同志社社会学研究　NO. 23, 2019

【セミナー】

統計的仮説検定の原理と実際
――社会調査の推測統計（4）――

小林　久高
KOBAYASHI Hisataka

1 はじめに

今回は統計的仮説検定の原理と実際の方法について解説する。検定の方法について解説した書籍は多い。多くの人は自分が解こうとしている問題に似た例をそういった解説書で見つけて検定を行っている。それはそれでいいのだが、何を行っているのかよくわからず、気持ちの悪さを感じている者も多いだろう。この章では、こういった気持ちの悪さを解消するために、基礎的な推測統計の原理からきちんと解説していこうと思う。

以下の解説を理解するために必要となるのは、(1) 確率変数とその合成、(2) 母数と標本統計量、(3) 正規分布の性質、(4) さまざまな確率分布の相互関係、についての知識である。これらの知識があるならば、統計的仮説検定の原理は無理なく理解できるし、それぞれの検定においてなぜある式が用いられるのかということもイメージできるようになる。上記 4 つの知識については、「母集団・標本・確率変数」「離散型確率変数とその分布」「連続型確率変数とその分布」（小林久高,2018a, 2018b, 2018c）でまとめておいたので、まだ読んでいない読者はそちらを先に読んだほうが理解が早いと思う。

4〜7 節ではさまざまな検定法について解説している。すべて読む時間がない場合は◆マークのついたところだけをきちんと読むといい。また、「補足」には理解を深めるための事柄が書いてある。必要に応じて読んでほしい。

具体的に検定を行うに際しては、確率分布の数表の読み方を知っておく必要がある。特に χ^2 分布と F 分布の数表の読み方には注意が必要だ。補に書いておいたので利用してほしい。なお、数表自体は小林（2019）に添付している。

2 統計的仮説検定の基本的な考え方

2.1 ボールの検査

統計的仮説検定は、次のような場合に使う方法だ。

あるスポーツ用品のメーカーに勤めている A さんは、取引先の担当者から「お宅のバレーボールのボール、重量 270 g と書いてあるけどちょっと軽めじゃないかな、使っている選手から軽いとよく言われるんだ」と言われた。それで工場長に伝えると、「ボールによって多少の違いはあると思うけど平均は厳密に 270g になるように作っているよ。運動選手は敏感だからそう感じる人もいるかもと思って、前に調べたデータをもとに『重量は平均で 270 g ですが、標準偏差が 6g あるため、商品によっては 270g と多少異なる場合があります』と書いた説明書も添付してる。丁寧だろ」と言う。そして「でもボールの重さはすごく大事なことだから、在庫で平均 270 g になっているか調べてくれないか」と言った。そこで、A さんは「いいですよ」とこの仕事を引き受けることにした。

Aさんが、この安請け合いを後悔したのは在庫のボールの山を見たときだった。2m四方のケースが10、ざっと見て1万個はある。いったいこの仕事にどれぐらいの時間がかかるのかAさんは計算しだした。「1個測って記録するのにがんばっても15秒はかかるな。ボールは1万個だとして、15×10000÷60÷60＝41.666。げっ、41時間もかかるのか。それにボールを取りに行ったり返したりするのに1個20秒はかかる。20×10000÷60÷60＝55.555時間。あわせて約100時間、週40時間働くとして2週間以上かかる。Aさんはぞっとした。こんなことはできないぞ、他になにもできなくなってしまう。かといってみんな忙しいから手伝ってもらうわけにもいかないし…。

そこで、Aさんは標本をとって重さを測ることを考えた。工場長もそれでいいと言っている。Aさんは大学時代に勉強した統計学をおさらいし、翌日作業することにした。

次の日、Aさんは山のようにあるボールからランダムに1つを選び、その重さを測って記録し、ボールをまた元に戻すという作業を25回行った。つまりとっては返すという復元抽出でサイズ25の標本を得たのである。この25回測った重さの平均、すなわち標本平均は267gだった。Aさんは「やっぱり270gより軽いのかな」と思った。そして少し慎重になって「でも、標準偏差が6gと書いてあったから在庫の平均は270かもね、昨日おさらいした統計学で検定をしてみるのがいいな」と考えた。「270gではない」という考えが正しいかどうかを検討する際に統計的仮説検定を行うと、「270gではない」あるいは「270gではないとは言えない」という結論を出すことができる。そして「270gではない」とする場合にも「ひょっとすると間違っているかもしれないが、その間違いの可能性は○％だ」などと言えるからだ。

2.2 考えの筋道

この検定について筋道を追って見ていこう。まず、ボールの重さをXとし、それが正規分布すると考えよう。すなわち、

$$X \sim N(\mu, \sigma^2)$$

そして、添付された説明書の世界を母集団と考える。すなわち、ボールの重さの平均は270gであり、標準偏差は6g（分散は36）という世界を基本に考えるのである。すると次のように表現できる。

$$X \sim N(270, 36)$$

ここで、母数と標本平均の次の関係を思い出そう。

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

すると、母平均270、母分散36の正規母集団からのサイズ25の標本の標本平均$\bar{X}$は次の式のように分布することになる。

$$\bar{X} \sim N\left(270, \frac{36}{25}\right)$$

すなわち、説明書の表記通りだとすると、サイズ25の標本におけるボールの重さの標本平均$\bar{X}$は、標本の取り方次第でさまざまな値になり得るが、母平均270を中心に分散36/25（標準偏差6/5＝1.2）の広がりをもって図1のように正規分布するのである。この図から、標本平均が270周辺の値になる確率は高いが、267.6より小さい値や272.4より大きな値になる確率は低いことがわかる。

図 1　標本平均の確率分布

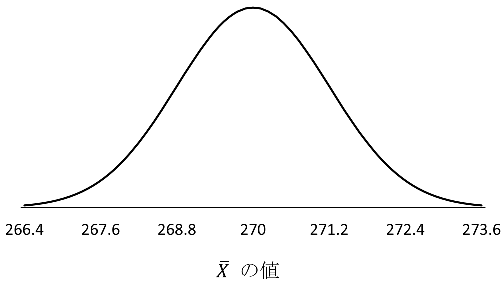
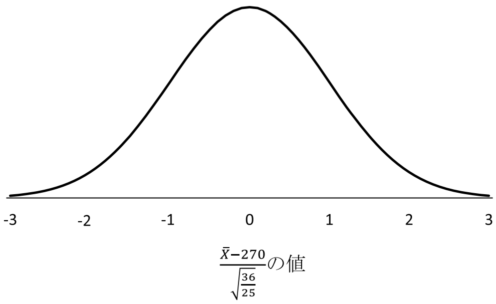


図 2　検定統計量の確率分布



実際に得られたサイズ 25 の標本の標本平均は 267g だった。この値は説明書に書かれた平均 270g、標準偏差 6g（分散 36g）の母集団から得られた標本ではあまり出ない値であることが図からわかる。

では、どれぐらい出にくい値なのだろうか。これを確率でとらえるためには、標本平均 $\overline{X}$ を標準正規分布に従う変数に変換する必要がある。この変換されたものを検定統計量と言い、それは下の式のようになる（X が正規分布に従うとき aX＋b も正規分布に従うという性質があるので、この検定統計量も正規分布に従う）。

$$\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

ボールに添付された説明書にあるように、母平均が 270、母分散が 36 であり、標本サイズが 25 の場合、この式は次のようになり、その確率分布は図 2 のようになる。

$$\frac{\overline{X}-270}{\sqrt{36/25}} \sim N(0,1)$$

数表によると、この分布において 1.96 の値に対応する上側確率は 0.025 である。正規分布は左右対称だから、ここから「－1.96 以下または＋1.96 以上」になる確率は 0.05（5%）になることがわかる（同様に「－2.58 以下または＋2.58 以上」となる確率は 0.01（1%）であることもわかる）。

さて、実際にとってきたサイズ 25 の標本の平均 267g であったが、この値がどれぐらい出にくいを明らかにするためには、上の式の $\overline{X}$ のところに実際の標本平均 267 を代入し、その位置を図 2 で確認するとともに、その確率を考えればいい。実際の標本平均を代入すると次のようになる。

$$\frac{267-270}{\sqrt{36/25}} = \frac{-3}{6/5} = \frac{(-3)\times 5}{6} = -2.5$$

すでに述べたように標準正規分布で値が「－1.96 以下または＋1.96 以上」となる確率は 0.05（5%）だ。標本から計算された－2.5 という値は－1.96 よりずっと小さい。平均 270 g 母集団から得た標本（サイズ 25）の平均がこんな値になる確率は 5% 以下なのである。

そこから次のように推論することができる。すなわち、「このサイズ 25 の標本は、平均 270g、標

準偏差 6g（分散 36）の母集団から得られたものではない。だから、実際に標本が得られた母集団（すなわち在庫全体）の平均は 270g ではないはずだ。ただし、このことは完全に正しいとは言えない。たまたまそんな値（−2.5）になるような標本が、平均 270 g の母集団から生じたかもしれないからだ。しかし、そんなことが起こる確率は 5%以下だ」。

A さんは工場長にこう伝えた。「在庫のボールの重さの平均は 270 g ではないです。やはり取引先の人が言うように軽いと思います。ただ私の言っていることは間違っていて、本当に在庫の平均は 270 g かもしれません。でもそんな可能性は 5%以下です」。これを聞いて工場長はボールの製作工程を見直すことにした。A さんはその日の午前中だけで作業を終えることができた。

以上が統計的仮説検定の基本的な考え方である。母平均と標本平均の関係さえ知っていれば、以上のことは容易に理解できるはずである。次節ではいくつかの用語を導入し、検定のプロセスを整理して述べていくことにする。

3 検定のプロセスと関連用語

3.1 帰無仮説と対立仮説

統計的仮説検定においては帰無仮説（H_0）と対立仮説（H_1）という 2 つの仮説をもとに検定が行われる。帰無仮説とは検定によって否定しうる命題のことであり、対立仮説とは帰無仮説の否定によって主張しうる命題のことである。さきのボールの例での帰無仮説は「母平均は 270g である」というものであり、対立仮説は「母平均は 270g ではない」というものである。

検定は、帰無仮説を否定（棄却）して対立仮説を主張するという方法をとる。ここで、なぜ直接対立仮説を主張するのではなく、そのような迂回的方法をとるのかという疑問が生じるだろう。ボールの例をもとにその理由について述べよう。

例では、「H_0：母平均は 270g である」という帰無仮説を否定することによって「H_1：母平均は 270g ではない」という対立仮説を主張した。ここで帰無仮説にもとづく標本平均の確率分布は 1 つに確定する。しかしながら、対立仮説にもとづく標本平均の確率分布は母平均が 271g の場合、272g の場合、273g の場合…と無数に存在する。検定は標本統計量の確率分布をもとに行われるので、それが確定しないと検定を行うことは難しい。それゆえ、標本統計量の確率分布が確定しない対立仮説を直接検討することはせず、分布が確定する帰無仮説を否定して対立仮説を採択するという方法をとるのである。さらに、帰無仮説を否定して対立仮説を採択するという方法をとれば、「本当は帰無仮説が正しいのに、誤って対立仮説を採用してしまう確率は 5%以下」などということも明らかにできる。これも帰無仮説を否定して対立仮説を採択するという方法をとる理由である。

母平均・母分散・母比率などを母数と言うが、帰無仮説は、1 つの母集団で「母数が特定の値であること」を述べたり（ボールの例）、2 つの母集団で「母数に差がないこと」を述べたりする。また、「2 変数間の相関がない（すなわち相関 0）こと」を述べるような帰無仮説もある。なぜこのような帰無仮説が立てられるのかというと、こういった帰無仮説の下では検定統計量についての特定の確率分布が描けるからである。帰無仮説が「母数が特定の値でない」ことを述べたり、「母数に差がある」ことを述べたり、「相関がある」ことを述べたりすることはない。それでは標本統計量について特定の確率分布は描けないからである。

3.2 定理と検定統計量

検定は母集団の母数と標本のサイズと標本統計量から作成される検定統計量についての定理をもとに行われる。それはたとえばこんなものである。

■定理
「母平均μ、母分散σ^2の正規母集団 N(μ, σ^2)からのサイズ n の標本について、標本平均を$\overline{X}$とすれば、$\dfrac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}$は、標準正規分布 N(0,1)に従う」

実はこれはボールの例で用いた定理である。定理にある次の式は検定で用いる統計量で、検定統計量と呼ばれる。ここでは Test と表記することにしよう。

■検定統計量

$$Test=\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}$$

この検定統計量は確率変数である標本平均$\overline{X}$と決まった値である母平均μ、母分散σ^2、標本サイズ n から合成された確率変数である。定理が述べているのは、母集団が正規分布に従うとき、この確率変数である検定統計量が標準正規分布に従うということである。

検定では、この検定統計量の式に現実の標本から得られた平均$\overline{x}$を代入して値を求め、検討を進める。それは次式のようになる。

■検定統計量の値

$$test=\frac{\overline{x}-\mu}{\sqrt{\sigma^2/n}}$$

ここでは小文字の test としてこの検定統計量の値を示している。大文字で始まる Test は確率変数であり確率分布に従う。それに対して、小文字で始まる test は現実の標本から得られる平均$\overline{x}$を用いているので決まった値になる（4 以降で種々の検定の解説を行うが、確率変数としての標本統計量（標本平均、標本分散、標本比率など）は大文字で、その現実の標本における値（標本平均、標本分散、標本比率など）は小文字で表すことを基本とする）。

3.3 検定統計量の確率分布

検定で用いる定理は行う検定の種類によって異なる。たとえば、上の定理は母分散がわかっている場合の母平均の検定についてのものだが、母分散が不明の場合には別の定理を用いることになる。そうなると、検定統計量も違ったものになり、検定統計量の確率分布も違ったものになる。

初歩的な推測統計で用いる確率分布は標準正規分布、χ^2分布、t 分布、F 分布である。後の解説を理解するためには、それらの確率分布相互の関係について思い出しておくほうがいい（小林,2018c）。すなわち、Z を標準正規分布に従う確率変数とし、K_nを自由度 n のχ^2分布に従う確率変数とし、T_nを自由度 n の t 分布に従う確率変数とし、F^m_nを自由度 m,n の F 分布に従う確率変数とすると次のような関係がある（すべての右辺の変数は独立）。

$$K_1=Z^2$$

$$K_n=\sum_{i=1}^{n}Z_i{}^2=Z_1^2+Z_2^2+\cdots+Z_n^2$$

$$T_n=\frac{Z}{\sqrt{K_n/n}}$$

$$F_n^m=\frac{K_m/m}{K_n/n}$$

また、X が正規分布に従うときは aX＋b も正規分布に従うこと、正規分布やχ^2分布には再生性という性質があることも思い出しておこう（小林,2018c）。

3.4 有意水準と棄却域

ボールの例では、次のように検定統計量の値を求めた。

$$test = \frac{\bar{x} - \mu}{\sqrt{\sigma^2/n}} = \frac{267 - 270}{\sqrt{36/25}} = -2.5$$

ここでこの値がどれくらいならば帰無仮説を棄却できるのかということが問題になる。検定統計量が標準正規分布に従う場合、この値が「－1.96 以下または＋1.96 以上」となる確率は 0.05（5%）となり、「－2.58 以下または＋2.58 以上」となる確率は 0.01（1%）となるので、これらの値を基準に帰無仮説を棄却するかどうかを決定することが多い。

この「－1.96 以下または＋1.96 以上」などといった「帰無仮説を棄却する値の範囲」のことを棄却域と言う。また、棄却域の範囲に対応する確率のことを有意水準と言う。「－1.96 以下または＋1.96 以上」という棄却域の有意水準は 5%である[1]。

棄却域と有意水準の関係については混乱することがある。棄却域は値の範囲の話であり、確率分布のグラフで言うと横軸の話、有意水準とは横軸の棄却域に対応する確率（面積）の話であるということをきちんと押さえておこう。

さて、ボールの例に戻ると、－2.5 という値は「－1.96 以下または＋1.96 以上」という有意水準 5%の棄却域にある。それゆえ、次のように言うことができる。

「有意水準 5%（両側）で、ボールの重さの平均は 270g だとする帰無仮説は棄却でき、ボールの重さの平均は 270g ではないと言える」

帰無仮説が棄却されるかどうかは有意水準によって変わってくる。たとえばボールの例では検定統計量は－2.5 だったのだが、それは有意水準 1%の棄却域内（－2.58 以下または＋2.58 以上）にはない。したがって、有意水準 1%で検定した場合には次のように言うことになる。

「有意水準 1%（両側）でボールの重さの平均は 270g であるという帰無仮説は棄却できず、ボールの重さの平均は 270g でないとは言えない」

検定ではこのように有意水準によって結果が異なるので有意水準を常に明示しておく必要がある。簡単に述べる場合も次のようにしておく必要がある[2]。

「平均は 270g ではない（p<0.05 両側）」

統計ソフトを利用している場合、有意水準の棄却域に検定統計量の値があるかどうかではなく、確率そのものが示されることが多い。正規分布や t 分布といった左右対称の確率分布を利用して検定

