側検定のｐ値は片側検定のそれの２倍となります。通常何も断らずに検定といえば、両側検定を意味します。両側検定では有意にならず、片側検定では有意となる微妙なケースでは、事後的に両側検定を片側検定に変えたい誘惑にかられますが、片側検定と両側検定のどちらを用いるかは、検定を行う前に決めておく必要があります」[5)]

※検定を行う前に「片側検定を使う」と宣言しておけば、それで OK という実務者の立場を理解した記述です。

＜芳賀先生の私見＞

厚労省の「承認申請のための試験では、片側検定の第１種の誤りを両側検定で慣例的に用いている値の半分に設定する方法が好ましい」という提言に対して、「これは“企業は自社に有利になるように両側・片側検定を設定する危険がある“という性悪説に基づいており、企業の統計担当者を無能力扱いしているように思われる。統計担当者が社内で権威ある立場を維持できるようになり、妥当な指針が出されるようになることが期待される」[7] と述べています。

※芳賀先生は片側検定を対立仮説に基づいて使用すべきであることを主張されていました。もちろん、後付けのような使い方ではなく、検定前に宣言しておくことは当たり前として、当局やレフリーの顔色を伺うような卑屈な態度はとるべきではないと。

＜χ^2検定は片側か両側か＞

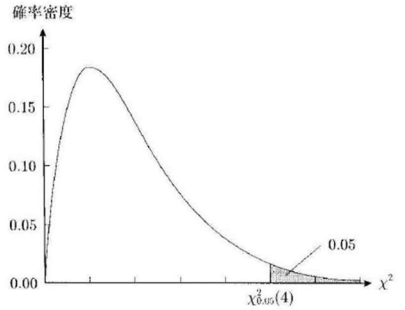


図５　χ^2 分布の密度関数(自由度 4)

「X^2 の実現値がこの分布の上側確率 0.05 の点より上の値であれば、仮説が正しければ頻繁には起こらないことが生じているとして、帰無仮説 H_0 を有意水準 0.05 で棄却する」[8]。

JMP アウトプット

検定

要因	自由度	(-1)*対数尤度	R2乗(U)
モデル	1	2.340874	0.1365
誤差	23	14.807371	
全体(修正済み)	24	17.148245	
N	25		

検定	カイ2乗	p値(Prob>ChiSq)
尤度比	4.682	0.0305
Pearson	4.573	0.0325

JMP の χ^2 検定は「Pearson」として出力される。

Pearson の方法は２項分布の正規近似を用いる両側検定であり，連続修正は加えられていない。標準正規分布を２乗したものが自由度１の χ^2 分布であり，標準正規分布が０で折り返されて２乗されるので，正規分布の両側確率が χ^2 分布の上側確率になる。

※χ^2 検定は χ^2 分布でみると上側確率 0.05 の片側検定です。しかし、通常正規近似で計算され JMP のように Pearson として両側検定で表記されます。どちらにしても、確率 0.05 には変わりはありません。でも、χ^2 検定として表記するのであれば、片側検定ということになります。

１）永田靖「統計的方法のしくみ」日科技連 1997 年

２）D.G. Altman「医学研究における実用統計学」サイエンティスト社 1999 年

３）丹後俊郎「医学への統計学」第 3 版　朝倉書店 2013 年

４）P. Armitage，G. Berry「医学研究のための統計的方法」サイエンティスト社 2001 年

５）浜田知久馬「学会・論文発表のための統計学」真興交易医書出版部 1999 年

６）佐藤敏彦、小西宏明「StatView4.0 日本語版」BNN1994 年

7）芳賀敏郎「医薬品開発のための統計解析」第１部　サイエンティスト社 2016 年

８）広津千尋「データ科学推論の基礎」岩波書店 2018 年