医学統計学演習：資料 7

芳賀昭弘 *

1 統計的仮説検定

前回、信頼区間の推定について考えました。この推定を**検定**に応用することが本章の目的です。つまり、実際のデータの実現値が、そのデータが本来従うべきと考えた確率分布の信頼区間（一般に 95% や 99% がよく使われる）に入っているかどうかを調べ、入っていない場合には、従うべきと考えた確率分布に属さないデータであるとみなします*1。標本データがある確率分布に従っていると仮定することを、**帰無仮説**と呼び、帰無仮説とは逆の仮説を**対立仮説**と呼びます。こうした仮説を予め設定し、その仮説が正しいとした上で、実際に観察された標本の統計量（平均値など）の信頼区間を求め、仮定した分布の統計量がその範囲内であるかどうかで仮説が正しいかどうかを判断する方法を**統計的仮説検定**と言います。統計的仮説検定は、仮定した確率分布と実際の標本データの間に矛盾がないという仮説から出発します。「矛盾がある」ことを示したい場合でも、まずは「矛盾がない」という仮説（帰無仮説）を立てて、それが矛盾することを証明することで「矛盾がある」ことを示すという論理で進めます。

と、このような言葉だけではなかなか解らないと思いますので、まずは以下に示す基本的な検定の流れを例題に当てはめることで、仮説検定で使われる用語を確認するとともに統計的仮説検定から実際に得られる結果について眺めたいと思います。

1.1 検定の基本的な流れ

手順	やること
1	母集団に関する帰無仮説 *a と対立仮説 *b の設定。問題の設定の仕方により両側検定 *c or 片側検定 *d が決まる。
2	検定統計量（z 値、t 値、χ^2 値、F 値など）を選ぶ（検定方法を選ぶ）
3	有意水準 α の値を決める
4	データの検定統計量の実現値を求める
5	検定統計量の実現値が棄却域に入れば帰無仮説を棄却して対立仮説を採択する。

a: 「差がない」「効果がない」という仮説を設けること。H_0 という記号を使う。
b: 帰無仮説とは逆（「差がある」「効果がある」）の仮説。H_1 という記号を使う。
c: 差があるかどうかを知ることが重要であり、どちら（どれ）が「大きい」or「小さい」のかは気にしない（図 1 参照）。
d: ある集団より別の集団の方が「大きい」or「小さい」ことを知りたい場合（図 1 参照）。

* Electronic address: haga@tokushima-u.ac.jp

*1 入っている場合には、従うべきと考えた確率分布に属さないデータとは言えない、と考えます。

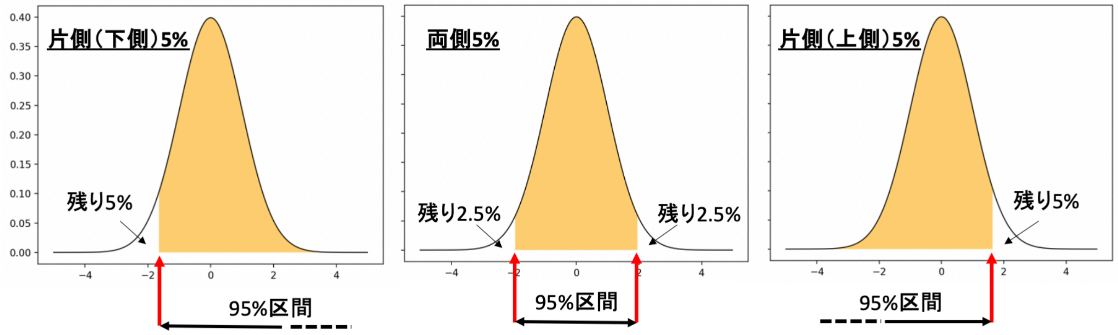

図 1　両側 5% と片側 5%（下側 2.5% と上側 2.5%）の意味

1.2　標準正規分布を用いた検定

母集団の性質（平均値や分散）がわかっている、もしくは母集団の分散が不明でもサンプル数が大きい場合に利用できる検定。

例題 1：あるゲームの得点について、平均点が 1024 で標準偏差が 256 である正規分布となることが調べられている。そのゲームで 4 回行った平均が 1250 点であった時、そのゲームについて得意・不得意であると言ってよいでしょうか？有意水準を 5% として検定せよ。

（手順）
帰無仮説 H_0：1250 点は平均 1024、標準偏差 256 の正規分布からサンプルされたものである（珍しくない、あり得る点数である）。（対立仮説 H_1：平均 1024、標準偏差 256 の分布のサンプルから得ることはできない（難しい））
検定統計量：z 値（標準正規分布の変数）。ここで $z = \frac{(x-\mu)}{\sigma/\sqrt{n}}$ で、μ、σ はそれぞれ母集団の平均値と標準偏差。n はサンプル数。
有意水準：両側 5% とする
検定統計量の実現値：

____

z 分布の両側 5% における限界値

____

p 値が____________であり、統計検定量は有意水準______である。
よって帰無仮説は______________________________。

例題 2：0 から 10000 番まで番号がついたアンケートを配布した。そのうち 100 枚のアンケートが回収され、番号の平均値が 4135 であった。回収されたアンケートはランダムにサンプルされていると言えるでしょうか？（ヒント：母集団 0~10000 を全て使って解析可能（平均値や分散も求めることができる）なので、母集団の性質がわかっている場合に利用できる検定を使う）