---

1 この 1%や 5%のことを危険率とも言う。

2 有意水準はこのように 1%や 5%とすることが一般的だが、いつもそうでなくてはならないわけではない。帰無仮説を棄却する方が安全性を考えて得策な場合には有意水準を大きくすることも必要である。ボールの例で、「平均が 270g でないことは即座に会社の存亡にかかわる」ということであれば、有意水準をより大きく 10%にして棄却域を広くして帰無仮説が棄却される可能性を高めるといい。そのとき言えるのは「間違えている確率は 10%あるが、平均は 270g ではない」ということである。

が行われる場合、ソフトでは検定統計量の値と符号を変えた検定統計量の値の外側の確率が示される。

ボールの例の検定統計量は－2.5 だった。

$$test = \frac{\overline{x} - \mu}{\sqrt{\sigma^2 / n}} = -2.5$$

数表を見ると－2.5 より小さな値の範囲の確率を求めると 0.0062 であり、符号を変えた＋2.5 より大きな値の範囲の確率も 0.0062 である。ソフトではこれら両者を足した 0.0124 が有意確率として表示されるのである。われわれはこの数値を見て、0.0124＜0.05 なので有意水準 5%（両側）で帰無仮説が棄却されると判断することになる（ソフトによっては足し算をせずに確率を表示するものがあるかもしれない）。

3.5 片側検定と両側検定

さて、上の説明で「両側」と書かれていることに気づかれただろうか。これはこの検定が両側検定の結果だということを意味している。

ボールの重さの平均についての検定結果は「270g ではない」というものだったのだが、「270g より軽い」と述べているわけではない。それゆえ「－1.96 以下または＋1.96 以上」というように棄却域を確率分布の左右両側において、検定を進めたのである（図 3）。

もし、「H_1：ボールの重さの平均は 270g より軽い」と主張したいならば、棄却域は片側にだけ設定すればいい。この場合、有意水準 5%の棄却域は、「－1.65 以下」となる（図 4）。

図 3 および図 4 の確率分布の左側だけを見ると、片側の棄却域「－1.65 以下」は両側の棄却域「－1.96 以下」より広いことがわかる。ここから、左側だけを考えるときには、片側検定の方が両側検定よりも帰無仮説が棄却されやすいということがわかる。そのことは主張したい対立仮説の採択にとって有利であることを意味する。しかしながら、一般的に言うと検定は厳しい方がいいので、「270g より軽い」と思っている場合でも、「270g ではない」という対立仮説を立てて両側検定をするのが普通である。

図 3　両側 5%棄却域

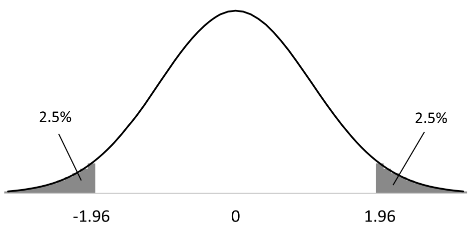
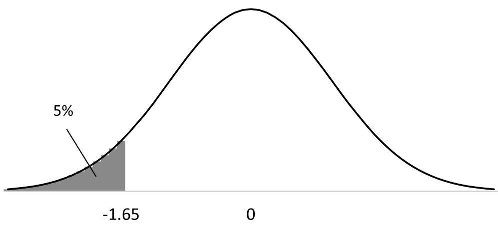


図 4　片側（左側）5%棄却域



両側検定と片側検定については注意しておくべきことが 2 つある。1 つめは確率分布の左側を見るか右側を見るかは検定統計量の式がどうなっているのかによるということである。ボールの場合の検定統計量の式からすると「270g より軽い」と言うためには分布の左側を見ればいい。しかし、もし

検定統計量が分布の左右を反転させた次のようなものだったら右側を見る必要がある。「小さいこと」を主張したいときは必ず左を見ると覚えておくのは危険だ。

$$Test' = \frac{\mu - \overline{X}}{\sqrt{\sigma^2/n}}$$

2 つめはやや専門的だが重要なことだ。それは「両側検定」と「片側検定」という用語が、論者によって（1）対立仮説について述べている場合と、（2）確率分布について述べている場合があるということである。ここでの例で言うと（1）の対立仮説のあり方の意味の場合、両側検定とは$\mu \neq 270$が検討されるということを意味し、片側検定とは$\mu < 270$または$\mu > 270$が検討されるということを意味する。（2）の分布のあり方から言うと両側検定とは、検定統計量の棄却域が確率分布の左右に設定されることを意味し、片側検定とは分布の右または左だけに設定されることを意味する。ここでの例のように標準正規分布や t 分布を用いた平均の検定では（1）と（2）は同じことになる。すなわち、（1）の意味で両側検定をしたければ、（2）の意味で両側に棄却域を設定すればいい。しかしχ^2分布や F 分布を用いた比率の検定などでは（1）と（2) にずれが生じうる。たとえば、比率について、χ^2分布を用いて（1）の意味での両側検定をする際には、（2）の意味での片側検定を行う必要があるのだ。このシリーズでは混乱を避けるため、両側検定と片側検定という用語を（2）ではなく、（1）の対立仮説のあり方の意味で用いるので注意してほしい。

3.6 帰無仮説が棄却されない場合

帰無仮説が棄却されないときどう考えるかということには注意が必要である。検定は帰無仮説を棄却し対立仮説を採用するという方向で進められる。ここでもし帰無仮説が棄却できなければどうなるのか。常識的に考えると帰無仮説が棄却されなければ「帰無仮説が正しい」という結論が導けそうだ。しかしそれは間違い。帰無仮説が棄却されないときも帰無仮説が正しいとは言えず、帰無仮説については何も言えないというのが正しい。

具体的に言うとこういうことだ。ボールの例の帰無仮説と対立仮説は次のようになっていた。

帰無仮説 H_0：ボールの重さの平均は 270g だ。
対立仮説 H_1：ボールの重さの平均は 270g ではない。

前の例とは異なり、サイズ 25 の実際の標本の平均が 269g だったとしよう。このとき、検定統計量の値は、

$$test = \frac{\overline{x} - \mu}{\sqrt{\sigma^2/n}} = \frac{269 - 270}{6/5} = -1.2$$

この値は、有意水準（両側）1%でも 5%でも棄却域に入らず、帰無仮説は棄却できない。このとき。「重さの平均は 270g である」と言うことはできないのである。言うことができるのは「『重さの平均は 270g でない』とは言えない」ということ、すなわち「対立仮説は採択できない」ということだけだ。検定の結果、正しいと言えるのは対立仮説のみであって、帰無仮説については正しさは主張できないのだ[3]。

[3]「棄却」という用語は帰無仮説にのみ用いられる。「採択」という用語は対立仮説にのみに用いられれば話はスッキリ

このような方針をとらないならば、実に不自然な事態が生じる。今、現実にある 1 つの標本を用いて 2 つの検定を行うとする。それらは「母平均は A」という帰無仮説による検定と「母平均は B」とする帰無仮説による検定である（B≠A）。2 つの検定の対立仮説はそれぞれ「母平均は A でない」「母平均は B でない」である。そしていずれの検定においても帰無仮説が棄却されないとする。

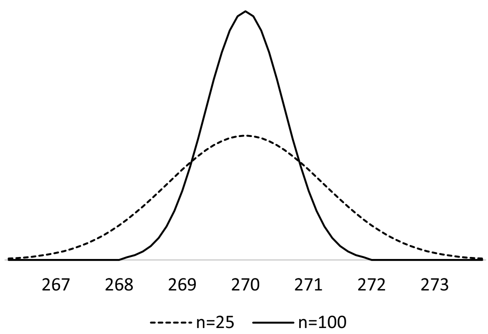
このときもし「帰無仮説が棄却されないならば帰無仮説が正しいと言える」という方針をとると、「母平均は A でありかつ B である」というおかしな結果になってしまう。一方「帰無仮説が棄却されないときには、対立仮説を主張できないだけであり、帰無仮説が正しいと言えるわけではない」という方針だと「母平均は A でないとは言えず、また B でないとも言えない」となり、おかしなところは何もない [4]。

3.7 標本サイズと検定

標本サイズは検定結果に大きく影響するということも押さえておく必要がある。標本平均の分散は σ^2/n だが、これは n が大きくなればなるほど小さくなる。ボールの例と同じ母平均 270、母分散 36 からの標本の場合、標本サイズによって標本平均の分布がどう変わるかを示したものが図 5 である。

この図を見ると、標本平均が 268 といった値になることはサイズ 25 の標本ではよくありそうだが、サイズ 100 の標本ではあまりなさそうだということがわかる。

図 5　標本平均の分布（母平均 270, 母分散 36）



このことは検定統計量の値に反映する。サイズ 25 の標本での検定統計量の値は

$$test = \frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}} = \frac{268-270}{\sqrt{36/25}} = -1.67$$

であるのに対し、サイズ 100 では次のようになる。

$$test = \frac{\bar{x}-\mu}{\sqrt{\sigma^2/n}} = \frac{268-270}{\sqrt{36/100}} = -3.33$$

したがって、実際の標本の平均が仮に 268 である場合、「母平均は 270 である」という帰無仮説は、サイズ 25 の標本では有意水準 5%（両側）でも棄却できないが、サイズ 100 の標本では有意水準 1%（両側）でも棄却できるのである。標本サイズが大きいことはこのようなメリットを持っている。

標本サイズが大きいことは別の点でも重要だ。ボールの例で用いた定理は次のようなものだった。

するのだが、現実には帰無仮説にも用いられる。したがって、帰無仮説が棄却されなかった場合、「帰無仮説が採択されるが帰無仮説は正しいとは言えない」という妙な表現が必要になる。

[4] 帰無仮説が真であるにもかかわらずそれを棄却する誤りを第一種の過誤と言い、普通その確率は 1%や 5%以内に抑えられる。これが有意水準（危険率）である。これに対して帰無仮説が偽であるにもかかわらずそれを採択する誤りを第二種の過誤と言い、その確率は通常計算できない。後で紹介するさまざまな検定の内、5.2 の等分散性の検定の失敗はこの第二種の過誤に関連する。

「母平均μ、母分散σ^2の正規母集団 N(μ, σ^2)からのサイズ n の標本について、標本平均を$\overline{X}$とすれば、$\dfrac{\overline{X}-\mu}{\sqrt{\sigma^2/n}}$は、標準正規分布 N(0,1)に従う」

この定理の中には母集団が正規分布に従うという条件が含まれている。では母集団が正規分布に従わないときはどうなるのか。

ここで思い出すべきことは、次の中心極限定理である。すなわち、

「n が大きいとき、同一の分布（平均μ,分散σ^2)に従う独立な確率変数X_1〜X_nの平均$\overline{X}$は、平均μ,分散σ^2/nの正規分布に従う」

したがって n が大きい時には母集団がどんな分布であっても標本平均は正規分布に従い、それを変換した検定統計量も正規分布に従うことになる。そこで、正規母集団でなくても、標本サイズが大きい時にはボールの検定に用いた定理が一般に利用できる。標本サイズが大きければ母集団の分布の形状に気遣う必要はなくなるのである。

社会調査で検討される変数にはさまざまなものがあり、収入などはよく議論の対象とされる。この収入は母集団では正規分布はせず、少数の大金持ちがいるため右側のすそ野が広い分布になることが多い。こんな母集団から標本をとり平均を検討する場合、標本サイズが小さければここで紹介した定理によって検定を進めるにはやや問題があるが、標本サイズが大きければ問題はないのである。

以上で統計的仮説検定の基本的な考え方についての解説は終わりである。次節から具体的な検定法について解説するが、おそらく読者はそこで母集団と標本について混乱することになると思う[5]。そこで、あらかじめこの問題について述べておきたい。

3.8 検定における母集団と標本

（1）社会調査データの母集団と標本

社会調査では一般にさまざまな事柄が同時に調べられる。このような調査データにおいて母集団や標本（理論的標本、現実の標本）はどのようなものになるのだろうか。

表 1　性別・統計学得点・調査法得点

ケース番号	性別	統計学	調査法
1	x_1 = 1	y_1 = 10	z_1 = 10
2	x_2 = 1	y_2 = 10	z_2 = 0
3	x_3 = 2	y_3 = 10	z_3 = 10
4	x_4 = 1	y_4 = 0	z_4 = 0
5	x_5 = 2	y_5 = 10	z_5 = 10
6	x_6 = 2	y_6 = 10	z_6 = 10
7	x_7 = 1	y_7 = 0	z_7 = 0
8	x_8 = 2	y_8 = 10	z_8 = 10
9	x_9 = 2	y_9 = 10	z_9 = 0
10	x_{10} = 2	y_{10} = 0	z_{10} = 10
11	x_{11} = 2	y_{11} = 0	z_{11} = 0

表 1 はある大学の学生に対する標本調査で明らかになった性別、統計学得点、調査法得点のデータだとする。性別は 1 が男性で 2 が女性である。

この「調査」の母集団は何であり標本は何かということに関しては、「ある大学の学生が母集団であり標本はそこから選ばれた 11 名からなる」と考えていい。しかし、「検定」まで射程に置くとき、標

[5] 平均や比率の差の検定などにおける「2 つの母集団とは何か」といったことや対応のある検定における「対応とは何か」といったことを理解する前提として、次項の内容を知っておく必要がある。

本と母集団についてはもっときちんと把握しておく必要がある。

推測統計における母集団は変数の値の集合であると前に述べた（小林,2018a）。標本についても同様で、標本は様々な属性（変数）を持ったケースの集合ではなく、ケースの属性（＝変数）の値の集合と考える必要がある。

このことを念頭に置き表 1 を見てみよう。真ん中の列には統計学のケースごとの得点が記載されている。それは次のようなものだ。

統計学得点の現実の標本

$$=\{y_1,y_2,y_3,y_4,y_5,y_6,y_7,y_8,y_9,y_{10},y_{11}\}$$
$$=\{10,10,10,0,10,10,0,10,10,0,0\}$$

実は、この得点の集合は調査データに含まれる 1 つの標本（現実の標本）と考えられるのである。すなわち、表 1 の調査データには多様な標本が含まれ、統計学得点の集合も標本の 1 つなのである。

この標本は実際の標本から得られた値を示しており、次の理論的標本の実現値ということになる。

統計学得点の理論的標本

$$=\{Y_1,Y_2,Y_3,Y_4,Y_5,Y_6,Y_7,Y_8,Y_9,Y_{10},Y_{11}\}$$

そして、この理論的標本や現実の標本の背後にある（ある大学の統計学得点の）母集団は、たとえば次のようなものである。

統計学得点の母集団

$$=\{10,0,0,10,0,10,10,0,0,0,10,10,0,\cdots\cdots\cdots\cdots\}$$

小林 2018a で示したように、これを比率の分布としてとらえた場合、これは母集団の統計学得点 Y の確率分布となる [6]。

われわれは、表 1 の統計学得点のデータをもとに母集団の統計学得点の平均を予測したりするのだが、そこで想定されている標本は表 1 のデータすべてではなく、ここに述べた統計学得点の標本である。また、その際、母集団とは統計学得点の母集団のことを意味している。

（2）調査データの中のさまざまな標本

さて、表 1 の調査データには、さまざなな標本があると述べた。統計学得点の標本だけでなくそこには性別の標本、調査法得点の標本もある。それらは次のものだ。

性別の標本

$$=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10},x_{11}\}$$
$$=\{1,1,2,1,2,2,1,2,2,2,2\}$$

調査法得点の標本

$$=\{z_1,z_2,z_3,z_4,z_5,z_6,z_7,z_8,z_9,z_{10},z_{11}\}$$
$$=\{10,0,10,0,10,10,0,10,0,10,0\}$$

では、調査データに含まれる標本はこれら 3 つだけなのかというとそうではない。われわれは分析に応じてさまざまな標本と母集団を設定するのである。

たとえば、今、データに基づいて男性の母集団での統計学の平均点を推測しようとしているとしよう。そのとき考えられている標本は次のものだ。

[6] 母集団はこのように値の集合であり、その要素の値と出現率の関係を示したものが（統計学得点 Y についての）母集団における（統計学得点 Y という）確率変数の確率分布である。本シリーズでは「Y についての母集団における Y という確率変数の確率分布」のことを省略して「Y の母集団の分布」や「母集団 Y の分布」と言うことがある。

男性統計学得点の標本
$=\{y_1, y_2, y_4, y_7\}=\{10,10,0,0\}$

すなわち、統計学得点の部分集合である男性統計学得点が標本として設定されるのである。これもまた調査データに含まれる標本であり、そこで設定される母集団は男性統計学得点の母集団である。

「標本において統計学の平均点は男女間で差があるが、母集団でも差があるといえるかどうか」ということを検討しているとしよう。そのとき考えられている標本は次の 2 つだ。

男性統計学得点の標本
$=\{y_1, y_2, y_4, y_7\}=\{10,10,0,0\}$

女性統計学得点の標本
$=\{y_3, y_5, y_6, y_8, y_9, y_{10}, y_{11}\}$
$=\{10,10,10,10,10,0,0\}$

