健康統計学 第14回

今回は、仮説検定の考え方と1標本の検定（83〜93ページ）について学習します。

テキスト

- 『やさしい保健統計学 改訂第5版』縣 俊彦著 (南江堂)

今回の内容

1. 仮説検定
    - 補足:仮説検定の考え方
2. 母平均の検定
3. 母比率の検定
4. 母相関係数の検定

仮説検定

仮説検定 (Statistical hypothesis testing)

- 統計的推定：標本の統計量から母数を、ある一定の値（点推定）、あるいは確率的な区間（区間推定）として推測する
- 統計的検定：母集団に関するある仮説（予測）のもとで、標本から得られる統計量を確率を求め、仮説の正当性を判断する
    - ただし、母集団の特性（母数）の程度を推測するわけではない

仮説検定の手順

母集団に関する仮説を立てて、その仮説が正しいという仮定のもとで、標本の統計量が得られる確率を求め、標本のようなことがめったに起こらない（偶然かどうか）を調べる。

1. 帰無仮説と対立仮説を立てる
2. 検定法を選択する
    - 分析目的にあった検定を分析者自身が選択する
3. 有意水準を決定する
    - 分析者自身が決定する
4. 検定統計量と有意点を算出する
    - 標本の値から検定統計量の値を算出する
    - 有意水準に対する有意点の値を算出する
5. 仮説の判定をする
    a. 検定統計量が棄却域に入っているかどうか調べる
    b. 検定統計量の値に対応するp値と有意水準を比較する

帰無仮説と対立仮説

検定では、まず検討したい母集団の特性に関する仮説を立てる。

一般には、明らかにしたい仮説を「**対立仮説**」(Alternative hypothesis) にして、それに対立する仮説を「**帰無仮説**」(Null hypothesis) とすることが多い。ただし、「仮説検定＝帰無仮説を棄却」することが目的ではないので注意すること。

- 帰無仮説 H_0
    - 帰無仮説に対立する仮説で、否定的な表現になる（「〜ではない」「〜と異なる」など）
    - 「母平均より標本平均のほうがおおきい」「2つの母集団の平均は等しくない」
- 対立仮説 H_1
    - 極論のような立証が難しそうで、肯定的な表現になる（「〜である」「〜と等しい」など）
    - 「母平均と標本平均は等しい」「2つの母集団の平均は等しい」

検定統計量

- 帰無仮説が正しいという仮定のもとで、標本から計算した値を、「**検定統計量**」という
- 検定統計量の計算式は、検定の方法によって決まっている
    - 標準正規分布やt分布のような、確率分布にしたがうような場合が多い

両側検定と片側検定

有意水準

- 正しい帰無仮説を棄却（採択しない）で、間違った対立仮説を採択するする確率を、「**有意水準**」(significance level) または「**危険率**」といい、α であらわす
    - 一般に（伝統的に）$\alpha = 0.05$ (5%) や $\alpha = 0.01$ (1%) が用いられる
- 有意水準 α から、検定法に応じて確率分布表から得られた値を「**有意点の値**」または「**限界値**」という
    - 帰無仮説か棄却されるかどうかは、検定統計量と有意点の値を比較して、判定する

棄却域と採択域による判定

検定統計量（標本から算出した値）と有意点（めったに起こるかどうかの確率から算出した値）を比較することで、帰無仮説の棄却・採択を判定する。

- それぞれの検定で用いる確率分布について、次のようにして判定する
    - 有意点より外側の部分を「**棄却域**」といい、検定統計量が棄却域に入っている場合は、帰無仮説は棄却される
    - 有意点より内側の部分を「**採択域**」といい、検定統計量が採択域に入っている場合は、帰無仮説は棄却されない

たとえば、帰無仮説 H_0 を「２つ母集団の母平均が等しい」とした場合、対立仮説は2通りの場合が考えられる。

1. 対立仮説「母平均は等しくない」 $H_1 : \mu_1 \neq \mu_2$
    - このような場合は「**両側検定**」を行う
    - 棄却域は、検定に用いる確率分布の両端にある
        - 「｜検定統計量｜＞有意水準 $\alpha/2$ に対応する有意点の値」なら帰無仮説は棄却
2. 対立仮説「どちらかの母平均が大きい（小さい）」 $H_1 : \mu_1 > \mu_2 \quad or \quad \mu_1 < \mu_2$
    - このような場合は「**片側検定**」を行う
    - 棄却域は、検定に用いる確率分布の片側にある
        - 「｜検定統計量｜＞有意水準 α に対応する有意点の値」なら帰無仮説は棄却

p値による判定

検定統計量と有意水準 α に対応する有意点の値を比較して判定する以外に、有意水準 α を直接使って判定する方法がある。

1. 検定統計量に対応する確率（p値）を確率分布から求める
2. p値と有意水準 α を比較する
    - 両側検定の場合
        - 「p値の2倍＜有意水準 α 」なら帰無仮説を棄却
    - 片側検定の場合
        - 「p値＜有意水準 α 」なら帰無仮説を棄却