（手順）

帰無仮説 H_0：回収されたアンケートはランダムに抽出された → 100 枚のアンケート番号の平均値は 5000 である。（対立仮説 H_1：100 枚のアンケート番号の平均値は 5000 ではない）
検定統計量：z 値（標準正規分布の変数）。ここで $z = \frac{(x-\mu)}{\sigma/\sqrt{n}}$ で、μ、σ はそれぞれ母集団の平均値と標準偏差。n はサンプル数。
有意水準：両側 $a\%$ とする
検定統計量の実現値：

____

p 値が__________であり、統計検定量は有意水準_____である。
よって帰無仮説は________________________。

1.3 t 分布を用いた検定

母集団の分散がわからずサンプルが少数である場合に利用できる検定。

例題３：A 社の放射線治療装置を購入した場合、１時間当たり 4 名以上の治療が可能であれば採算がとれるという。近隣の 5 つの病院では、A 社の放射線治療装置を用いて１時間当たり平均 4.6 名、標準偏差 0.6 名という結果が得られていた。A 社の放射線治療装置を購入すべきかどうかを判定せよ。
（考え方） 5 つの病院（標本）での平均が 4.6 名、標準偏差 0.6 名という結果は、統計学的に確かに 4 名以上（大もとの母集団の平均が 4 名以上）で治療可能と言えるのか？

（手順）
帰無仮説 H_0：A 社の新型放射線治療装置による治療可能件数は１時間当たり 4 名である（対立仮説 H_1：4 名を超える）
検定統計量：t 値（t 分布の変数）。ここで自由度 $n-1$ の t 分布は（n はサンプル数）

$$t = \frac{x-\mu}{s/\sqrt{n-1}} = \frac{x-\mu}{\hat{\sigma}/\sqrt{n}} \tag{1}$$

である（μ は母集団の平均、s は標本標準偏差）。これを使って母集団の平均を推定する。
有意水準：片側 5% とする
統計量の実現値：
まず片側 5% の確率となる t 値の限界値は

$> t_{5\%}$ = qt(***,4,lower.tail = FALSE)

で与えられる（***には何が入りますか？）。この値を上の式に代入して μ を求めると、

> 4.6-$t_{5\%}$*0.6/sqrt(4)

____

これは、サンプルの標本平均値と標本標準偏差から推定された下側 95% 信頼区間の下限の治療人数である。この推定から採算が取れると確実に言えるでしょうか？（4.6 名というのは、どの程度の "確実" な結果なのだろうか？ → 演習）

相関係数の検定に利用できる検定。

例題４：密度がわかっている７つの物質を CT で撮影したところ、CT 値（画像上での濃淡を表す信号値）は次のようなデータとなった。
密度 0.26, 0.52, 0.93, 1.0, 1.25, 1.46, 1.92
CT 値 -817, -522, -90, 0, 491, 741, 1547
密度と CT 値に相関があると言えるか？

（手順）
帰無仮説 H_0：$\rho = 0$、すなわち母相関はない（対立仮説 H_1：$\rho \neq 0$、すなわち母相関がある）
検定統計量：t 値。なぜなら r を相関係数とすると、$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ は自由度 $n-2$ の t 分布となる。
有意水準：両側 $a\%$ とする

```
> dens = c(0.26, 0.52, 0.93, 1, 1.25, 1.46, 1.92)
> ct = c(-817, -522, -90, 0, 491, 741, 1547)
> plot(dens,ct)
> r = cor(dens,ct)
> n = length(dens)
> tvalue = ___________
> pt(tvalue, n-2, lower.tail=FALSE)*2
```

____

p 値が___________であるので、統計検定量は有意水準_____である。
よって帰無仮説は___________________________。

なお、相関係数の検定には、cor.test 関数が用意されている：

```
> cor.test(dens,ct, conf.level=0.95)
```

上と同じ結果が与えられていることを確認できたでしょうか？

2 演習

1. 例題３で有意水準をいくらに設定すると、４名という結果が得られるか？
   （つまり、p 値はいくらであるか？）

2. ある血液の μL 辺りの白血球数を５回測定したところ、
   5623, 5887, 5054, 6289, 6003
   という観測値が得られた。真の白血球数 [個 $/\mu L$] の 95% 信頼区間及び 99% 信頼区間を求めよ。
   ヒント：
   自由度 n-1 の t 分布は

$$t = \frac{x-\mu}{s/\sqrt{n-1}} = \frac{x-\mu}{\hat{\sigma}/\sqrt{n}}$$

である（μ は母集団平均、s は標本標準偏差）。
95% 信頼区間を与える t 値（上側）は、

```
> p95level = qt(0.025,4,lower.tail = FALSE)
```

99% 信頼区間を与える t 値（上側）は、

```
> p99level = qt(0.005,4,lower.tail = FALSE)
```

したがって、

$$-\text{p95level} < \frac{x-\mu}{s/\sqrt{n-1}} < \text{p95level}$$

となる x が白血球数 [個 $/\mu L$] の 95% 信頼区間（99% 信頼区間も同様に求める）。

3. 過去のデータから、あるテストの平均値は 60 の正規分布に従うことがわかっている。今回 20 名でそのテストを実施したところ、平均値が 65、標本分散が 8 であった。この 20 名は、過去の人たちと同じ母集団からの無作為標本と考えてもよいか？
（手順）
帰無仮説 H_0：無作為標本と考えて良い。
検定統計量：t 値（母集団の標準偏差が不明でサンプルもそれほど多くはない）。
有意水準：両側 5% とする。