われわれはこれらの 2 つの標本からまず得点平均に差があること確認し、次いで 2 つの母集団でも差があると言えるのかを検討する。このとき想定されている母集団は、（ある大学の）男性統計学得点の母集団と女性統計学得点の母集団である。

「統計学得点が 0 点の者と 10 点の者では調査法得点に差があると母集団でも言えるのか」ということを検討する際の標本は次の 2 つになる。

統計学0点の調査法得点の標本
$=\{z_4, z_7, z_{10}, z_{11}\}=\{0,0,10,0\}$

統計学10点の調査法得点の標本
$=\{z_1, z_2, z_3, z_5, z_6, z_8, z_9\}$
$=\{10,0,10,10,10,10,0\}$

そして、それぞれに対応する母集団が想定されるのである。

標本の値や母集団の値をそれぞれセットとして考えるような場合もある。たとえば、標本における統計学得点と調査法得点の相関から、母集団における統計学得点と調査法得点の相関を推測するようなとき、次のような 2 次元の標本が設定される。

統計学得点と調査法得点のセットの標本
$=\{(y_1, z_1), (y_2, z_2), (y_3, z_3), \ldots, (y_{11}, z_{11})\}$
$=\{(10,10), (10,0), (10,10), \ldots, (0,0)\}$

そして、この標本の背後に、セットとしての値を持つ 2 次元の母集団をわれわれは考えるのだ。

さらに、調査データをすべて含んだ次のような 3 次元の標本も考えられる [7]。

性別の値, 統計学得点, 調査法得点のセットの標本
$=\{(x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_{11}, y_{11}, z_{11})\}$
$=\{(1,10,10), (1,10,0), \ldots, (2,0,0)\}$

そしてここではこれに対応する 3 次元の母集団が想定される。

以上見てきたように、表 1 のような社会調査のデータはその内部に多様な標本を含み、検定や推定においてはそれに対応する母集団が想定されるのである。次節からの具体的な検定法の解説を読む際には、母集団と標本がこのような意味を持っ

[7] 本シリーズで登場するのは 2 次元までの標本だが、多変量解析においてはこのようにセットの要素数がもっと多い多次元の標本をもとに分析がなされる。

ているということを念頭に置いてほしい。

4 母平均に関する検定

4.1 母平均の検定（母分散既知）♦

（1）目的

母分散が既知のとき、母平均μが何らかの特定の値 m ではない（より大きい、より小さい）と言いたい。

（2）定理（検定統計量とその確率分布）

母平均μ、母分散σ^2の正規母集団 $N(\mu, \sigma^2)$からのサイズ n の標本について、標本平均を$\overline{X}$とすれば、下の検定統計量は標準正規分布 N(0,1)に従う。

$$Test = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\overline{x} - \mu}{\sqrt{\sigma^2/n}}$$

（4）帰無仮説

$\mu = \mathrm{m}$

（5）対立仮説

$\mu \neq \mathrm{m}$：両側検定（両側棄却域、5%で±1.96 の外で棄却）

$\mu > \mathrm{m}$：片側検定（右側棄却域、5%で＋1.65 の外で棄却）

$\mu < \mathrm{m}$：片側検定（左側棄却域、5%で－1.65 の外で棄却）

（6）解説

μに特定の値 m を、σ^2に母分散を、$\overline{x}$に現実の標本の平均を、n に標本サイズを代入して test の値（検定統計量の値）を算出する。Test の分布（標準正規分布）における test の位置から帰無仮説が棄却できるかどうかを判断する。

この検定は前節のボールの例で述べたものである。シンプルにこの検定結果を示すと次のようになる。

（7）例

（問）分散が 36 の正規母集団からのサイズ 25 の標本において平均が 267 のとき、「母集団の平均が 270 ではない」と言えるか。有意水準 5%の両側検定で検討せよ。

（答）

H_0：母平均は 270 である（帰無仮説）。

H_1：母平均は 270 ではない（対立仮説）。

$$test = \frac{\overline{x} - \mu}{\sqrt{\sigma^2/n}} = \frac{267 - 270}{\sqrt{36/25}} = -2.5$$

$$z\left(\frac{0.05}{2}\right) = z(0.025) = 1.96$$

$$|test| > z\left(\frac{0.05}{2}\right)$$

したがって、「母平均は 270」という帰無仮説が有意水準 5%の両側検定で棄却でき、「母平均は 270 ではない」と言える。

4.2 母平均の検定（母分散未知）♦

（1）目的

母分散が未知のとき、母平均 μ が何らかの特定の値 m ではない（より大きい、より小さい）と言いたい。

（2）定理（検定統計量とその確率分布）

母平均 μ 、母分散 σ^2（未知）の正規母集団からのサイズ n の標本について、標本平均を $\overline{X}$ 、不偏分散を U^2 とすれば、下の検定統計量は自由度 n-1 の t 分布に従う。

$$Test = \frac{\overline{X} - \mu}{\sqrt{U^2/n}} \sim t_{n-1}$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\overline{x} - \mu}{\sqrt{u^2/n}}$$

（4）帰無仮説

μ ＝m

（5）対立仮説

μ ≠m：両側検定（両側棄却域）
μ ＞m：片側検定（右側棄却域）
μ ＜m：片側検定（左側棄却域）

（6）解説

特定の値 m を μ に、現実の標本の平均と不偏分散を $\overline{x}$ と u^2 に、標本サイズを n に代入して test の値（検定統計量の値）を算出する。Test の分布（t 分布）における棄却域に test があるならば、帰無仮説が棄却できる。

母平均がわからないときには母分散もわからないのが普通だ。そのときにはこの定理を用いた検定を行う。

この定理の検定統計量は前の検定統計量の母分散のかわりに標本の不偏分散を用いたものだ。母分散はわからないので、その不偏推定量である標本の不偏分散を用いるのである。不偏分散を用いたこの検定統計量は、もはや標準正規分布には従わない。それは、標準正規分布に似た左右対称の t 分布に従うことになる。

（7）例

(問)正規母集団からのサイズ 25 の標本において、平均が 267、不偏分散が 36 のとき、「母集団の平均が 270 ではない」と言えるか。有意水準 5%の両側検定で検討せよ。

（答）
H_0：母平均は 270 である（帰無仮説）。
H_1：母平均は 270 ではない（対立仮説）。

$$test = \frac{\overline{x} - \mu}{\sqrt{u^2/n}} = \frac{267-270}{\sqrt{36/25}} = -2.5$$

$$t_{24}\left(\frac{0.05}{2}\right) = t_{24}(0.025) = 2.064$$

$$|test| > t_{24}\left(\frac{0.05}{2}\right)$$

したがって、「母平均は 270」という帰無仮説が有意水準 5%の両側検定で棄却でき、「母集団は 270 ではない」と言える。

（8）補足

なぜこの定理が成り立つのかということについて気になる読者もいると思うので補足しておく。

まず、不偏分散を含んだ次の式が自由度 n－1 の

χ^2分布に従うことを明らかにする。

■目標

$$\frac{(n-1)U^2}{\sigma^2} \sim \chi^2_{n-1}$$

■導き方

$$\sum\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\sum(X_i-\mu)^2$$

$$= \frac{1}{\sigma^2}\sum(X_i-\bar{X}+\bar{X}-\mu)^2$$

$$= \frac{1}{\sigma^2}\sum(X_i-\bar{X})^2 + \frac{2}{\sigma^2}(\bar{X}-\mu)\sum(X_i-\bar{X})$$

$$+\frac{n}{\sigma^2}(\bar{X}-\mu)^2$$

$$= \frac{1}{\sigma^2}\sum(X_i-\bar{X})^2 + \frac{2}{\sigma^2}(\bar{X}-\mu)\left(\sum X_i-\sum\bar{X}\right)$$

$$+\frac{n}{\sigma^2}(\bar{X}-\mu)^2$$

$$= \frac{1}{\sigma^2}\sum(X_i-\bar{X})^2 + \frac{2}{\sigma^2}(\bar{X}-\mu)(n\bar{X}-n\bar{X})$$

$$+\frac{n}{\sigma^2}(\bar{X}-\mu)^2$$

$$= \frac{1}{\sigma^2}\sum(X_i-\bar{X})^2 + \left(\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right)^2$$

$$\therefore \sum\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{1}{\sigma^2}\sum(X_i-\bar{X})^2 + \left(\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right)^2$$

ここで左辺は標準正規分布に従う変数の二乗の n 個の和だから自由度 n の χ^2 分布に従う。右辺第 2 項は標準正規分布に従う変数の二乗だから自由度 1 の χ^2 分布に従う。したがって、右辺第 1 項は、χ^2分布の再生性より自由度 n－1 の χ^2 分布に従うことになる[8]。

ここで、右辺第 1 項は目標で示した次の式と同じである。

$$\frac{(n-1)U^2}{\sigma^2}$$

というのは

$$\frac{1}{\sigma^2}\sum(X_i-\bar{X})^2 = \frac{n-1}{\sigma^2(n-1)}\sum(X_i-\bar{X})^2$$

$$= \frac{(n-1)U^2}{\sigma^2}$$

だからである。したがって、下の式は自由度 n－1 の χ^2 分布に従う。

$$\frac{(n-1)U^2}{\sigma^2} \sim \chi^2_{n-1}$$

■検定統計量が自由度 n－1 の t 分布に従うわけ

上のことを押さえたうえで、検定統計量についての式を変形していくと次のようになる。

$$Test = \frac{\bar{X}-\mu}{\sqrt{U^2/n}} = \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\cdot\frac{\sqrt{\sigma^2}}{\sqrt{U^2}} = \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\Big/\sqrt{\frac{U^2}{\sigma^2}}$$

$$= \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\Bigg/\sqrt{\frac{(n-1)U^2}{\sigma^2}\Big/(n-1)}$$

[8] 左辺についての説明からわかるように、ここで明らかにしようとしていること、すなわち「$(n-1)U^2/\sigma^2$ が自由度 n-1 の χ^2 分布に従う」ことは、母集団が正規分布に従うことを前提としている。それゆえここでの定理が成り立つためには母集団の正規性が必要になる。このことは、よく使われる 4.4 の t 検定においても言える。

なお、ここでの議論は確率変数相互の独立性の検討を含んでいないので厳密な証明とは言えない。以下の「補足」でもそのような検討を省略し「基本的な考え方」を中心に解説しているので注意してほしい。

ここで、

$$\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$
$$\frac{(n-1)\,U^2}{\sigma^2} \sim \chi^2_{n-1}$$

だから、検定統計量 Test の式は「確率分布の相互関係」で用いたシンボルで表現すると次の形になっている。

$$\frac{Z}{\sqrt{K_{n-1}/(n-1)}} = T_{n-1}$$

したがって、検定統計量は自由度 n－1 の t 分布に従うことになる。

4.3 母平均の差の検定（母分散既知）♦

（1）目的

2 つの母集団において母分散が既知であるとき、母平均に差があると言いたい[9]。

（2）定理（検定統計量とその確率分布）

2 つの正規母集団 $N(\mu_A, \sigma_A^2)$、$N(\mu_B, \sigma_B^2)$よりサイズ n_A と n_B の標本を取り出したとき、2 つの母平均 μ_A と μ_B が等しいならば、下の検定統計量は標準正規分布 N(0,1)に従う。$\overline{X}_A$ と $\overline{X}_B$ は標本平均。

$$Test = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}}} \sim N(0,1)$$

[9] 2 つの母集団や 2 つの標本の意味については、3.8 を参照。

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}}}$$

（4）帰無仮説

$\mu_A = \mu_B$

（5）対立仮説

$\mu_A \neq \mu_B$：両側検定（両側棄却域）
$\mu_A > \mu_B$：片側検定（右側棄却域）
$\mu_A < \mu_B$：片側検定（左側棄却域）

（6）解説

2 つの母分散、ならびに現実の標本の 2 つ平均と標本サイズとをもとに test の値を算出し、その値の Test の分布（標準正規分布）における位置から帰無仮説が棄却できるかどうかを検討する。

2 つの母分散がわかっているとき、この定理を用いて平均の差の検定ができるのだが、母平均はわからないのに母分散はわかっているという状況は現実にはあまりない。したがって、この検定の利用頻度は多いとはいえない。しかし、この検定は平均の差の検定の基本的な考え方を表しているので重要だ。母分散が不明なときは、次に説明する t 検定などを用いる。

（7）例

（問）ある会社の常勤職員 22 名とパート職員 20 名について、職業満足度テストを行ったところ平均はそれぞれ 14.12 と 9.96 だった。母分散はそれぞれ 18.04 と 13.34 であることが分かっているとして、

母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言えるか。有意水準1%の両側検定で検討せよ。

（答）

H_0：常勤職員とパート職員で職業満足度テストの得点に差はない（帰無仮説）。

H_1：常勤職員とパート職員で職業満足度テストの得点に差がある（対立仮説）。

$$test = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}}} = \frac{14.12 - 9.96}{\sqrt{\frac{18.04}{22} + \frac{13.34}{20}}} = 3.41$$

$$z\left(\frac{0.01}{2}\right) = z(0.005) = 2.5758$$

$$|test| > z\left(\frac{0.01}{2}\right)$$

したがって、「差がない」という帰無仮説が有意水準１％の両側検定で棄却でき、母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言える。

（8）補足

母平均と標本平均の関係、ならびに正規分布の再生性についての知識があればこの定理は理解できる。

$$X_A \sim N\left(\mu_A, \sigma_A{}^2\right)$$
$$X_B \sim N\left(\mu_B, \sigma_B{}^2\right)$$

このとき、母平均と標本平均の関係から、

$$\bar{X}_A \sim N\left(\mu_A, \frac{\sigma_A{}^2}{n_A}\right)$$

$$\bar{X}_B \sim N\left(\mu_B, \frac{\sigma_B{}^2}{n_B}\right)$$

ここで、正規分布の再生性より、

$$\bar{X}_A - \bar{X}_B \sim N\left(\mu_A - \mu_B, \frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}\right)$$

さらに標準化して、

$$\frac{\left(\bar{X}_A - \bar{X}_B\right) - \left(\mu_A - \mu_B\right)}{\sqrt{\frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}}} \sim N(0, 1)$$

ここで、$\mu_A = \mu_B$ ならば、

$$Test = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{\sigma_A{}^2}{n_A} + \frac{\sigma_B{}^2}{n_B}}} \sim N(0, 1)$$

これが検定統計量である。

4.4 母平均の差の検定（母分散未知だが等しい）◆

（1）目的

2 つの母集団において、母分散が未知ではあるが等しいと考えられる場合、母平均に差があると言いたい。

（2）定理（検定統計量とその確率分布）

2 つの正規母集団 $N(\mu_A, \sigma^2)$、$N(\mu_B, \sigma^2)$ よりサイズ n_A と n_B の標本を取り出したとき、2 つの母平均 μ_A と μ_B が等しいならば、下の検定統計量は自由度 $n_A + n_B - 2$ の t 分布に従う。$\bar{X}_A$ と $\bar{X}_B$ は標本平

均、U_A^2 と U_B^2 は不偏分散、U^{2*}は合併分散と呼ばれるもの。

$$Test = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{U^{2*}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} \sim t_{n_A+n_B-2}$$

$$\left(U^{2*} = \frac{(n_A-1)U_A{}^2 + (n_B-1)U_B{}^2}{n_A+n_B-2} \right)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{u^{2*}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$

$$\left(u^{2*} = \frac{(n_A-1)u_A{}^2 + (n_B-1)u_B{}^2}{n_A+n_B-2} \right)$$

（4）帰無仮説

$\mu_A = \mu_B$

（5）対立仮説

$\mu_A \neq \mu_B$：両側検定（両側棄却域）
$\mu_A > \mu_B$：片側検定（右側棄却域）
$\mu_A < \mu_B$：片側検定（左側棄却域）

（6）解説

現実の標本における 2 つの不偏分散ならびに 2 つの標本サイズをもとに合併分散 u*を算出する。この合併分散と、標本平均、標本サイズから test の値を算出し、その値の Test の分布（t 分布）における位置から帰無仮説が棄却できるかどうかを検討する。

この検定は t 検定（スチューデントの t 検定）と呼ばれるよく利用される検定法である。標本サイズが大きく異なり（2 倍程度が限界）、分散も大きく異なるとこの方法は使えない。その時は、次項のウェルチの検定を行う。したがって検定の前に等分散性の検定を行うことがある。

（7）例

（問）ある会社の常勤職員 22 名とパート職員 20 名について、職業満足度テストを行ったところ平均はそれぞれ 14.12 と 9.96、不偏分散は 18.04 と 13.34 であった。母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言えるか。正規母集団であることならびに母分散は等しいことを仮定して有意水準１％の両側検定で検討せよ。

（答）

H_0：常勤職員とパート職員で職業満足度テストの得点に差はない（帰無仮説）。

H_1：常勤職員とパート職員で職業満足度テストの得点に差がある（対立仮説）。

$$test = \frac{14.12-9.96}{\sqrt{\frac{(22-1)18.04+(20-1)13.34}{22+20-2}}\sqrt{\frac{1}{22}+\frac{1}{20}}}$$

$$= 3.39$$

$$t_{40}\left(\frac{0.01}{2}\right) = t_{40}(0.005) = 2.704$$

$$|test| > t_{40}\left(\frac{0.01}{2}\right)$$

したがって、「差がない」という帰無仮説が有意水準１％の両側検定で棄却でき、母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言える。