母比率の検定では、確率の計算から直接p値を計算することができる。

判定結果の表現

- 帰無仮説を棄却する場合
    - 「帰無仮説を棄却する」
    - 「有意である」
    - 「対立仮説は正しい」（「絶対に」ではない）

- 「〜でない」

- 帰無仮説を採択する場合
  - 「帰無仮説を棄却できない」
  - 「有意でない」
  - 「帰無仮説は誤っているとはいえない」
  - 「〜でないとはいえない」

帰無仮説を採択する場合に、「帰無仮説を棄却できない」といい、「帰無仮説は正しい」といえないのは、誤った帰無仮説を採択してしまう可能性が少なからずあるためである。（後述の第2種の過誤のため）

第1種の過誤と第2種の過誤

- 正しい帰無仮説を棄却してしまうことを、「**第一種の過誤**」という
  - その確率を α（つまり有意水準）であらわす
  - 仮説検定では、第1種の過誤を重視（α と β の両方を小さくすることは簡単ではない）
- 間違った帰無仮説を採択してしまうことを、「**第二種の過誤**」という
  - その確率を β であらわす
  - 間違った仮説を正しく棄却する確率を $1-\beta$ であらわし、「検定力」という
- 第一種の過誤 α を小さくすると第二種の過誤 β が大きくなり、第二種の過誤 β を小さくすると第一種の過誤 α が大きくなる関係がある
  - 第一種の過誤 α も第二種の過誤 β も小さくするには、標本数をできるだけ大きくするしかない

	検定の結論	
	帰無仮説を棄却	**帰無仮説を採択**
帰無仮説は正しい	第一種の過誤 （確率 α）	正しい判断 （確率 $1-\alpha$）
帰無仮説は間違い	正しい判断 （確率 $1-\beta$）	第二種の過誤 （確率 β）

仮説検定の考え方

仮説検定の考え方について、簡単な例を用いて考えてみましょう。

コインを投げて表が出た回数を数える

例として、コインを10回投げて、表が出た回数を数えることを考えてみましょう。

ゆがみがない、いわゆる「かたよりのない」コインであれば、コインを1回投げた結果、表が出る確率も裏が出る確率も、ちょうど半分の $\frac{1}{2}=0.5$ と考えて問題ないでしょう。つまり、10回投げた結果として表が出る回数は5回ぐらいが最も多いと考えられます。

そこで、「**あるコイン**」を10回投げたところ、表が9回も出たとします。この「あるコイン」は「かたよりがない」コインでしょうか？ それとも「かたよりがある」コインでしょうか？

「コインにはかたよりがない」という仮説を立てる

仮説検定では、母集団に対するある**仮説**を立てます。そして、母集団から取り出した一部分、つまり標本を使って、その結果が偶然のものなのか必然なのかを確率的に調べて、仮説が正しいかどうかを判断する方法です。

今、「コインを10回投げたうち9回表がでた」ことについて考えています。コインにかたよりがある可能性がありそうです。そのことを仮説検定で確かめてみましょう。

仮説検定では、どちらかというと主張したいことに反対の仮説をまず考えます。これが、「母集団に対するある仮説」になります。この仮説を「無に帰することを予定した」という意味で、「**帰無仮説**」といいます。そして、帰無仮説に対立する仮説、つまり、どちらかといえば主張したい仮説を「**対立仮説**」といいます。

もしコインにかたよりがあるとしても、どの程度かたよっているかまではわかりません。そこで、帰無仮説として「**コインにかたよりはない**」という仮説を立てることにします。

まとめると、帰無仮説と対立仮説は次のようになります。

- 帰無仮説：「コインにはかたよりが**ない**」
- 対立仮説：「コインにはかたよりが**ある**」

コインの表が出る回数の確率を求める

次に、帰無仮説として立てた仮説のもとでの確率を求めて、その確率をもとに、今考えている事象（コインを10回投げたら9回表が出た）が偶然起こったことか必然的に起こったことかを判断してみましょう。

コインのように表または裏の2種類の結果を考えるには、第9回で学習した、二項分布の考え方を利用します。二項分布は、ある独立な試行について事象 A が起こる確率を p 、起こらない確率を $q(=1-p)$ とすると、この試行を独立に n 回繰り返したときに、事象 A が起こる回数を確率変数 X としたとき、 $X=x$ （つまり x 回起こる）となる確率は次のようになります。

$$\begin{aligned} \mathrm{P}(X=x) &= {}_n\mathrm{C}_x p^x q^{n-x} \\ &= \frac{n!}{x!(n-x)!} p^x q^{n-x} \\ &= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \end{aligned}$$

