数理統計学　前回の解答

第 8 回分

次のように作成した x が 9 個, y が 10 個のデータに対して、母分散が未知の場合の母平均の差の 95% 信頼区間を求めよ。

```
> set.seed(学籍番号)
> x <- round(rnorm(9,15,3.5),digits=1)
> y <- round(rnorm(10,17,3.5),digits=1)
```

解答例：学籍番号が 1825200 の場合

データから標本平均・標本分散を求めると

```
> set.seed(1825200)
> x <- round(rnorm(9,15,3.5),digits=1)
> y <- round(rnorm(10,17,3.5),digits=1)
> c(mean(x),mean(y))
 [1] 14.75556 18.60000
> c(var(x),var(y))
 [1]  8.510278 31.522222
> c(sd(x),sd(y))
 [1] 2.917238 5.614465
```

共通な母分散 σ^2 が未知なので、推定値を求めると

$$s^2 = \frac{1}{(9-1)+(10-1)}\{(9-1)\times 8.510 + (10-1)\times 31.52\} = 20.691\cdots$$

さらに、s の値は $\sqrt{20.69} = 4.5486\cdots$ である。

$$T = \frac{(\bar{X}-\bar{Y})-(\mu_x-\mu_y)}{s\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

で変換すると、T は自由度 17(=(9-1)+(10-1)) の t 分布に従うことがわかる。ここで、

$$\Pr\{-t_\alpha \leqq T \leqq t_\alpha\} = 0.95 = 1-0.05$$

となる t_α を R を使って求めると

```
> qt(1-0.05/2,17)
 [1] 2.109816
```

であることが求められる。以上のことから、母平均 μ の 95%信頼区間は

$$-2.110 \leqq \frac{(14.76-18.60)-(\mu_x-\mu_y)}{4.549\times\sqrt{\frac{1}{9}+\frac{1}{10}}} \leqq 2.110$$

$$-3.84 - 2.110\times 2.090 \leqq \mu_x-\mu_y \leqq -3.84 + 2.110\times 2.090$$

$$-8.2499 \leqq \mu_x-\mu_y \leqq 0.5699$$

$$-8.25 \leqq \mu_x-\mu_y \leqq 0.570$$

であるから、$[-8.25, 0.570]$ である。《精密に求めると $-8.2541\cdots \leqq \mu \leqq 0.565290\cdots$》

学籍番号による解答一覧

学籍番号	推定した s の値	下限	上限	下限 (3 桁)	上限 (3 桁)
1825200	4.549	-8.25418	0.5652909	-8.25	0.565
1523051	2.967	-5.627497	0.1252748	-5.63	0.125
1723016	3.127	-4.647195	1.4160842	-4.65	1.42
1823019	2.860	-4.089907	1.4543518	-4.09	1.45
1823038	2.724	-3.495966	1.7848546	-3.50	1.78
1823058	3.424	-4.32495	2.3138391	-4.32	2.31
1823067	3.335	-3.626939	2.8380498	-3.63	2.84
1823068	3.440	-4.959216	1.7103269	-4.96	1.71
1823070	3.949	-6.122572	1.5336827	-6.12	1.53
1823072	4.445	-6.668938	1.9489379	-6.67	1.95
1823076	2.397	-5.286798	-0.6398682	-5.29	-0.640
1823091	3.365	-7.085293	-0.561374	-7.09	-0.561
1823100	3.876	-5.834415	1.6810816	-5.83	1.68
1823102	3.685	-4.314009	2.8295647	-4.31	2.83
1823109	3.533	-6.318107	0.5314405	-6.32	0.531
1823131	2.724	-3.542104	1.7398816	-3.54	1.74