（8）補足

検定統計量について少し解説しておく。分散が

既知の場合の検定統計量の式は次の式だった。

$$\frac{\bar{X}_A-\bar{X}_B}{\sqrt{\frac{\sigma_A{}^2}{n_A}+\frac{\sigma_B{}^2}{n_B}}}$$

ここで母分散が等しいと仮定すると次のようになる。

$$\frac{\bar{X}_A-\bar{X}_B}{\sqrt{\sigma^2\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$

この母分散がわからないので次の合併分散 U^{2*} というものを考え、それをこの検定では用いているのである（$U_A{}^2$ と $U_B{}^2$ はそれぞれの不偏分散）。

$$U^{2*}=\frac{\sum_{i=1}^{nA}\left(X_{Ai}-\bar{X}_A\right)^2+\sum_{j=1}^{nB}\left(X_{Bj}-\bar{X}_B\right)^2}{\left(n_A-1\right)+\left(n_B-1\right)}$$
$$=\frac{\left(n_A-1\right)U_A^2+\left(n_B-1\right)U_B^2}{n_A+n_B-2}$$

なぜこのような得体のしれない合併分散 U^{2*} なるものを用いるのかというと、それが 2 つの母集団で等しい母分散 σ^2 の不偏推定量だからである。

■準備

$$E\left(\Sigma\left(X_i-\bar{X}\right)^2\right)=E\left(\left(n-1\right)U^2\right)$$
$$=\left(n-1\right)E\left(U^2\right)=\left(n-1\right)\sigma^2$$

■U^{2*} が σ^2 の不偏推定量になるわけ

$$E\left(U^{2*}\right)=E\left(\frac{\sum_{i=1}^{nA}\left(X_{Ai}-\bar{X}_A\right)^2+\sum_{j=1}^{nB}\left(X_{Bj}-\bar{X}_B\right)^2}{n_A+n_B-2}\right)$$
$$=\frac{E\left(\sum_{i=1}^{nA}\left(X_{Ai}-\bar{X}_A\right)^2\right)+E\left(\sum_{j=1}^{nB}\left(X_{Bj}-\bar{X}_B\right)^2\right)}{n_A+n_B-2}$$

ここで「準備」をあてはめると（2 つの母分散は等しいことに注意）、

$$=\frac{\left(n_A-1\right)\sigma^2+\left(n_B-1\right)\sigma^2}{n_A+n_B-2}=\sigma^2$$

検定統計量が自由度（n_A+n_B-2）の t 分布に従うことについても明らかにしておこう。

■Test が自由度（n_A+n_B-2）の t 分布に従うわけ

$$Test=\frac{\bar{X}_A-\bar{X}_B}{\sqrt{U^{2*}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$
$$=\frac{\bar{X}_A-\bar{X}_B}{\sqrt{\frac{\left(n_A-1\right)U_A{}^2+\left(n_B-1\right)U_B{}^2}{n_A+n_B-2}}\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}}$$

分子分母を $\sqrt{\sigma^2\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}$ で割ると、

$$=\frac{\bar{X}_A-\bar{X}_B\Big/\sqrt{\sigma^2\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}{\sqrt{\frac{\left(n_A-1\right)U_A{}^2+\left(n_B-1\right)U_B{}^2}{n_A+n_B-2}}\Big/\sqrt{\sigma^2}}$$

$$=\frac{\bar{X}_A-\bar{X}_B\Big/\sqrt{\dfrac{\sigma^2}{n_A}+\dfrac{\sigma^2}{n_B}}}{\sqrt{\dfrac{(n_A-1)U_A{}^2+(n_B-1)U_B{}^2}{\sigma^2}}\sqrt{\dfrac{1}{n_A+n_B-2}}}$$

$$=\frac{\bar{X}_A-\bar{X}_B\Big/\sqrt{\dfrac{\sigma^2}{n_A}+\dfrac{\sigma^2}{n_B}}}{\sqrt{\dfrac{(n_A-1)U_A{}^2}{\sigma^2}+\dfrac{(n_B-1)U_B{}^2}{\sigma^2}}\Big/\sqrt{n_A+n_B-2}}$$

ここで分子は標準正規分布に従い（$\mu_1-\mu_2=0$ だから）、分母左根号内の左項は自由度 n_A-1 の χ^2 分布に、右項は自由度 n_B-1 の χ^2 は分布に従う（4.3（8）参照）。χ^2 は分布には再生性があるので、左根号内全体は自由度 n_A+n_B-2 の χ^2 は分布に従うことになる。このことからこの式は「確率分布の相互関係」で用いたシンボルで表現すると次のような形になっていることがわかる。

$$\frac{Z}{\sqrt{K_{nA+nB-2}/(n_A+n_B-2)}}=T_{nA+nB-2}$$

したがって、自由度（n_A+n_B-2）の t 分布に従うことになる。

4.5 母平均の差の検定（ウェルチの検定）

（1）目的

2 つの母集団において、母分散が未知であり、等しいとも考えられない場合、母平均に差があると言いたい。

（2）定理（検定統計量とその確率分布）

2 つの正規母集団 $N(\mu_A,\sigma^2)$、$N(\mu_B,\sigma^2)$ よりサイズ n_A と n_B の標本を取り出したとき、2 つの母平均 μ_A と μ_B が等しいならば、下の検定統計量は下の自由度の t 分布に従う。$\bar{X}_A$ と $\bar{X}_B$ は標本平均、$U_A{}^2$ と $U_B{}^2$ は不偏分散。

$$Test=\frac{\bar{X}_A-\bar{X}_B}{\sqrt{\dfrac{U_A{}^2}{n_A}+\dfrac{U_B{}^2}{n_B}}}\sim t_{df}$$

この検定統計量の式は、4.3 の母分散既知の場合の式の母分散を不偏分散に変えたものだ。自由度は次の式に最も近い整数である。

$$df=\frac{\left(\dfrac{U_A{}^2}{n_A}+\dfrac{U_B{}^2}{n_B}\right)^2}{\dfrac{(U_A{}^2/n_A)^2}{n_A-1}+\dfrac{(U_B{}^2/n_B)^2}{n_B-1}}$$

（3）現実の標本を参照した検定統計量の値

$$test=\frac{\bar{x}_A-\bar{x}_B}{\sqrt{\dfrac{u_A{}^2}{n_A}+\dfrac{u_B{}^2}{n_B}}}$$

（4）帰無仮説

$\mu_A=\mu_B$

（5）対立仮説

$\mu_A\neq\mu_B$：両側検定
$\mu_A>\mu_B$：片側検定
$\mu_A<\mu_B$：片側検定

（6）解説

現実の標本における 2 つの平均、2 つの不偏分散ならびに 2 つの標本サイズをもとに test の値を算出し、その値の Test の分布（t 分布）における位置

から帰無仮説が棄却できるかどうかを検討する。

前項の t 検定の改良版である。t 検定とウェルチの検定のどちらを行うかは分散が同一かどうかによる。したがって検定の前に等分散性の検定を行うことがある（5.2 参照）[10]。ウェルチの検定も t 分布を用いるので t 検定（ウェルチの t 検定）と呼ばれることがある。

（7）例

（問）ある会社の常勤職員 22 名とパート職員 20 名について、職業満足度テストを行ったところ平均はそれぞれ 14.12 と 9.96、不偏分散は 18.04 と 13.34 であった。母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言えるか。ウェルチの方法を用いて有意水準１％の両側検定で検討せよ。

（答）

H_0：常勤職員とパート職員で職業満足度テストの得点に差はない（帰無仮説）。

H_1：常勤職員とパート職員で職業満足度テストの得点に差がある（対立仮説）。

$$test=\frac{\overline{x}_A-\overline{x}_B}{\sqrt{\frac{u_A{}^2}{n_A}+\frac{u_B{}^2}{n_B}}}=\frac{14.12-9.96}{\sqrt{\frac{18.04}{22}+\frac{13.34}{20}}}=3.41$$

$$df=\frac{\left(\frac{u_A{}^2}{n_A}+\frac{u_B{}^2}{n_B}\right)^2}{\frac{\left(u_A{}^2/n_A\right)^2}{n_A-1}+\frac{\left(u_B{}^2/n_B\right)^2}{n_B-1}}$$

$$=\frac{\left(\frac{18.04}{22}+\frac{13.34}{20}\right)^2}{\frac{\left(18.04/22\right)^2}{22-1}+\frac{\left(13.34/20\right)^2}{20-1}}=39.88$$

$$t_{40}\left(\frac{0.01}{2}\right)=t_{40}\left(0.005\right)=2.704$$

$$|test|>t_{40}\left(\frac{0.01}{2}\right)$$

したがって、「差がない」という帰無仮説が有意水準１％の両側検定で棄却でき、母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言える。

（8）補足

この定理の証明は推測統計の入門レベルを超えている。

4.6 母平均の差の検定（大標本）♦

（1）目的

大標本を用いて、2 つの母集団において母平均に差があると言いたい。

---

[10] スチューデントの t 検定は分散が等しい場合に用いられるが、ウェルチの検定は分散が等しくても等しくなくても利用できる。だったらいつもウェルチの検定を行えばいいということになりそうだが、検定を行う者にとって複雑な自由度は言わばブラックボックスなのでできれば避けたいという気分になる。スチューデントの t 検定は 2 つの標本のサイズが大きく異なっていない限り（おおよそ 2 倍以内）、分散がある程度異なっていても利用できるとされているのでわかりやすい t 検定をしたくなるのである。ただ、t 検定は母集団の正規性を前提としている。その意味で個人的には 4.6 の大標本での検定がいいと思うのだが、統計ソフトでは t 検定やウェルチの検定がデフォルトになっていることが多い。

（2）定理（検定統計量とその確率分布）

2 つの正規母集団 N(μ_A, σ^2)、N(μ_B, σ^2)よりサイズ n_A と n_B の大標本を取り出したとき、2 つの母平均 μ_A と μ_B が等しいならば、下の検定統計量は標準正規分布に近似的に従う。$\overline{X}_A$ と $\overline{X}_B$ は標本平均、$S_A{}^2$ と $S_B{}^2$ は標本分散。

$$Test = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\frac{S_A{}^2}{n_A} + \frac{S_B{}^2}{n_B}}} \sim N(0,1)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\frac{s_A{}^2}{n_A} + \frac{s_B{}^2}{n_B}}}$$

（4）帰無仮説

$\mu_A = \mu_B$

（5）対立仮説

$\mu_A \neq \mu_B$：両側検定（両側棄却域）
$\mu_A > \mu_B$：片側検定（右側棄却域）
$\mu_A < \mu_B$：片側検定（左側棄却域）

（6）解説

現実の標本における 2 つの平均、2 つの分散ならびに 2 つの標本サイズをもとに test の値を算出し、その値の Test の分布（標準正規分布）における位置から帰無仮説が棄却できるかどうかを検討する。

大標本のときに使える検定法である（安田・原, 1982）。検定統計量の式は、4.1 の母分散が既知の場合の検定統計量の式の母分散を標本分散に変えたものだ。大標本だから標本分散を用いても不偏分散を用いても結果はほとんどかわらない。不偏分散を用いると、この検定統計量の式はウェルチの検定における検定統計量の式と同じになる[11]。

（7）例

（問）ある会社の常勤職員 220 名とパート職員 200 名について、職業満足度テストを行ったところ平均はそれぞれ 14.12 と 9.96、標本分散は 18.04 と 13.34 であった。母集団においても常勤職員とパート職員で職業満足度テストに得点差があると言えるか。大標本と考え、有意水準 1 ％の両側検定で検討せよ。

（答）

H_0：常勤職員とパート職員で職業満足度テストの得点に差はない（帰無仮説）。

H_1：常勤職員とパート職員で職業満足度テストの得点に差がある（対立仮説）。

$$test = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\frac{s_A{}^2}{n_A} + \frac{s_B{}^2}{n_B}}} = \frac{14.12 - 9.96}{\sqrt{\frac{18.04}{220} + \frac{13.34}{200}}} = 10.79$$

$$z\left(\frac{0.01}{2}\right) = z(0.005) = 2.5758$$

$$|test| > z\left(\frac{0.01}{2}\right)$$

したがって、「差がない」という帰無仮説が有意水準 1 ％の両側検定で棄却でき、母集団においても常勤職員とパート職員で職業満足度テストに得点

[11] 大標本についての検定だから、t 検定より使えない検定法かというとそういうわけでもない。t 検定は母集団の正規性が前提とされるが、この検定では正規性は強く要求されるわけではない。それゆえ利用できる範囲は t 検定よりも広いとも考えられる。

差があると言える。

5 母分散に関する検定

5.1 母分散の検定

（1）目的

母分散 σ^2 が何らかの特定の値 q^2 ではない（より大きい、より小さい）と言いたい。

（2）定理（検定統計量とその確率分布）

正規母集団 $N(\mu, \sigma^2)$ よりサイズ n の標本を取り出したとき、下の検定統計量は自由度 n－1 の χ^2 分布に従う。ただし、U^2 は不偏分散。

$$Test = \frac{(n-1)U^2}{\sigma^2} \sim \chi^2_{n-1}$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{(n-1)u^2}{\sigma^2}$$

（4）帰無仮説

$\sigma^2 = q^2$

（5）対立仮説

$\sigma^2 \neq q^2$：両側検定（両側棄却域）
$\sigma^2 > q^2$：片側検定（右側棄却域）
$\sigma^2 < q^2$：片側検定（左側棄却域）

（6）解説

test の σ^2 に特定の値 q^2 を、u^2 に現実の標本から得られた不偏分散を、n に標本サイズを代入して検定統計量の値を算出し、その値の Test の分布（χ^2 分布）における位置から帰無仮説が棄却できるかどうかを検討する。

（7）例

（問）正規母集団から復元抽出されたサイズ 20 の標本の不偏分散が 20 だった。母分散は 25 ではないと言えるかどうか、有意水準 5%の両側検定をせよ。

（答）
H_0：母分散は 25 である（帰無仮説）。
H_1：母分散は 25 ではない（対立仮説）。

$$test = \frac{(n-1)u^2}{\sigma^2} = \frac{(20-1)20}{25} = 15.2$$
$$\chi^2_{19}(0.025) = 32.852$$
$$\chi^2_{19}(0.975) = 8.907$$
$$\chi^2_{19}(0.975) < test < \chi^2_{19}(0.025)$$

したがって、「母分散は 25 である」という帰無仮説は有意水準 5%の両側検定で棄却できず、母分散は 25 でないとは言えない。

（8）補足

定理がなぜ成り立つのかということについては、4.2（8）を参照のこと。

5.2 母分散の比の検定（等分散性の検定）

（1）目的

2 つの母分散 σ_A^2 と σ_B^2 が等しいかどうか判定したい。

（2）定理（検定統計量とその確率分布）

2 つの正規母集団 N(μ_A, σ_A^2)、N(μ_B, σ_B^2)よりサイズ n_A と n_B の標本を取り出したとき、2 つの母分散が等しいならば、下の検定統計量は自由度（n_A－1, n_B－1）の F 分布に従う。ただし、U_A^2 と U_B^2 は不偏分散。

$$Test = \frac{{U_A}^2}{{U_B}^2} \sim F_{n_B-1}^{n_A-1}$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{{u_A}^2}{{u_B}^2} \qquad \left(u_A^2 \geq u_B^2\right)$$

分散の大きい方を分子にする。

（4）帰無仮説

$\sigma_A^2 = \sigma_B^2$

（5）対立仮説

$\sigma_A^2 \neq \sigma_B^2$：両側検定（5%両側検定の場合、上側 F(0.025)以上を棄却域とする）
$\sigma_A^2 > \sigma_B^2$：片側検定（5%片側検定の場合、上側 F(0.05)以上を棄却域とする）
両側検定でも上側棄却域だけを見ればいい。

（6）解説

現実の標本から得られた 2 つの不偏分散をもとに test の値を計算する。このとき、必ず分子を分母より大きくする。test の Test の分布における位置から帰無仮説を棄却するかどうか決定する。

この検定は等平均仮説の検定の準備として行われることが多い。帰無仮説が棄却されなければ、一応、母分散は等しいと考えて等平均仮説の検定に向かうことが一般的である。「棄却されなくても帰無仮説が正しいとはいえない」というこれまでの話と少しずれるのだが、母分散が異なると言えない限り母分散は大きく異ならないと考えて t 検定をすることになるのである。t 検定は標本サイズに大きな差がない限り（2 倍まで）、分散が同じであることをそれほどきびしく要求するものではないと言われている。

（7）例

（問）ある会社の常勤職員 22 名とパート職員 20 名について、職業満足度テストを行ったところ平均はそれぞれ 14.12 と 9.96、不偏分散は 18.04 と 13.34 であった。母分散は等しいと仮定して平均の差の検定を進めていいか検討せよ。