今回は、10回のうち表がでる回数を確率変数 X として考えます。表が出る確率も表が出ない（裏が出る）確率も同じで $p=q=\frac{1}{2}$ となりますから、表が x 回でる確率は次のようになります。

$$\begin{aligned} \mathrm{P}(X=x) &= {}_{10}\mathrm{C}_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x} \\ &= {}_{10}\mathrm{C}_x\left(\frac{1}{2}\right)^{10} \\ &= \frac{{}_{10}\mathrm{C}_x}{1024} \end{aligned}$$

この式を計算した結果を、確率分布として次のようにまとめておきます。

X	0	1	2	3	4	5	6	7	8	9	10
確率	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001

仮説から求めた確率をもとに判断でする

今考えている仮説は「コインはかたよりがない」です。では、10回中9回表が出るコインにかたよりがないのかあるのかを判断するには、どうすればよいでしょう。

ここで、「コインにかたよりがないという仮説のもとで、まれなこと（ある一定の確率以下の出来事）が起きた場合は、そのコインはかたよりがないとは見なせない」としましょう。これが仮説検定では重要な考え方です。

この「まれなこと」が起きたと判断する基準を、**有意水準**といいます。有意水準はあらかじめ決めておきます。一般には5%（0.05）か1%（0.01）が使われます。今回は有意水準を5%としておきましょう。

> 有意水準は、「100回のうち5回以下しか起こらない」というどのくらい稀（まれ）なことが起こるかの判断基準となります。
> しかし、正しい仮説であっても、100回のうち5回以下しか起こらないことが起こってしまう場合も考えられます。つまり、「100回のうち5回未満は間違った判断をして、正しい仮説を捨ててしまう」可能性があることになります。このような誤った判断をする危険があるため、有意水準を「**危険率**」とも呼びます。そして、「正しい帰無仮説を捨ててしまい、対立仮説を採択してしまう」ことを**第一種の過誤**といいます。

表が9回はでる確率は、「表が9回でた」場合と「表が10回でた」場合の確率を足し合わせたものになります。少なくとも9回は出た、と考えます。確率分布の表から、表が9回は出る確率は、0.010+0.001=0.011となります。有意水準を5%(0.05)と考えると、それより小さい確率です。

有意水準より小さい確率で起きてしまったことを、仮説検定では「**仮説では起こるはずのないことが起こった**」と見なします。このことを「**帰無仮説を棄却する**」といいます。もし、有意水準より大きい確率だった場合は、「仮説で起こるはずのないことが起こらなかった」とみなして、「帰無仮説を棄却できなかった」といいます。今回の場合、表が9回でるのはめったに起こらないことが起こったので、帰無仮説を棄却し、「**コインにはかたよりがある**」という判断になります。

ちなみに、表が8回でたコインが別にあったとしましょう。そのコインにかたよりがあるかどうかを考えると、8回以上表が出る確率は、確率分布から求めると、0.044+0.010+0.001=0.055となり、有意水準5%を超えることになります。つまり「コインにかたよりがない**とはいえない**」という判断になります。

母平均の検定

母平均の検定では、「**母平均と標本平均との差の程度**」を調べる。

- 帰無仮説 H_0 は「母平均と標本平均が等しい」：$\mu = \bar{x}$
- 対立仮説 H_1 は「母平均と標本平均が等しくない」：
  - 両側検定の場合は $\mu \neq \bar{x}$
  - 片側検定の場合は $\mu < \bar{x}$ または $\mu > \bar{x}$

母分散が既知の場合（z検定）

- 母分散 σ^2 を使う（めったにないことだが…）
- 母平均を μ、標本平均を $\bar{x}$、標本の大きさを n とする
- 標準正規分布にしたがう、検定統計量 z_0 を次の式から算出する

$$z_0 = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- 検定統計量 z_0 と、有意水準 α の有意点の値（標準正規分布表などから求める）を使って、判定をする
  - 片側検定
    - 帰無仮説 H_0 を棄却：$|z_0| \geq z_{(\alpha)}$
    - 帰無仮説 H_0 を採択：$|z_0| < z_{(\alpha)}$
  - 両側検定
    - 帰無仮説 H_0 を棄却：$|z_0| \geq z_{(\alpha/2)}$
    - 帰無仮説 H_0 を採択：$|z_0| < z_{(\alpha/2)}$