レポート作成についての注意点

・解答については、必ず途中計算もしくはプログラムを記載すること。R の結果を使う場合は、どこの部分をしようしたのかわかるようにすること。

《解答のみは点数を半分といたします》

例．１．(4) の答えは、R の結果の $\boxed{1}$ の部分のプログラムと結果から ⋯ である。

・[1] (6) の正規分布・t 分布のグラフについては手書き不可とする。

《R で直接印刷するか、Word にコピーして印刷する》

・レポートは次の順番とすること

- [1] の解答　　《**答えは手書き。R の結果は印刷可**》

  (1)∼(3) は考え方や途中式も記載すること

  (4),(5) はプログラムを記載すること（最終的な解答はわかるように書くこと）

- [1] (6) のグラフ　　《**グラフは必ず印刷**》

  プログラムは 3 行なので、印刷または手書きで書いておくこと

  どちらが正規分布・t 分布のグラフかわかるように記載すること

- [2] の解答　　《**必ず手書き**》

  R でのプログラム・結果は別紙（次ページ）とし、このページは解答のみとする。

- [2] の R のプログラム・結果

  手書きで書く場合は解答と一緒に書いてもかまわない

- [3] の解答　　《**必ず手書き**》
- [3] の R のプログラム・結果

  必ずすべてのデータ ($x_1, \cdots, x_{30}$ の値) を記載すること。

[1] (6) のグラフ作成についてのヒント

・標準正規分布・t 分布の確率密度関数をあらわす関数は

`dnorm(`x`)` と `dt(`x, 自由度`)`

《`pnorm(`x`)` や `pt(`x, 自由度`)` は分布関数をあらわす関数なので×》

・x の範囲（定義域）は $[-4, 4]$ または $[-5, 5]$ として下さい。

・プログラムは 3 行で書けます。

（正規分布を書く／上書き命令／t 分布を書く）

１．仮説検定とは

　今までは、母集団の母平均や母分散等を標本から推定することを行ってきたが、仮説検定は、事前に立てた母集団に対する性質の仮定が正しいかどうかを判定する一つの方法である。もちろん推定と同様に、母集団にあるすべてのデータを調べることができれば正しい性質も分かるが、実際には一部のデータ（標本）しか得ることができないので、標本から判断せざるを得ない。しかしながら、すべてのデータが手に入らないということは、正しい判定ができないことを意味しており、仮説検定は確率を用いて、なるべく誤った判定を少なくする方法を用いている。

　実際の仮説検定の方法はおおよそ次のとおりである。

１）仮説を立てる。
　調べたいことに対して、２つの仮説を立てる。帰無仮説 (H_0) と対立仮説 (H_1) と呼ばれており、示したい（証明したい）仮説を対立仮説としておき、対立仮説を否定したものを帰無仮説としておく。

２）母集団から標本を抽出する。
　基本的は無作為抽出でよいが、調べたい事柄によっては層別抽出などを用いる。

３）帰無仮説を真と仮定したときの、標本の出現確率を計算する。
　仮説検定は背理法の考え方を用いているので、正しくないと思われる仮説を真として、標本の出現確率を調べる。

４）３）の出現確率が極めて低いとき、帰無仮説を棄却して対立仮説が正しいと判定する。そうでないときは、判断を保留する。
　３）の出現確率が低い場合、めったに起こらないことが起きたと考えるより、最初の仮説が間違っていたと考えたほうが妥当である。そこで、帰無仮説が否定されるので、対立仮説が正しいと判断できる。しかし、棄却できない場合、どちらが正しいかという積極的な理由が無いため、判断を保留する（背理法と同じ）。

前にも述べたように、仮説検定では確率で帰無仮説を棄却するかどうかを判断しているために、判断を誤る可能性がある。どのようなケースがあるかを考えると次のような表になる。

	帰無仮説を棄却しない	帰無仮説を棄却する
帰無仮説が真	正しい判定	誤った判定 (第１種の過誤)
対立仮説が真	誤った判定 (第２種の過誤)	正しい判定

できれば第１種の過誤の確率と第２種の過誤の確率をできるだけ小さくしたいが、両方の確率を制御することは非常に困難である。仮説検定では、４）の帰無仮説を棄却する確率が第１種の過誤の確率と一致するので、この値を固定した上で第２種の過誤の確率が小さくなるような方法を用いる。一般に“1−(第２種の過誤の確率) =（対立仮説が真のとき帰無仮説を棄却する確率）”の値を検出力と呼び、この値が大きいほど良い検定である。

2019年度 数理統計学 D科　小テスト解答用紙　2019.11.29

・帰無仮説と対立仮説　《仮説検定》

　ガソリンに添加剤を加えた場合に燃費が良くなるかどうかを検定する場合、
データとして

「(添加後の燃費) − (添加前の燃費)」　《単位は km/ℓ》

が得られたときの帰無仮説および対立仮説を示せ。

2019年度神奈川工科大学 数理統計学 演習問題	学科	学年	組	学　籍　番　号	氏　　　名	

提出先：K3-3309号室前　20番のボックス　提出期限：１２月　２日（月）１７時頃まで

学籍番号別：データの確認

学籍番号	$x[1]$	$y[10]$
1825200	16.0	16.2
1523051	16.9	20.4
1723016	11.9	20.5
1823019	18.1	12.0
1823038	19.0	17.3
1823058	12.3	11.5
1823067	17.6	23.2
1823068	17.8	17.4
1823070	9.7	24.0
1823072	22.2	15.2
1823076	16.2	15.6
1823091	16.2	18.4
1823100	13.6	21.3
1823102	14.1	14.4
1823109	20.9	20.8
1823131	22.1	21.4