（答）
H_0：常勤職員とパート職員で職業満足度テストの母分散は同じ（帰無仮説）。
H_1：常勤職員とパート職員で職業満足度テストの母分散は異なる（対立仮説）。

$$test = \frac{18.04}{13.34} = 1.352$$

$$2.452 < F_{19}^{21}\left(0.025\right) < 2.509$$

（数表に F(21,19)がないのでこのようにする）

$$test < F_{19}^{21}\left(0.025\right)$$

母分散は同じであるという帰無仮説は棄却できず、母分散が異なるとは言えない（有意水準 5%両側）。したがって、母分散は等しい（大きくは異ならない）と考えて平均の差の検定を進めていい。

（8）補足

定理について説明する。次のことが成り立つ（4.2（8）参照）。

$$\frac{(n_A-1)U_A^2}{\sigma_A^2} \sim \chi^2_{n_A-1}$$
$$\frac{(n_B-1)U_B^2}{\sigma_B^2} \sim \chi^2_{n_B-1}$$

したがって、次の式はF分布に従う。

$$\frac{\left(\frac{(n_A-1)U_A^2}{\sigma_A^2}\right)\Big/(n_A-1)}{\left(\frac{(n_B-1)U_B^2}{\sigma_B^2}\right)\Big/(n_B-1)} \sim F^{n_A-1}_{n_B-1}$$

というのは、「確率分布の相互関係」で用いたシンボルで表現すると次の形になっているからだ。

$$\frac{K_m/m}{K_n/n} = F_n^m$$

ここで、母分散が等しいとき、

$$\sigma_A^2 = \sigma_B^2$$

これも用いて分子分母を整理すると、

$$\frac{\left(\frac{(n_A-1)U_A^2}{\sigma_A^2}\right)\Big/(n_A-1)}{\left(\frac{(n_B-1)U_B^2}{\sigma_B^2}\right)\Big/(n_B-1)} = \frac{U_A^2}{U_B^2}$$

したがって、

$$\frac{U_A^2}{U_B^2} \sim F^{n_A-1}_{n_B-1}$$

このように、検定統計量はF分布に従う。

ところで、「分子を分母に比べてつねに大きくして検定統計量を計算する」というこの検定では上側 F(0.025)以上を棄却域としたときに、5%両側検定となる。その理由を説明しておこう。

上側確率をαとする場合、F分布の「もとの分布」とその「自由度を逆転させた分布」の間には次のような関係がある。

$$F_n^m(\alpha) = \frac{1}{F_m^n(1-\alpha)}$$

たとえば、下側確率 0.025（＝上側確率 0.975）については次のようになる。

$$F_n^m(0.975) = \frac{1}{F_m^n(0.025)}$$

ここから、ある値 x（x は正）が左辺の値より小さいということと、その値の逆数（1/ x）が右辺の分母より大きいということは同じであることがわかる。つまり、

$$x < F_n^m(0.975) \Leftrightarrow x < \frac{1}{F_m^n(0.025)}$$
$$\Leftrightarrow \frac{1}{x} > F_m^n(0.025)$$

したがって、ある値が F^m_n(0.975)より小さいことと、その値の逆数が自由度を入れ替えた F^n_m (0.025)より大きいこととは同じなのである。

以上のことを押さえて、5%の両側検定がどう行

われるのかを考えよう。2 つの正規母集団 N(μ_A, σ_A^2)、N(μ_B, σ_B^2)よりサイズ n_A+1 と n_B+1 の標本を取り出し、その不偏分散が u_A^2 と u_B^2 だったとする。この場合、$p<0.05$ の両側検定は、細かく見れば次のようになされている。

まず、分散の大きな方を分子にすると決めない場合について。これを A の方法とするとき、[A1]たまたま $u_A^2 > u_B^2$ であるならば、$u_A^2/u_B^2 > F^m_n(0.025)$ の検討がなされ、[A2]たまたま $u_A^2 < u_B^2$ であるならば、$u_A^2/u_B^2 < F^m_n(0.975)$の検討がなされる。これで、5%の両側検定がなされるのである。

次に、分散の大きな方を分子にすると決める場合について。これを B の方法とするとき、[B1]たまたま $u_A^2 > u_B^2$ であるならば、$u_A^2/u_B^2 > F^m_n(0.025)$ の検討がなされ、[B2]たまたま $u_A^2 < u_B^2$ であるならば、分子分母と自由度を逆にして $u_B^2/u_A^2 > F^n_m(0.025)$ の検討がなされる。

ところで、上で見たように、$u_A^2 < u_B^2$ の場合について、[A2]の $u_A^2/u_B^2 < F^m_n(0.975)$の検討と、[B2]の $u_B^2/u_A^2 > F^n_m(0.025)$の検討は同じことである。ここから、A の方法も B の方法も実質的には同じであることがわかる。A の方法は 5%の両側検定であるが、B の方法もまた 5%の両側検定なのである。したがって、結局、分散の大きな方を分子にすると決めて上側 F(0.025)以上を棄却域とするとき、5%両側検定を行っていることになる。

6 母比率と関連性の検定

6.1 母比率の検定◆

（1）目的

母比率 p が何らかの特定の値 m ではない（より大きい、より小さい）と言いたい。

（2）定理（検定統計量とその確率分布）

ある性質についての母比率が p であるような母集団から、サイズ n の大標本を取り出すとき、下の検定統計量は近似的に標準正規分布 N(0, 1)に従う。ただし、P’は標本比率 [12]。

$$Test = \frac{P' - p}{\sqrt{\dfrac{p(1-p)}{n}}} \sim N(0,1)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{p' - p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

（4）帰無仮説

p＝m

（5）対立仮説

p≠m：両側検定（両側棄却域）
p＞m：片側検定（右側棄却域）
p＜m：片側検定（左側棄却域）

（6）解説

p に「特定の値 m」を、p’に現実の標本の比率を、n に標本サイズを代入して test の値を算出し、この値の Test の分布（標準正規分布）における位置から帰無仮説を棄却するかどうかを検討する。

[12] 本シリーズでは小文字の p は母比率、大文字の P’は理論的な標本における確率変数としての標本比率、小文字の p’は現実の標本での標本比率を表す（大文字の P は確率を示すときに用いる）。

（7）例

（問）母集団から復元抽出されたサイズ 200 の標本において支持政党のないものの比率は 0.4 であった。母集団における支持政党なしの比率が 0.3 ではないと言えるかどうか、有意水準 5%で両側検定せよ。

（答）
H_0：母比率は 0.3 である（帰無仮説）。
H_1：母比率は 0.3 ではない（対立仮説）。

$$test = \frac{0.4 - 0.3}{\sqrt{\frac{0.3(1-0.3)}{200}}} = 3.086$$

$$z\left(\frac{0.05}{2}\right) = z(0.025) = 1.96$$

$$|test| > z\left(\frac{0.05}{2}\right)$$

したがって、「母比率は 0.3 である」という帰無仮説は有意水準 5%の両側検定で棄却でき、母比率は 0.3 ではないと言える。

（8）補足

定理については小林（2018c）で述べたが、ここでも簡単に説明しておく。今、X が二項分布に従うとする。

$$X \sim Bin(n, p)$$

その期待値と分散は np と np(1－p)である。ここで n が大きいなら、ラプラスの定理より

$$X \sim N(np, np(1-p))$$

ここで生じる回数ではなく生じる比率に焦点を置き、その比率を P’とする。P’は注目している事象の標本比率を示す確率変数だ。この確率分布は次のようになる（V(cX)＝c^2V(X)であることに注意）。

$$P' = \frac{X}{n} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

ここからさらに標準正規分布に従う確率変数を求めると次のようになる。

$$Test \frac{P' - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

これが検定統計量だ。

6.2 母比率の差の検定（2 つの母集団）

（1）目的

2 つの母集団において特定の性質をもつものの母比率に差があると言いたい。

（2）定理（検定統計量とその確率分布）

ある性質についての母比率が p_A と p_B であるような 2 つの母集団から、サイズ n_A と n_B の大標本を取り出すとき、2 つの母比率 p_A と p_B が等しいならば下の検定統計量は近似的に標準正規分布 N(0, 1)に従う。ただし、P’は標本比率、P*は合併した標本での標本比率。

$$Test = \frac{P'_A - P'_B}{\sqrt{P^*(1-P^*)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim N(0,1)$$

$$\left(P^* = \frac{n_A P'_A + n_B P'_B}{n_A + n_B}\right)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{p'_A - p'_B}{\sqrt{p^*(1-p^*)\left(\frac{1}{n_A}+\frac{1}{n_A}\right)}}$$

$$\left(p^* = \frac{n_A p'_A + n_B p'_B}{n_A + n_B}\right)$$

（4）帰無仮説

pA＝pB

（5）対立仮説

pA≠pB：両側検定（両側棄却域）
pA＞pB：片側検定（右側棄却域）
pA＜pB：片側検定（左側棄却域）

（6）解説

現実の標本の比率 p’A、p’B と標本サイズ nA、nB から合併した標本での標本比率 p*を計算する。この p*と p’A、p’B、nA、nB をもとに test の値を算出し、この値の Test の分布（標準正規分布）における位置から帰無仮説を棄却するかどうかを検討する。

（7）例

（問）ある会社の常勤職員 100 名とパート職員 120 名について、仕事に満足しているか調べたところ、満足していると答えたものはそれぞれ 61 名と 60 名で、比率にするとそれぞれ 0.61 と 0.50 であった。母集団においても常勤職員とパート職員で満足している者の比率は違うと言えるか。有意水準 5%で検討せよ。

（答）

H0：常勤職員とパート職員で満足している者の比率に差はない（帰無仮説）。

H1：常勤職員とパート職員で満足している者の比率に差がある（対立仮説）。

$$p^* = \frac{n_A p'_A + n_B p'_B}{n_A + n_B} = \frac{61+60}{100+120} = 0.55$$

$$test = \frac{p'_A - p'_B}{\sqrt{p^*(1-p^*)\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$

$$= \frac{0.61-0.50}{\sqrt{0.55(1-0.55)\left(\frac{1}{100}+\frac{1}{120}\right)}} = 1.633$$

$$z\left(\frac{0.05}{2}\right) = z(0.025) = 1.96$$

$$|test| < z\left(\frac{0.05}{2}\right)$$

したがって、「差がない」という帰無仮説は有意水準 5%の両側検定で棄却できず、母集団において常勤職員とパート職員で満足している者の比率に差があるとは言えない。

（8）補足

この母比率の差の検定と同じことは 6.5 の「独立性検定（χ^2 検定）」でもできる。一般的には後者を使うことの方が多い（補参照）。

定理について説明しておこう。

p を母比率、P’を標本比率とするとき、

$$P' \sim N\left(p, \frac{p(1-p)}{n}\right)$$

だから

$$P_A' - P_B'$$
$$\sim N\left(p_A - p_B, \frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}\right)$$

ここで母比率が等しいとすると

$$P_A' - P_B' \sim N\left(p - p, \frac{p(1-p)}{n_A} + \frac{p(1-p)}{n_B}\right)$$
$$P_A' - P_B' \sim N\left(0, p(1-p)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)\right)$$

これを標準化すると、

$$\frac{P_A' - P_B' - 0}{\sqrt{p(1-p)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} \sim N(0,1)$$

母比率 p はわからないので、合併した標本の比率 P*を使うと。

$$\frac{P_A' - P_B'}{\sqrt{P^*(1-P^*)\left(\frac{1}{n_A} + \frac{1}{n_A}\right)}} \sim N(0,1)$$

これが検定統計量である。

合併した標本の比率 P*は、2 つの母集団で等しい比率 p の不偏推定量である。

$$E(P^*) = E\left(\frac{n_A P_A' + n_B P_B'}{n_A + n_B}\right)$$
$$= E\left(\frac{n_A P_A'}{n_A + n_B}\right) + E\left(\frac{n_B P_B'}{n_A + n_B}\right)$$
$$= \frac{n_A}{n_A + n_B} E(P_A') + \frac{n_B}{n_A + n_B} E(P_B')$$
$$= \frac{n_A}{n_A + n_B} p_A + \frac{n_B}{n_A + n_B} p_B$$
$$= \frac{n_A}{n_A + n_B} p + \frac{n_B}{n_A + n_B} p = p\left(\frac{n_A + n_B}{n_A + n_B}\right) = p$$

6.3 母比率の差の検定（同一母集団内）

（1）目的

1 つの母集団においてある性質をもつものの比率と別の性質をもつものの比率は異なると言いたい。

（2）定理（検定統計量とその確率分布）

1 つの母集団において、その構成比が p_A, p_B, p_C（$p_A+p_B+p_C=1$）かつ $p_A = p_B$ であるとき、ここから抽出したサイズ n の大標本において、p_A, p_B に相当する標本比率を P_A’, P_B’ とすると、次の標本統計量は標準正規分布に近似する。

$$Test = \frac{(P_A' - P_B')}{\sqrt{\frac{(P_A' + P_B')}{n}}} \sim N(0,1)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{(p_A' - p_B')}{\sqrt{\frac{(p_A' + p_B')}{n}}}$$

分母と分子に n を掛けて次のように変形すると使いやすい。

$$test = \frac{(np_A' - np_B')}{\sqrt{(np_A' + np_B')}}$$

（4）帰無仮説

$p_A = p_B$

（5）対立仮説

$p_A \neq p_B$：両側検定（両側棄却域）
$p_A > p_B$：片側検定（右側棄却域）
$p_A < p_B$：片側検定（左側棄却域）

（6）解説

現実の標本の比率 p'_A、p'_B と標本サイズ n をもとに test の値を算出し、この値の Test の分布（標準正規分布）における位置から帰無仮説を棄却するかどうかを検討する。

下の例を見るとわかると思うが、1 つの母集団における標本の比率 P'_A と P'_B は独立でない。その時には 6.2 の検定ではなくこの検定をする。

（7）例

（問）選挙を控えたある町の 200 名についてどの候補者を支持するかを尋ねたところ、A 候補者が 60 名、B 候補者が 40 名、支持する候補者がいない人が 100 名だった。比率で言うとそれぞれ 0.3, 0.2, 0.5 である。母集団において A 候補者と B 候補者の支持率に差があるか。有意水準 5%で検討せよ。

（答）

H_0：A 候補者と B 候補者の支持率には差はない（帰無仮説）。

H_1：A 候補者と B 候補者の支持率には差がある（対立仮説）。

$$test = \frac{\left(np'_A - np'_B\right)}{\sqrt{\left(np'_A + np'_B\right)}} = \frac{60-40}{\sqrt{60+40}} = 2$$

$$z\left(\frac{0.05}{2}\right) = z\left(0.025\right) = 1.96$$

$$\left|test\right| < z\left(\frac{0.05}{2}\right)$$

したがって、「差がない」という帰無仮説は有意水準 5%の両側検定で棄却でき、母集団においても支持率に差があると言える。

（8）補足

「同一母集団内の母比率の差」の検定は 6.4 の「適合度検定」でも可能である（補参照）。

6.4 適合度検定

（1）目的

カテゴリカルなデータで実測された比率が理論的な比率と適合していないと言いたい。

（2）定理（検定統計量とその確率分布）

相互に排反する 1～k のカテゴリーがあり、それらのカテゴリーの比率が母集団において $p_1, p_2, \cdots, p_k$ だとする。また、標本においてそれぞれのカテゴリーに含まれる個体数が、$f_1, f_2, \cdots, f_k$（全部で n）だとする（表 2）。

表 2　各カテゴリーの母比率と標本個体数

カテゴリー	1	2	…	k	計
母比率	p_1	p_2	…	p_k	1
標本での個体数	f_1	f_2	…	f_k	n

そのとき、$F_i \geqq 5$ という条件の下で、下の検定統計量は、近似的に自由度(k－1)の χ^2 分布をする。ただし $F_i = p_i n$ である。

$$Test = \sum_{i=1}^{k} \left(f_i - F_i\right)^2 \Big/ F_i \sim \chi^2_{k-1}$$

（3）現実の標本を参照した検定統計量の値

$$test = \sum_{i=1}^{k} \left(f_i - F_i \right)^2 \Big/ F_i$$

検定統計量 Test の f_i は確率変数だが、これに現実の標本の値を代入したものが test である。

（4）帰無仮説

母集団の各カテゴリーの比率＝標本の各カテゴリーの比率。

（5）対立仮説

母集団と標本で比率の異なったカテゴリーがある（右側棄却域）。

（6）解説

母比率と標本サイズから計算される F_i と、現実の標本の個体数 f_i をもとに test の値を算出し、その値の Test の分布（χ^2 分布）における位置から帰無仮説が棄却できるかどうかを検討する。つねに右側検定で、有意水準は上側確率で決める。