母分散が未知の場合（1標本t検定）

- 母分散 σ^2 の代わりに、不偏分散 s^2 を使う

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})$$

- 母平均を μ、標本平均を $\bar{x}$、標本の大きさを n とする
- 自由度 $df = n-1$ のt分布にしたがう、検定統計量 t_0 を次の式から算出する

$$t_0 = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

- 検定統計量 t_0 と、自由度 $df = n-1$、有意水準 α の有意点の値（t分布表などから求める）を使って、判定をする
  - 片側検定
    - 帰無仮説 H_0 を棄却：$|t_0| \geq t_{(\alpha)}(n-1)$
    - 帰無仮説 H_0 を採択：$|t_0| < t_{(\alpha)}(n-1)$
  - 両側検定
    - 帰無仮説 H_0 を棄却：$|t_0| \geq t_{(\alpha/2)}(n-1)$
    - 帰無仮説 H_0 を採択：$|t_0| < t_{(\alpha/2)}(n-1)$

母比率の検定

母比率の検定では、「**母比率と標本比率との差の程度**」を調べる。

- 帰無仮説 H_0 は「母比率と標本比率が等しい」：$p=\hat{p}$
- 対立仮説 H_1 は「母比率と標本比率が等しくない」：
  - 両側検定の場合は $p\neq\hat{p}$
  - 片側検定の場合は $p<\hat{p}$ または $p>\hat{p}$

二項検定

- 二項定理を使って、母比率に対する標本比率の統計値を直接計算し、有意水準と比較する
- 理論的には正当な方法だが計算が複雑なため、コンピュータによる統計処理が登場するまでは、正規分布に近似する方法（後述）などが使われていた。

考え方

- 母比率 p_0 の事象を、n 回試行するとき、
- r 回起きる確率は次のようになる

$$P_r={}_nC_r{p_0}^r(1-p_0)^{n-r}$$

- 「r 回以上」起きる確率は、次のような確率の和から算出できる

$$\begin{aligned}P&={}_nC_r{p_0}^r(1-p_0)^{n-r}+{}_nC_{r+1}{p_0}^{r+1}(1-p_0)^{n-r+1}+\cdots+{}_nC_n{p_0}^n(1-p_0)^0\\&=\sum_{i=r}^{n}{}_nC_i{p_0}^i(1-p_0)^{n-i}\\&=1-\sum_{i=0}^{r-1}{}_nC_i{p_0}^i(1-p_0)^{n-i}\end{aligned}$$

- 算出した確率（p値）と有意水準を比較する
  - 片側検定
    - 帰無仮説 H_0 を棄却：$P<\alpha$
    - 帰無仮説 H_0 を採択：$P\geq\alpha$
  - 両側検定
    - 帰無仮説 H_0 を棄却：$2P<\alpha$
    - 帰無仮説 H_0 を採択：$2P\geq\alpha$

正規分布に近似

- 特定の条件の時にだけ使える方法である
  - 標本数を n、母比率を p_0 とするとき、
  - $np_0>5,\ p_0<1-p_0$、または、$n>25$ の場合

考え方

- 標本の値を x、標本比率を $\hat{p_0}=\frac{x}{n}$ とする

- 標準正規分布に近似される、検定統計量 z_0 を次の式から算出する

$$z_0 = \frac{\widehat{p_0} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- 検定統計量 z_0 を使って判定をする

連続補正をする場合

- 二項分布は離散型の分布であるため、正規分布のような連続型の分布に近似すると、その精度はあまりよくない
- そこで、連続補正（イエーツ(Yates)の補正）をすることで、精度をよくする

$$z_{0c} = \frac{|x - np_0| - 0.5}{\sqrt{np_0(1-p_0)}}$$

- 補正した検定統計量 z_{0c} を使って判定をする

母相関係数の検定

母相関係数の検定では、「**母相関係数と標本相関係数との差の程度**」を調べる。

- 帰無仮説 H_0 は「母相関係数と標本相関係数が等しい」: $\rho = r$
- 対立仮説 H_1 は「母相関係数と標本相関係数が等しくない」:
  - 両側検定の場合は $\rho \neq r$
  - 片側検定の場合は $\rho < r$ または $\rho > r$

相関係数のz変換

母相関係数 ρ と標本相関係数 r をz変換する。

- 正規分布で近似させるために、フィッシャー(Fisher)のz変換で変換する

$$z_\rho = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right)$$

$$z_r = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$$

- $\ln$ は自然対数で $\log_e$ をあらわす

相関係数の検定

1. z変換した標本相関係数の標準誤差 s_{zr} を次の式から算出する

$$s_{zr} = \frac{1}{\sqrt{n-3}}$$

2. 検定統計量 z_0 を次の式から算出する

$$z_0 = \frac{z_r - z_\rho}{s_{zr}}$$

3. 検定統計量 z_0 を使って判定をする
   - 片側検定
     - 帰無仮説 H_0 を棄却 : $|z_0| \geq z_{(\alpha)}$
     - 帰無仮説 H_0 を採択 : $|z_0| < z_{(\alpha)}$
   - 両側検定
     - 帰無仮説 H_0 を棄却 : $|z_0| \geq z_{(\alpha/2)}$
     - 帰無仮説 H_0 を採択 : $|z_0| < z_{(\alpha/2)}$