たとえばさいころが歪んでいるかどうかを検定するときにはこの適合度検定を行う。

（7）例 1

（問）あるさいころを 60 回振ったところ、1 が 20 回、2、3、4、5 がそれぞれ 10 回出て、6 は出なかった。このさいころは歪んださいころだろうか。

（答）

H_0 : 歪んださいころではない。すなわち、標本のカテゴリー比率は正しい比率（各 1/6）を持つ母集団からの標本抽出で得られたものだ（帰無仮説）。

H_1 : 歪んださいころである。すなわち、標本のカテゴリー比率は正しい比率（各 1/6）を持つ母集団からの標本抽出で得られたものではない（対立仮説）。

$$test = \sum_{i=1}^{k} \left(f_i - F_i \right)^2 \Big/ F_i$$
$$= \left[(20-10)^2 \big/ 10 \right] + \left[(10-10)^2 \big/ 10 \right] \times 4$$
$$+ \left[(0-10)^2 \big/ 10 \right] = 20$$
$$\chi_5^2 (0.05) = 11.070$$
$$test > \chi_5^2 (0.05)$$

したがって、「標本のカテゴリー比率は正しい比率（各 1/6）を持つ母集団からの標本抽出で得られたものだ」という帰無仮説は有意水準 5%で棄却でき、「歪んださいころだ」と言える。

（8）例 2

（問）選挙を控えたある町の 100 名についてどの候補者を支持するかを尋ねたところ、A 候補者が 60 名、B 候補者が 40 名だった。母集団において A 候補者と B 候補者の支持率に差があるか。有意水準 5%で検討せよ。

（答）

H_0 : A 候補者と B 候補者の支持率には差はない（帰無仮説）。

H_1 : A 候補者と B 候補者の支持率には差がある（対立仮説）。

$$test = \sum_{i=1}^{k} \left(f_i - F_i \right)^2 \Big/ F_i$$
$$= \left[(60-50)^2 \big/ 50 \right] + \left[(40-60)^2 \big/ 50 \right] = 4$$
$$\chi_1^2 (0.05) = 3.841$$
$$test > \chi_5^2 (0.05)$$

したがって、「差がない」という帰無仮説は有意水準5%で棄却でき、母集団においても支持率に差があると言える。

(9) 補足

6.3 (7) の例での検定と上の 6.4 (8) 例2での検定は実質的に同じものである（補参照）。6.3 (7) では検定統計量の値は 2 であったが、その上側確率は0.2275であり両側確率は0.0455となる。すぐ上の例 2 では検定統計量の値は 4 であり上側確率は 0.0455 である。値相互のこのような関係を理解するためには、一方では標準正規分布をもとに検定がなされ、もう一方ではχ^2分布をもとに検定がなされていることに気づく必要がある。自由度1のχ^2分布は標準正規分布に従う確率変数の2乗の分布なので、標準正規分布で−2と2の外側の確率の合計はχ^2分布で4の上側確率になるのである。

定理についての証明は難しい[13]。k＝2 の場合についてのみ書いておく。

$$Test = \frac{(f_1 - F_1)^2}{F_1} + \frac{(f_2 - F_2)^2}{F_2}$$

$$= \frac{(f_1 - np_1)^2}{np_1} + \frac{(f_2 - np_2)^2}{np_2}$$

$$= \frac{(f_1 - np_1)^2}{np_1} + \frac{((n - f_1) - n(1 - p_1))^2}{n(1 - p_1)}$$

$$= \frac{(f_1 - np_1)^2}{np_1} + \frac{(n - f_1 - n - np_1)^2}{n(1 - p_1)}$$

$$\frac{(f_1 - np_1)^2}{np_1} + \frac{(-f_1 + np_1)^2}{n(1 - p_1)}$$

$$= \frac{(1 - p_1)(f_1 - np_1)^2 - p_1(f_1 - np_1)^2}{np_1(1 - p_1)}$$

$$= \frac{(f_1 - np_1)^2}{np_1(1 - p_1)}$$

さて、ここでいったん頭を切り替え f_1 は母集団からのサイズnの標本でカテゴリー1という注目属性をもつものの数であることに気づこう。そうすると、f_1 は二項分布に従う変数だということがわかる。

$$f_1 \sim Bin(n, p_1)$$

このとき、f_1 の期待値は np_1、分散は $np_1(1-p_1)$である。nが大きいときには、二項分布は正規分布に近似できるので、下のようになる。

$$f_1 \sim N(np_1, np_1(1 - p_1))$$

さらに、

$$\frac{f_1 - np_1}{\sqrt{np_1(1 - p_1)}} \sim N(0, 1)$$

標準正規分布に従う確率変数 m 個の二乗和の分布が自由度 m のχ^2分布である。したがって、上の式を2乗すると、それは自由度1のχ^2分布に従うことになる。

$$\frac{(f_1 - np_1)^2}{np_1(1 - p_1)} \sim \chi_1^2$$

これは、上で見た検定統計量の証明の最後の式と同じである。したがって、k＝2 のとき、検定統計量は自由度 k−1（＝1）のχ^2分布に従う。

---

[13] 証明は稲垣（2003）など参照。

$$Test = \frac{\left(f_1 - F_1\right)^2}{F_1} + \frac{\left(f_2 - F_2\right)^2}{F_2}$$
$$= \frac{\left(f_1 - np_1\right)^2}{np_1\left(1 - p_1\right)} \sim \chi_1^2$$

6.5 独立性検定（χ^2検定）♦

（1）目的

カテゴリカルな値をとる 2 変数が関連していると言いたい。

（2）定理（検定統計量とその確率分布）

表 3 のような s×t 表からなる標本があるとする（f は個体数）。

表 3　s×t 表

	b_1	b_2	$\cdots$	b_j	$\cdots$	b_t	計
a_1	f_{11}	f_{12}	$\cdots$	f_{1j}	$\cdots$	f_{1t}	$T_{1\cdot}$
a_2	f_{21}	f_{22}	$\cdots$	f_{2j}	$\cdots$	f_{2t}	$T_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
a_i	f_{i1}	f_{i2}	$\cdots$	f_{ij}		f_{it}	$T_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
a_s	f_{s1}	f_{s2}	$\cdots$	f_{sj}	$\cdots$	f_{st}	$T_{s\cdot}$
計	$T_{\cdot 1}$	$T_{\cdot 2}$		$T_{\cdot j}$		$T_{\cdot t}$	T

このとき、母集団において a_1～a_s のカテゴリカルな値をとる変数 A と b_1～b_t のカテゴリカルな値をとる変数 B の関係が独立（関連がない）ならば、$F_{ij} \geqq 5$ という条件の下で、下の検定統計量は、近似的に自由度(s－1)(t－1)の χ^2 分布をする。ただし $F_{ij} = T_{i\cdot} T_{\cdot j} / T$ である。

$$Test = \sum_1^s \sum_1^t \left(f_{ij} - F_{ij}\right)^2 / F_{ij} \sim \chi^2_{(s-1)(t-1)}$$

（3）現実の標本を参照した検定統計量の値

$$test = \sum_1^s \sum_1^t \left(f_{ij} - F_{ij}\right)^2 / F_{ij}$$

検定統計量 Test の f_{ij} は確率変数だが、これに現実の標本の値を代入したものが test である。

（4）帰無仮説

母集団で A と B には関係がない。

（5）対立仮説

母集団で A と B には関係がある（右側棄却域）。

（6）解説

現実の標本をもとに test の値を算出し、その値の Test の分布（χ^2 分布）における位置から帰無仮説が棄却できるかどうかを検討する。F_{ij} は、A と B が独立している場合の期待度数と見ることができる。常に右側検定で、有意水準は上側確率で決める。

前項の適合度検定も χ^2 検定と言うことがあるが、一般にはこの検定のことを χ^2 検定と言う。

（7）例 1

（問）ある村のある期間における小学校卒業生について、その小学校成績と、離村しているか残留しているかとの関係を調べたところ表 4 のようであった。

表 4　小学校成績と離村・残留

小学校成績	離村者	残留者	計
A	20	13	33
B	18	12	30
C	29	11	40
計	67	36	103

小学校成績と離村するかどうかということの間に関連があるか。有意水準 5%で検討せよ（安田・原,

1982:255 改変）。

（答）

H_0：小学校成績と離村するかどうかは関連がない（帰無仮説）。

H_1 : 小学校成績と離村するかどうかは関連する（対立仮説）。

表 5　検定統計量計算表

f_{ij}	$T_{i\cdot}T_{\cdot j}$	$F_{ij}=\frac{T_{i\cdot}T_{\cdot j}}{T}$	$f_{ij}-F_{ij}$	$(f_{ij}-F_{ij})^2$	$\frac{(f_{ij}-F_{ij})^2}{F_{ij}}$
20	2211	21.5	-1.5	2.25	0.105
18	2010	19.5	-1.5	2.25	0.115
29	2680	26	3	9	0.346
13	1188	11.5	1.5	2.25	0.196
12	1080	10.5	1.5	2.25	0.214
11	1440	14	-3	9	0.643
計					1.619

表 5 より

$$test=\sum_{1}^{s}\sum_{1}^{t}\left(f_{ij}-F_{ij}\right)^2/F_{ij}=1.619$$

$$\chi_2^2\left(0.05\right)=5.992$$

$$test<\chi_2^2\left(0.05\right)$$

したがって、「関連しない」という帰無仮説は棄却できず、母集団において小学校成績と離村するかどうかということの間に関連があるとはいえない（有意水準 5%）。

（8）例 2

（問）ある会社の常勤職員 100 名とパート職員 120 名について、仕事に満足しているか調べたところ、満足していると答えたものはそれぞれ 61 名と 60 名で､比率にするとそれぞれ 0.61 と 0.50 であった。母集団においても常勤職員とパート職員という雇用形態と満足しているかどうかということに関係があると言えるか。有意水準 5%で検討せよ。

表 6　常勤職員とパート職員の職業満足感

	常勤	パート	計
満足している	61	60	121
満足していない	39	60	99
計	100	120	220

（答）

H_0 : 常勤・パートの雇用形態と満足しているかどうかには関連はない（帰無仮説）。

H_1 : 常勤・パートの雇用形態と満足しているかどうかには関連がある（対立仮説）。

$$test=\sum_{1}^{s}\sum_{1}^{t}\left(f_{ij}-F_{ij}\right)^2/F_{ij}=2.667$$

$$\chi_2^2\left(0.05\right)=5.992$$

$$test<\chi_2^2\left(0.05\right)$$

したがって、「関連はない」という帰無仮説は有意水準 5%で棄却できず、母集団において常勤・パートの雇用形態と満足しているかどうかには関連があるとは言えない。

（9）補足

この定理の証明は推測統計の入門レベルを超えている。興味のある読者は「数理統計学」の専門書を参照してほしい。

例 2 にあるような 2×2 の表を四分表というが、四分表についての χ^2 検定は 6.2 で述べた「母比率の差の検定（2 つの母集団）」と実質的に同じである（補参照）。

6.2（7）の例では検定統計量の値は 1.633 であっ

たが、その上側確率は 0.051 であり、両側確率は 0.102 となる。すぐ上の例 2 では検定統計量の値は 2.667 であり、その上側確率は 0.102 である。自由度 1 の χ^2 分布は標準正規分布に従う確率変数の 2 乗の分布なので、標準正規分布で－1.633 と 1.633 の外側の両側確率は自由度 1 の χ^2 分布で 2.666（$=1.633^2$）の上側確率に等しくなり、ともに 0.102 となる。

これに直接関連することだが、χ^2 分布を用いて比率の差の検定をする場合、上側棄却域だけを見ているからといって片側検定と判断するのは間違いである。対立仮説は「比率が異なる」というものであり、「比率が大きい」「比率が小さい」といったものではない。すなわちここで行われているのは両側検定なのである。この両側検定を χ^2 分布に片側棄却域を設けて行っていると考えなければならない [14]。

6.6 無相関検定♦

（1）目的

2 つの変数には相関があると言いたい。

（2）定理（検定統計量とその確率分布）

X,Y に関する 2 次元正規母集団より大きさ n の標本を抽出したとき、母相関係数 $\rho=0$ であれば、下の検定統計量は、自由度 n－2 の t 分布に従う。ただし、R は標本相関係数。

$$Test=\frac{R\sqrt{n-2}}{\sqrt{1-R^2}}\sim t_{n-2}$$

（3）現実の標本を参照した検定統計量の値

$$test=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

（4）帰無仮説

$\rho=0$

（5）対立仮説

$\rho\neq 0$：両側検定（両側棄却域）

（6）解説

現実の標本における相関係数ならびに標本サイズから test の値を算出し、その値の Test の分布（t 分布）における位置から帰無仮説が棄却できるかどうかを検討する。

小標本でも大標本でも使うことができる検定方法である。大標本の場合は次の項の検定を用いることができる。

（7）例

（問）標本サイズが 27、標本相関係数が－0.215 であるとき、母集団においても相関があるといえるか。有意水準 5%で検定をせよ。

（答）

H_0：母集団において相関はない（帰無仮説）。

---

[14] 分散を χ^2 分布を用いて検定するような場合にはこのことは成り立たたず、両側検定を行う場合には分布の両側に棄却域を設ける必要がある。このとき、両方の確率の合計が有意水準となる（5.1）。F 分布を用いた分散の比の検定においても同様、両側検定を行う場合には分布の両側に棄却域を設ける必要があり、それらに対応する確率の合計が有意水準ととなる（5.2）。ただしこの分散の比の検定の場合、「分子を大きくして上側 2.5%に対応する棄却域を見ると両側 5%の有意水準の検定になる」というさらにややこしい事態が成立するので注意が必要だ。

H_1：母集団において相関がある（対立仮説）。

$$test=\frac{-0.215\sqrt{27-2}}{\sqrt{1-(-0.215)^2}}=-1.101$$

$$t_{25}\left(\frac{0.05}{2}\right)=t_{25}(0.025)=2.060$$

$$|test|<t_{25}\left(\frac{0.05}{2}\right)$$

したがって、有意水準5%（両側）で「相関がない」という帰無仮説が棄却できない。したがって、母集団において相関があるとはいえない。

（8）補足

この定理の証明は推測統計の入門レベルを超えている。

6.7 無相関検定（大標本）

（1）目的

大標本を用いて、2つの変数には相関があると言いたい。

（2）定理（検定統計量とその確率分布）

XYに関する2次元正規母集団より大きさnの大標本を無作為に抽出したとき、母相関係数$\rho=0$であれば、下の検定統計量は近似的に標準正規分布をなす。ただし、Rは標本相関係数。

$$Test=R\sqrt{n-1}\sim N(0,1)$$

（3）現実の標本を参照した検定統計量の値

$$test=r\sqrt{n-1}$$

（4）帰無仮説

$\rho=0$

（5）対立仮説

$\rho\neq 0$

（6）解説

現実の標本における相関係数ならびに標本サイズからtestの値を算出し、その値のTestの分布（標準正規分布）における位置から帰無仮説が棄却できるかどうかを検討する。

大標本で使える検定方法である（安田・原, 1982）。

（7）例

（問）標本サイズが4901、標本相関係数が0.211であるとき、母集団においても相関があるといえるか。有意水準1%で検定せよ。

（答）
H_0：母集団において相関はない（帰無仮説）。
H_1：母集団において相関がある（対立仮説）。

$$test=0.211\sqrt{4901-1}=14.7$$

$$z\left(\frac{0.01}{2}\right)=z(0.005)=2.5759$$

$$|test|>z\left(\frac{0.01}{2}\right)$$

したがって、有意水準１%（両側）で「相関がない」という帰無仮説が棄却でき、母集団においても相関があるといえる。

（8）補足

この定理の証明は推測統計の入門レベルを超え

ている。

7 対応がある場合の平均と比率の検定

7.1 対応のありなしとは◆

（1）平均の場合

平均や比率の検定においては「対応のある場合の検定」が必要な場合がある。ここではこういった検定が必要となる標本とそこから想定される母集団の関係について解説し、対応のある場合の平均と比率の検定を紹介する。

ある調査で表 7 のようなデータが得られたとしよう。ここで、性別は 1 が男、2 が女、統計学と調査法はテストの得点が示されている。

表 7　対応についての解説データ

ケース番号	性別	統計学	調査法
1	1	10	10
2	1	10	0
3	2	10	10
4	1	0	0
5	2	10	10
6	2	10	10
7	1	0	0
8	2	10	10
9	2	10	0
10	2	0	10
11	2	0	0

ここで行に注目しよう。ケース 1 からケース 11 までは無作為抽出で選ばれているので独立したものだ。たとえばケース 1 の統計学の得点はケース 2 の統計学の得点には影響しない。次に列に注目。ケース 1 の統計学の得点はケース 1 の調査法の得点に影響することは十分ありうる。1 人のケースについての得点だから、点数がランダムに生じるとは考えられないのである [15]。

このことを押さえて平均の検定について考えよう。まず、男女別の統計学の平均得点の差の検定について（表 8）。

表 8　統計学平均の男女差

	男	女	全体
統計学平均	5.00	7.14	6.36
n	4	7	11

この検定はこれまで述べてきた平均の差の検定で行ってかまわない。というのは、男性の統計学得点と女性の統計学得点にはケースの重なりがなく、2 つ独立した母集団より得られた標本と考えられるからである [16]。

統計学は 10 点で合格とするとき、調査法の得点平均について表 9 が得られる。

ここでも統計学合格者の調査法得点と不合格者の調査法得点にはケースの重なりがなく 2 つの独立した母集団からの標本と考えられる。それゆえ、

[15] この表に記されているものは、性別 X の分布をもつ母集団から抽出される確率変数の集合としての標本（X_1〜X_{11}）の実現値（x_1〜x_{11}）、統計学得点 Y の分布をもつ母集団から抽出される確率変数の集合としての標本（Y_1〜Y_{11}）の実現値（y_1〜y_{11}）、調査法得点 Z の分布をもつ母集団から抽出される確率変数の集合としての標本（Z_1〜Z_{11}）の実現値（z_1〜z_{11}）である。ここで述べていることをきちんと言うと、X_1〜X_{11} は相互に独立であり、Y_1〜Y_{11} は相互に独立であり、Z_1〜Z_{11} も相互に独立であるが、X,Y,Z は独立ではないということである。さらに付け加えると、復元抽出の場合、X の分布と X_1〜X_{11} の分布は等しく、Y の分布と Y_1〜Y_{11} の分布は等しく、Z の分布と Z_1〜Z_{11} の分布は等しい。そして X,Y,Z が独立ならば X_1,Y_1,Z_1 等も独立だが、X,Y,Z が独立でないならば X_1,Y_1,Z_1 等も独立ではない。

[16] きちんと言うと、統計学得点について、男性の統計学得点 Y_A の母集団と女性の統計学得点 Y_B の母集団という形で別々にとらえたとき、Y_A と Y_B という 2 つの確率変数は独立ということである。

これまで述べてきた検定法で平均の差の検定を行えばいい[17]。

表 9　統計学合格者・不合格者別の調査法平均得点

	統計学 合格者	統計学 不合格者	全体
調査法平均	7.14	2.50	5.45
n	7	4	11

しかし、統計学と調査法の得点平均に差があるかどうかを検定する場合には事情がちょっと異なってくる（表 10）。

表 10　統計学と調査法の平均得点

	統計学	調査法
得点平均	6.36	5.45
n	11	11

この場合、統計学においても調査法においても全データである 11 ケースを使って平均が算出されている。統計学の 11 のデータと調査法の 11 のデータは統計学得点の母集団と調査法得点の母集団から得られた別の標本と考えられるが、相互に独立とは言えず、ケースごとの対応関係がある[18]。このようなときに用いるのが対応のある場合の平均の検定である。

（2）比率の場合

比率の場合も事態は平均の場合とパラレルである。10 点を合格点と考え調査法と統計学についての四分表を作成しておこう（表 11）[19]。

表 11　統計学と調査法の合否についての四分表

		統計学 合格	統計学 不合格	計
調査法	合格	5	1	6
	不合格	2	3	5
計		7	4	11

この四分表を使うと、統計学の合格者と不合格者それぞれの調査法の合格率はそれぞれ 5/7=0.71 と 1/4=0.25 になる（表 12）。

表 12　統計学合格者と不合格者の調査法合格比率

	統計学 合格者	統計学 不合格者	全体
調査法合格率	0.71	0.25	0.55
n	7	4	11

この合格率の違いを検討するためには、統計学の合格者と不合格者ごとに調査法の合否の値を示す母集団があると考え、6.2 の「母比率の差の検定（2 つの母集団）」をすればいい。また、表 11 の四分表の χ^2 検定でも同等のことできる（6.5（9）ならびに補参照）。

しかしながら、次の表 13 のような母比率の差の検定、すなわち、全員を対象とした「統計学合格率

---

[17] すなわち、統計学合格者の調査法得点 Z_A の母集団と調査法不合格者の調査法得点 Z_B の母集団を考えるとき、Z_A と Z_B という 2 つの確率変数は独立ということである。

[18] 統計学得点 Y の母集団と調査法得点 Z の母集団について、X と Y は相互に独立とは言えないということである。

[19] この四分表をもとに前もって述べておくと、調査法の合格率に焦点を置くとき、「調査法の合格率は不合格率と異なるか」という問いは 6/11 と 5/11 の比較になり、その検定は 6.3 の「母比率の差の検定（同一母集団内）」である。「調査法の合格率は統計学の合格者と不合格者で異なるのか」という問いは 5/7 と 1/4 の比較になり、その検定は 6.2 の「母比率の差の検定（2 つの母集団）」である。「調査法の合格率は統計学の合格率と異なるか」という問いは 6/11 と 7/11 の比較になるが、その検定はこれから紹介する 7.3 の「母比率の差の検定（対応がある場合）」である。

0.64（=7/11）と調査法の合格率 0.55（=6/11）に差があるか」の検定ではこういった方法をとることはできない。というのは、統計学の得点と調査法の得点は独立しておらず対応関係があるからである。

表 13　統計学と調査法の合格率

	統計学	調査法
合格率	0.64	0.55
n	11	11

以下、このように対応がある場合の母平均の差の検定と母比率の差の検定について述べる。

7.2 母平均の差の検定（対応のある場合）

（1）目的

対応のあるデータをもとに、2 つの母平均に差があると言いたい。

（2）定理（検定統計量とその確率分布）

2 つの正規母集団よりそれぞれサイズ n の相互に独立でない標本を抽出するとき、2 つの母平均 μ_A と μ_B が等しいならば、下の検定統計量は近似的に標準正規分布 N(0,1)に従う。ただし $\overline{X}_A$ と $\overline{X}_B$ は標本平均、S_A^2 と S_B^2 は標本分散、R は標本相関係数。

$$Test = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\dfrac{S_A^2 + S_B^2 - 2RS_AS_B}{n}}} \sim N(0,1)$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\dfrac{s_A^2 + s_B^2 - 2rs_As_B}{n}}}$$

（4）帰無仮説

$\mu_A = \mu_B$

（5）対立仮説

$\mu_A \neq \mu_B$：両側検定（両側棄却域）
$\mu_A > \mu_B$：片側検定（右側棄却域）
$\mu_A < \mu_B$：片側検定（左側棄却域）

（6）解説

$\overline{x}_A$，$\overline{x}_B$, s_A, s_B, r, n を代入して test の値（検定統計量の値）を算出する。Test の分布（標準正規分布）における test の位置から帰無仮説が棄却できるかどうかを判断する。

（7）例

（問）ある学年の男性 100 人の冬休み前の体重と冬休み後の体重を測ったところ、冬休み前では平均 60kg、標準偏差 10kg であり、冬休み後では平均 65kg、標準偏差 12kg であった。両者の相関係数は 0.60 だった。母集団でも冬休み中に体重が増えていると言えるか。有意水準 5%の両側検定で検討せよ。

（答）

H_0：母平均は等しい（帰無仮説）。

H_1：母平均は異なる（対立仮説）。

$$test = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\dfrac{s_A^2 + s_B^2 - 2rs_As_B}{n}}}$$

$$= \frac{65-60}{\sqrt{\dfrac{10^2 + 12^2 - 2\times 0.60 \times 10 \times 12}{n}}} = 5$$

$$z\left(\frac{0.05}{2}\right) = z(0.025) = 1.96$$

$$\left|test\right| > z\left(\frac{0.05}{2}\right)$$

したがって、「母平均に差はない」という帰無仮説が有意水準 5%の両側検定で棄却でき、「母平均に差がある」と言える。

（8）補足

この検定統計量と下の 4.6 の「母平均の差の検定（大標本）◆」の検定統計量を見比べておいてほしい。

$$Test = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{S_A{}^2}{n_A} + \frac{S_B{}^2}{n_B}}} \sim N(0,1)$$

この検定統計量の n_A と n_B を 2 標本とも n とし、相関に関わる項 $2rs_As_B$ を付け加えたのがここでの検定統計量になる。

統計ソフトではこの定理ではなく、n が小さい時に使う次の定理を用いて対応のある平均の差の検定を行うことが多い。

■定理
2 つの正規母集団に属する個体が 1 対 1 の対応を持ち、かつ母平均が等しいならば、n 組の対（X_i, Y_i）からなる標本において次の検定統計量は自由度 n-1 の t 分布に従う。

$$Test = \frac{\sum D_i / n}{\sqrt{\frac{n\sum D_i^2 - \left(\sum D_i\right)^2}{n^2(n-1)}}} \sim t_{n-1}$$

$$D_i = X_i - Y_i$$

これは D_i（$=X_i-Y_i$）について、その母平均が 0 ではないと主張したい場合の t 検定（下式）の検定統計量と同じものである（4.2 参照）。

$$Test = \frac{\bar{X} - \mu}{\sqrt{U^2/n}} \sim t_{n-1}$$

7.3 母比率の差の検定（対応がある場合）

（1）目的

対応のあるデータをもとに、2 つの母比率に差があると言いたい。

（2）定理（検定統計量とその確率分布）

母集団に含まれる個体が 2 つの属性を持っており、属性 1 は A か a、属性 2 は B か b の値を示すとする。この母集団からの標本が、個体数 n を示した表 14 のような四分表で表現されるとき、属性 A をもつ個体の比率 p_A（= n(A)/N）と属性 B をもつ個体の比率 p_B（= n(B)/N）が母集団において等しいならば、下の検定統計量は近似的に自由度 1 の χ^2 分布に従う。

表 14　対応のある比率の検定のための四分表

	B	b	計
A	n(AB)	n(Ab)	n(A)
a	n(aB)	n(ab)	n(a)
計	n(B)	n(b)	N

$$Test = \frac{\left(\left|n(Ab) - n(aB)\right| - 1\right)^2}{n(Ab) + n(aB)} \sim \chi_1^2$$

（3）現実の標本を参照した検定統計量の値

$$test = \frac{\left(\left|n(Ab) - n(aB)\right| - 1\right)^2}{n(Ab) + n(aB)}$$

Test の n(Ab)等は確率変数だが、test の n(Ab)等は実際の標本より得られた値である。

（4）帰無仮説

$p_A = p_B$

（5）対立仮説

$p_A \neq p_B$：両側検定（右側棄却域）
χ^2 分布の右側を見るだけで両側検定の意味を持つ[20]。

（6）解説

現実の標本のサイズと n(Ab)、n(aB)の値を代入して test の値（検定統計量の値）を算出する。Test の分布（自由度 1 の χ^2 分布）における test の位置から帰無仮説が棄却できるかどうかを判断する。

（7）例

（問）ある町で町内会活動やボランティア活動に参加したいかどうかを調査したところ、次のような四分表を得た。母集団において町内会活動とボランティア活動で参加希望者の比率が異なるかどうかを有意確率 5%の両側検定で検討せよ。

表 15　町内会とボランティア活動への参加希望者

		ボランティア		計
		参加	不参加	
町内会	参加	20	10	30
	不参加	30	30	60
計		50	40	90

（答）
H_0：ボランティアと町内会活動の参加希望者の母比率は等しい（帰無仮説）。
H_1：ボランティアと町内会活動の参加希望者の母比率は異なる（対立仮説）。

$$test = \frac{\left(\left|n\left(Ab\right) - n\left(aB\right)\right| - 1\right)^2}{n\left(Ab\right) + n\left(aB\right)} = \frac{\left(\left|10 - 30\right| - 1\right)^2}{10 + 30}$$
$$= 9.025$$
$$\chi_1^2\left(0.05\right) = 3.841$$
$$test > \chi_5^2\left(0.05\right)$$

したがって、「母比率は等しい」という帰無仮説が有意水準 5%の両側検定で棄却でき、母集団でも「町内会活動とボランティア活動の参加希望者の比率は異なる」と言える。

（8）補足

この検定をマクネマー検定と言う。

8 おわりに

今回は統計的仮説検定の原理と実際の検定の方法について説明した。いろいろ述べてきたが、最も重要なのは、母集団における帰無仮説が正しい場合の検定統計量の確率分布がまず描かれるということだ。帰無仮説が検定統計量の確率分布を決定するのである。

一方、現実の標本のデータを参照して得られるのは、当該標本についての検定統計量の「値」である。この値は検定統計量の確率分布の中にまず位置づけられ、次いで帰無仮説下においてどのぐらいの確率でその値が生じうるのかが検討される。これが統計的仮説検定で行われていることなので

[20] 6.5（9）とそこにある注を参照されたい。

ある。

検定についての定理がなぜ成り立つのかということに関しては、基本的なものは押さえておいた方がいいだろう。そこで重要になるのは、母集団Xの平均がμ、分散がσ^2のときの標本平均の期待値はμ、分散はσ^2/nということだ。基本は本当に大事なのである。

いくつかの定理の証明は入門的レベルを超えていた。そういった問題についてはとりあえず置いておいておこう。ただ、定理の証明は難しいとしても、定理が何を意味しているのかはきちんと理解しておく必要はある。そして、定理の下でどのような確率分布が現れ、どんな筋道で検定がなされるのかということも理解しておかねばならない。

こういったことの理解があれば、知らない検定に直面しても、「帰無仮説は何か」「帰無仮説の下でどんな検定統計量がどのように分布するのか」といった基本的な問いが生まれ、その検定の意味を理解できるようになる。このことによって読者は「何をやっているのかわからないが結果はこうだ」という立場に別れを告げることができるのである。

補1：数表の読み方

（1）標準正規分布

標準正規分布の数表には2種類のものがある。1つは値zから確率αを求めるものである。これを抜粋したものが表 16 である。この数表には図 6 のような絵が描かれており、どの部分の確率の表なのかがわかるようになっている。ここで示されているのは、z より大きな値の範囲に対応する確率（上側確率）の表だということである。

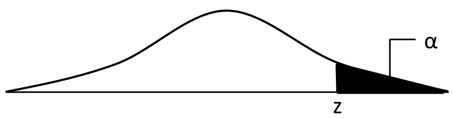
たとえば値が 1.96 のときの上側確率を求めるためには次のようにする。まず一番左の列から 1.9 を探し、そこから一番上の行を見ながら 0.06 の列まで右に移動していく。するとそこに示されている確率 0.0250 が 1.96（1.9＋0.06）という値の上側確率となる。

表 16　標準正規分布の数表（値→確率）

標準正規分布の値zに対応する上側確率α

z	0.00	…	0.06	…	0.09
0.0	0.5000	…	0.4761	…	0.4641
0.1	0.4602	…	0.4364	…	0.4247
:	:	…	:	…	:
1.9	0.0287	…	0.0250	…	0.0233
:	:	…	:	…	:
3.0	0.0013	…	0.0011	…	0.0010

図 6　どの範囲の確率かを示す絵



もう 1 つの数表は確率αから値 z を求めるものであり、表 17 のようなものである。

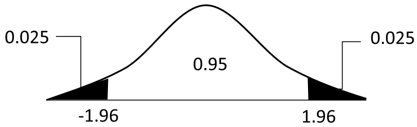
表 17　標準正規分布の数表（確率→値）

標準正規分布の上側確率αに対応する点の値z(α)

α	0.000	…	0.005	…	0.009
0.00		…	2.5758	…	2.3656
0.01	2.3263	…	2.1701	…	2.0749
0.02	2.0537	…	1.9600	…	1.8957
0.03	1.8808	…	1.8119	…	1.7624
0.04	1.7507	…	1.6954	…	1.6546
0.05	1.6449	…	1.5982	…	1.5632

こちらの表で上側確率 0.025 のときの値を求めるには、まず左の端の列を見て 0.02 を探し、そこから一番上の行を見ながら 0.005 の列まで右へ移動していく。そこに書かれている 1.96 が上側確率 0.025 に対応する値である。

図 7　正規分布の両側確率 5%の値の範囲と確率



標準正規分布は 0 を中心に左右対称だから、1.96 の上側確率が 0.025 であるなら－1.96 の下側確率も 0.025 になる。「－1.96 以下または＋1.96 以上」言い換えれば「－1.96 と＋1.96 の外側」の確率は合わせて 0.05（0.025×2）になるが、この値の範囲を両側確率 5%（0.05）に対応する値の範囲と言う。－1.96 から＋1.96 の値の範囲に対応する確率は 0.95（1－0.025×2）である（図 7）[21]。

数表からは上側確率が 0.005 になる値は 2.58 ということも読み取れる。したがって両側確率が 1%（0.01）となる値の範囲は「－2.58 以下または＋2.58 以上」である。－2.58 から＋2.58 の範囲に対応する確率は 0.99（1－0.005×2）となる。1.96 や 2.58 という数値は検定や推定でよく利用される。

（2）χ^2 分布

χ^2 分布の値と確率の関係については、各自由度の χ^2 分布の上側確率を記載した数表が利用される（表 18）。

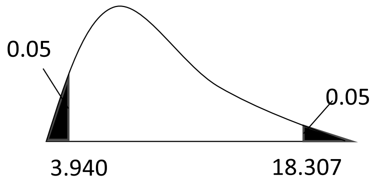
表をもとに自由度 10 の χ^2 分布の上側確率 0.05 と下側確率 0.05 に対応する値を求めると 18.307 と 3.940 となる。下側確率 0.05 に対応する値は上側確率 0.95 に対応する値になることに注意しよう。これを図に示すと次のようになる（図 8）。

χ^2 分布は標準正規分布に従う確率変数の 2 乗の和（自由度個の和）の分布だった。したがって、ある確率変数 X が標準正規分布に従うとき、その変数の 2 乗である X^2 は自由度 1 の χ^2 分布に従うことになる。そこからたとえば、標準正規分布の－1.96 と＋1.96 の外側の確率（0.05）は、自由度 1 の χ^2 分布の 1.96^2＝3.841 の上側確率（0.05）に等しい、といったことが起こる。

表 18　χ^2 分布の数表

自由度（df）	上側確率α				
	…	0.95	…	0.05	…
1	…	0.004	…	3.841	…
2	…	0.103	…	5.991	…
:	…	:	…	:	…
10	…	3.940	…	18.307	…
:	…	:	…	:	…
240	…	205.135	…	277.138	…

図 8　χ^2 分布の両側確率 10%の値の範囲と確率（df=10）



（3）t 分布

t 分布の値と確率の関係については、各自由度の t 分布の上側確率を記載した数表が利用される（表

[21] 値は横軸上に位置する点、値の範囲は横軸上の区間である。確率は値の範囲（区間）に対応する面積であり点に対応する面積ではないので「1.96 に対応する確率は 0.025」という表現はおかしい。それでは確率は 0 になってしまう。「1.96 より大きい値の範囲に対応する確率は 0.025」、あるいは「1.96 の上側確率は 0.025」という表現をしないといけない。

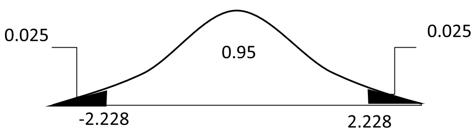
19）。分布は 0 を中心に左右対称なので、下側確率 0.025 の値を知りたければ、上側確率 0.025 に対応する値を求めマイナスを付ければいい。

表をもとに自由度 10 の t 分布の上側確率 0.025 と下側確率 0.025 に対応する値を求めると 2.228 と－2.228 となる。これを図に示すと図 9 になる。

表 19　t 分布の数表

自由度（df）	上側確率 α				
	…	0.05	0.025	…	0.005
1	…	6.314	12.706	…	63.657
2	…	2.920	4.303	…	9.925
:		:	:		:
10	…	1.812	2.228	…	3.169
:		:	:		:
240	…	1.651	1.970	…	2.596

図 9　t 分布の両側確率 5%の値の範囲と確率（df=10）



（4）F 分布

自由度を 2 つとる F 分布の数表は巨大なものになる。したがって、テキストなどには上側確率 0.05（5%）や上側確率 0.025（2.5%）の数表だけが通常掲載される。表 20 は上側確率 0.025 の F 分布の表である。

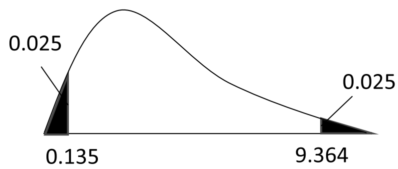
自由度（5,4）の上側確率 0.025 に対応する値を知りたいときは、この上側確率 0.025 の表を用い、m が 5、n が 4 になる値を見ればいい。その値は 9.364 である。

表 20　F 分布の数表

上側確率α=0.025

n ＼ m	1	…	4	5	…
1	647.789	…	899.583	921.848	…
2	38.506	…	39.248	39.298	…
3	17.443	…	15.101	14.885	…
4	12.218	…	9.605	9.364	…
5	10.007	…	7.388	7.146	…
:	:	…	:	:	…

図 10　F 分布の両側確率 5%の値の範囲と確率（df=5,4）



下側確率に対応する値を知りたいときには、F 分布についての次の性質を利用する。

$$F_n^m(\alpha)=\frac{1}{F_m^n(1-\alpha)}$$

この式をもとに、たとえば自由度（5,4）の F 分布の下側確率 0.025（＝上側確率 0.975）に対応する値を求めたいときには次の式を利用する。

$$F_4^5(0.975)=\frac{1}{F_5^4(0.025)}$$

すなわち、自由度を（4,5）と逆転させた上側確率 0.025 に対応する値を求め（7.388）、次いでこの逆数を求めるのである。

$$F_4^5(0.975)=\frac{1}{7.388}=0.135$$

これを図に示すと図 10 になる。

補 2 : 同一母集団内の母比率の差の検定と適合度検定

同一母集団内の母比率の差の検定（6.3）は適合度検定（6.4）と実質的に同じであることを明らかにしよう。

n を標本サイズとし、比率の差を検討する 2 つのカテゴリーを 1 と 2（添字）として、それぞれの度数を f_1, f_2、期待度数を F_1, F_2、標本比率を p'_1, p'_2 とするとき、2 つの母比率が等しい（$p_1=p_2$）場合、それぞれのカテゴリーの期待度数と度数は下の表のようになる。

表 21　同一母集団内の度数と期待度数

期待度数	$F_1=(f_1+f_2)/2$	$F_2=(f_1+f_2)/2$
度数	$f_1=np'_1$	$f_2=np'_2$

この分布における適合度検定の検定統計量は次のようになる。

$$\frac{(f_1-F_1)^2}{F_1}+\frac{(f_2-F_2)^2}{F_2}$$

$$=\frac{(f_1-F)^2}{F}+\frac{(f_2-F)^2}{F}$$

$$=\frac{\left(f_1-\frac{f_1+f_2}{2}\right)^2}{\frac{f_1+f_2}{2}}+\frac{\left(f_2-\frac{f_1+f_2}{2}\right)^2}{\frac{f_1+f_2}{2}}$$

これを変形していく。

$$\frac{\left(f_1-\frac{f_1+f_2}{2}\right)^2}{\frac{f_1+f_2}{2}}+\frac{\left(f_2-\frac{f_1+f_2}{2}\right)^2}{\frac{f_1+f_2}{2}}$$

$$=\frac{2}{f_1+f_2}\left(\frac{f_1-f_2}{2}\right)^2+\frac{2}{f_1+f_2}\left(\frac{f_1-f_2}{2}\right)^2$$

$$=\frac{(f_1-f_2)^2}{f_1+f_2}=\left(\frac{np'_1-np'_2}{\sqrt{np'_1+np'_2}}\right)^2$$

$$=\left(\frac{(p'_A-p'_B)}{\sqrt{\frac{(p'_A+p'_B)}{n}}}\right)^2$$

最後の式は、6.3 の母比率の差の検定（同一母集団内）における検定統計量の 2 乗の式である。したがって、同一母集団内の母比率の差の検定（6.3）は適合度検定（6.4）と実質的に同じであることがわかる。

自由度 1 の χ^2 分布は標準正規分布に従う確率変数の 2 乗の分布なので、標準正規分布で－k と＋k（k は任意の実数）の外側の確率の合計は χ^2 分布で k^2 の上側確率になる。比率の差の両側検定を適合度検定の方法を用いて行う際には χ^2 分布の上側だけに棄却域が設けられるので注意が必要だ（上側 5%の棄却域を設定すると 5%の両側検定となる）。

補 3 : 2 つの母集団の母比率の差の検定と独立性検定

2 つの母集団の母比率の差の検定（6.2）は四分表の独立性検定（6.5）と実質的に同じである。したが

って、母比率の差の検定をしたいときはχ^2検定をしてもいい。その際、5%の両側検定をしたいときにはχ^2分布の上側確率 5%に対応する範囲に棄却域を設ける必要がある。すなわち両側検定をしたいときでも片側に棄却域を設けるのである。

四分表の独立性検定（χ^2検定）とは「2 つの標本におけるある属性の比率（p_A と p_B）が標本で異なるとき、母集団でも異なるかどうか」を検定するものだ。そこでは表 22 ような度数の記載された四分表をもとに検定が進められる。f は度数、F は期待度数、n_A, n_B は標本サイズ、p_A, p_B は標本比率である[22]。

表 22　四分表

$f_{11}=n_A p_A$ $F_{11}=n_A p$	$f_{12}=n_B p_B$ $F_{12}=n_B p$	$(n_A+n_B)p$
$f_{21}=n_A(1-p_A)$ $F_{21}=n_A(1-p)$	$f_{22}=n_B(1-p_B)$ $F_{22}=n_B(1-p)$	$(n_A+n_B)(1-p)$
n_A	n_B	n_A+n_B

p はここでは 2 つの標本を合併したときの対象の比率、すなわち「合併比率」の値を表している。2 つの母集団の母比率に差がない場合、標本から得られた合併比率は 2 つの等しい母比率の不偏推定値となる。

合併比率の値は次の式で表される。

$$p=\frac{n_A p_A+n_B p_B}{n_A+n_B}$$

これらをもとに、「四分表の独立性検定の検定統計量」が「2 つの母集団の母比率の差の検定における検定統計量」の 2 乗であることを示し、双方の検定が実質的に同じであることを明らかにしよう。われわれは検定統計量の値についての下の等式の左辺から出発する。目標の式は右辺である。

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(f_{ij}-F_{ij}\right)^2}{F_{ij}}=\frac{\left(p_A-p_B\right)^2}{p(1-p)\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}$$

では出発。

$$\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(f_{ij}-F_{ij}\right)^2}{F_{ij}}=\frac{\left(f_{11}-F_{11}\right)^2}{F_{11}}$$
$$+\frac{\left(f_{12}-F_{12}\right)^2}{F_{12}}+\frac{\left(f_{21}-F_{21}\right)^2}{F_{21}}+\frac{\left(f_{22}-F_{22}\right)^2}{F_{22}}$$

$$=\frac{\left(n_A p_A-n_A p\right)^2}{n_A p}+\frac{\left(n_B p_B-n_B p\right)^2}{n_B p}$$
$$+\frac{\left(n_A(1-p_A)-n_A(1-p)\right)^2}{n_A(1-p)}$$
$$+\frac{\left(n_B(1-p_B)-n_B(1-p)\right)^2}{n_B(1-p)}$$

$$=\frac{\left(n_A(p_A-p)\right)^2}{n_A p}+\frac{\left(n_B(p_B-p)\right)^2}{n_B p}$$
$$+\frac{\left(n_A(p_A-p)\right)^2}{n_A(1-p)}+\frac{\left(n_B(p_B-p)\right)^2}{n_B(1-p)}$$

22 p_A や p_B は標本比率の値であるので、本シリーズの表記法では本来「'」(ダッシュ）が必要なのだが、式の変形が見づらくなるのでここでは省く。また、p は本来母比率を表すが、ここでは合併比率の値を表している。母比率は合併比率の値から推定される。

$$= \frac{n_A\left(p_A - p\right)^2}{p} + \frac{n_B\left(p_B - p\right)^2}{p}$$
$$+ \frac{n_A\left(p_A - p\right)^2}{\left(1-p\right)} + \frac{n_B\left(p_B - p\right)^2}{\left(1-p\right)}$$

$$= \frac{\begin{pmatrix} \left(1-p\right)n_A\left(p_A - p\right)^2 + \left(1-p\right)n_B\left(p_B - p\right)^2 \\ + pn_A\left(p_A - p\right)^2 + pn_B\left(p_B - p\right)^2 \end{pmatrix}}{p\left(1-p\right)}$$

目標の式の分母 p(1-p)が出てきた。

$$= \frac{\begin{pmatrix} n_A\left(p_A - p\right)^2\left(1-p+p\right) \\ + n_B\left(p_B - p\right)^2\left(1-p+p\right) \end{pmatrix}}{p\left(1-p\right)}$$
$$= \frac{n_A\left(p_A - p\right)^2 + n_B\left(p_B - p\right)^2}{p\left(1-p\right)}$$

簡単になったが分子に目標の式にない p があるので消そう。

$$= \frac{\begin{pmatrix} n_A\left(p_A - \dfrac{n_A p_A + n_B p_B}{n_A + n_B}\right)^2 \\ + n_B\left(p_B - \dfrac{n_A p_A + n_B p_B}{n_A + n_B}\right)^2 \end{pmatrix}}{p\left(1-p\right)}$$
$$= \frac{\begin{pmatrix} n_A\left(\dfrac{\left(n_A + n_B\right)p_A - n_A p_A - n_B p_B}{n_A + n_B}\right)^2 \\ + n_B\left(\dfrac{p_B\left(n_A + n_B\right) - n_A p_A - n_B p_B}{n_A + n_B}\right)^2 \end{pmatrix}}{p\left(1-p\right)}$$

$$= \frac{n_A\left(\dfrac{n_B p_A - n_B p_B}{n_A + n_B}\right)^2 + n_B\left(\dfrac{n_A p_B - n_A p_A}{n_A + n_B}\right)^2}{p\left(1-p\right)}$$

$$= \frac{\begin{pmatrix} n_A\left(\dfrac{1}{n_A + n_B}\left(n_B p_A - n_B p_B\right)\right)^2 \\ + n_B\left(\dfrac{1}{n_A + n_B}\left(n_A p_B - n_A p_A\right)\right)^2 \end{pmatrix}}{p\left(1-p\right)}$$
$$= \frac{\begin{pmatrix} n_A\left(\dfrac{1}{n_A + n_B}\right)^2\left(n_B p_A - n_B p_B\right)^2 \\ + n_B\left(\dfrac{1}{n_A + n_B}\right)^2\left(n_A p_B - n_A p_A\right)^2 \end{pmatrix}}{p\left(1-p\right)}$$
$$= \frac{\left(\dfrac{1}{n_A + n_B}\right)^2\begin{Bmatrix} n_A\left(n_B p_A - n_B p_B\right)^2 \\ + n_B\left(n_A p_B - n_A p_A\right)^2 \end{Bmatrix}}{p\left(1-p\right)}$$

目標の式を作るのに必要な $1/(n_A+n_B)$が出てきた。

$$= \frac{\left(\dfrac{1}{n_A + n_B}\right)^2\begin{Bmatrix} n_A\left(n_B^2 p_A^2 - 2n_B^2 p_A p_B + n_B^2 p_B^2\right) \\ + n_B\left(n_A^2 p_B^2 - 2n_A^2 p_A p_B + n_A^2 p_A^2\right) \end{Bmatrix}}{p\left(1-p\right)}$$

$$= \frac{\left(\dfrac{1}{n_A + n_B}\right)^2\begin{Bmatrix} n_A n_B\left(n_B p_A^2 - 2n_B p_A p_B + n_B p_B^2\right) \\ + n_A n_B\left(n_A p_B^2 - 2n_A p_A p_B + n_A p_A^2\right) \end{Bmatrix}}{p\left(1-p\right)}$$

$$= \frac{\left(\dfrac{1}{n_A + n_B}\right)^2 n_A n_B\begin{Bmatrix} \left(n_B p_A^2 - 2n_B p_A p_B + n_B p_B^2\right) \\ + \left(n_A p_B^2 - 2n_A p_A p_B + n_A p_A^2\right) \end{Bmatrix}}{p\left(1-p\right)}$$

目標の式を作るのに必要な $n_A n_B$ が出てきた。

$$=\frac{\left(\frac{1}{n_A+n_B}\right)^2 n_A n_B \left\{ \begin{array}{l} n_B\left(p_A{}^2-2p_A p_B+p_B{}^2\right) \\ +n_A\left(p_B{}^2-2p_A p_B+p_A{}^2\right) \end{array} \right\}}{p(1-p)}$$

$$=\frac{\left(\frac{1}{n_A+n_B}\right)^2 n_A n_B \left\{ \begin{array}{l} n_B\left(p_A-p_B\right)^2 \\ +n_A\left(p_A-p_B\right)^2 \end{array} \right\}}{p(1-p)}$$

$$=\frac{\left(\frac{1}{n_A+n_B}\right)^2 n_A n_B \left(p_A-p_B\right)^2\left(n_A+n_B\right)}{p(1-p)}$$

目標の式の分子の$(p_A-p_B)^2$が出てきた。

$$=\frac{\left(\frac{1}{n_A+n_B}\right) n_A n_B \left(p_A-p_B\right)^2}{p(1-p)}$$

$$=\frac{\left(\frac{n_A n_B}{n_A+n_B}\right)\left(p_A-p_B\right)^2}{p(1-p)}$$

$$=\frac{\left(p_A-p_B\right)^2}{p(1-p)\left(\frac{n_A+n_B}{n_A n_B}\right)}$$

$$=\frac{\left(p_A-p_B\right)^2}{p(1-p)\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}$$

おわり。

したがって、「四分表の独立性検定」と「2 つの母集団の母比率の差の検定」とは実質的に同じ検定なのである。

文献

馬場敬之・久池井茂, 2003『確率統計キャンパスゼミ』マセマ

稲垣宣生, 2003『数理統計学 (数学シリーズ)』裳華房.

小林久高,2018a「母集団・標本・確率変数」『同志社社会学研究』22.

小林久高,2018b「離散型確率変数とその分布」『同志社社会学研究』22.

小林久高,2018c「連続型確率変数とその分布」『同志社社会学研究』22.

小林久高,2019「統計的推定の原理と実際」『同志社社会学研究』23.

小寺平治,2002『ゼロから学ぶ統計解析』講談社

蓑谷千凰彦, 2003『統計分布ハンドブック』朝倉書店

安田三郎・原純輔, 1982『社会調査ハンドブック（第 3 版）』有斐閣