数理統計学　前回の解答

新しい電球の消費電力を測定したところ下記のような結果を得た（データは学籍番号を使って下記のように作成すること）。

```
> set.seed(学籍番号)
> x <- round(rnorm(8,23.5,1.23),digits=1)
```

（例．1825200 の場合　23.8, 22.1, 24.8, 24.0, 23.6, 24.7, 22.3, 22.4）

同じ明るさの旧型の電球の消費電力の平均が $\mu = 25.1$ であるとき、新しい電球の消費電力が下がったとしてよいか、有意水準５％で検定せよ。

（ただし、消費電力は正規分布になると仮定する）≪解答例（ 1825200 の場合）≫

（１）帰無仮説と対立仮説を立てる。

新しい電球の消費電力が旧型の平均 $\mu = 25.1$ よりも小さくなることを証明したいので

帰無仮説：新しい電球の消費電力は旧型に比べて高いか同じ　　H_0: $\mu \geqq 25.1$

対立仮説：新しい電球の消費電力は旧型に比べて低い　　H_1: $\mu < 25.1$

（２）標本の抽出　R で 8 個のデータを作成

```
> set.seed(1825200)
> x <- round(rnorm(8,23.5,2.12),digits=1)
> x
 [1] 23.8 22.1 24.8 24.0 23.6 24.7 22.3 22.4
```

（３）帰無仮説が真の場合の（２）で得られた標本の出現確率を求める。

まず、標本平均と標本分散を計算すると

```
> c(mean(x),var(x),sd(x))
 [1] 23.462500  1.154107  1.074294
```

であるから、$\bar{x} = 23.46, s^2 = 1.154, s = 1.074$ である。今回の検定は、消費電力が低くなることを証明するため、帰無仮説 $\mu \geqq 25.1$ は $\mu = 25.1$ のみ考えれば十分である（$\mu = 25.1$ のときに帰無仮説を棄却できれば、25.1 より小さいことが証明できるので）。したがって、母平均を $\mu = 25.1$ とし、母分散が未知なので t で変換すると、

$$t = \frac{23.46 - 25.1}{\frac{1.074}{\sqrt{8}}} = -4.3190\cdots$$

である。また、$\mathrm{P}(T < -4.319)$ の値は自由度 $8 - 1 = 7$ の t 分布に従うので、`pt(-4.319,7)` より $0.0017422\cdots$ である。

（４）判定

この問題は小さいときのみ帰無仮説を棄却する片側検定なので、（３）で求めた値が、５％より小さい場合に帰無仮説を棄却する。実際 $0.00174 < 0.05$ なので、帰無仮説を棄却することができる。つまり、新しい電球の消費電力は旧型に比べて下がったといえる。

ちなみに棄却域を使う場合は、`qt(0.05,7)` の値が -1.895 であることを使い、$t < -1.895$ なので、帰無仮説を棄却することが出来る。

《t.test を使う場合》対立仮説が “<” なので、alternative="less"を使う。

```
> t.test(x,mu=25.1,alternative="less")

        One Sample t-test

data:  x
t = -4.3113, df = 7, p-value = 0.001759
alternative hypothesis: true mean is less than 25.1
                          ⋮
```

のような結果を得る。t.test の場合、p-value の値と有意水準を直接比較すればよいので、帰無仮説を棄却する。つまり、新しい電球の消費電力は旧型に比べて下がったといえる。

結果一覧

学籍番号	$\bar{y}$	s	t	確率	判定
1825200	23.46	1.074	-4.311	0.001759	棄却する
1523051	23.75	0.7329	-5.210	0.000620	棄却する
1723016	23.69	1.152	-3.468	0.005216	棄却する
1823019	23.70	1.276	-3.103	0.008626	棄却する
1823038	23.64	0.9724	-4.254	0.001887	棄却する
1823058	23.58	1.244	-3.467	0.005224	棄却する
1823067	23.69	1.049	-3.807	0.003327	棄却する
1823068	24.41	1.201	-1.620	0.074659	棄却できない
1823070	23.10	1.451	-3.898	0.002956	棄却する
1823072	23.66	1.970	-2.064	0.038942	棄却する
1823076	23.39	1.056	-4.586	0.001263	棄却する
1823091	23.26	1.165	-4.462	0.001465	棄却する
1823100	23.28	1.617	-3.193	0.007607	棄却する
1823102	24.21	1.142	-2.198	0.031953	棄却する
1823109	23.29	1.723	-2.975	0.010333	棄却する
1823131	24.25	1.045	-2.301	0.027444	棄却する

注）1622023 の確率は 8.10×10^{-08} です。

数理統計学 二標本の平均値の差の検定

2017.12.13

推定のときと同じように、二標本の平均の差に関する検定を行うことができる。たとえば、２種類の釘の長さに違いがあるかどうかを調べたいのであれば、

１）仮説を立てる（有意水準１％）

調べることは２種類の釘の長さが「異なるかどうか」であるから、釘Ａの平均を μ_x とし、釘Ｂの平均を μ_y とすれば、２つの仮説は次の通り

帰無仮説 H_0：同じである $(\mu_x = \mu_y \;\Rightarrow \mu_x - \mu_y = 0)$

対立仮説 H_1：異なる $(\mu_x \neq \mu_y \;\Rightarrow \mu_x - \mu_y \neq 0)$

２）母集団から標本を抽出する

それぞれの釘を無作為抽出し釘の長さを１５個調べた。

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

実際のデータから標本平均、標本分散を計算すると次の通り。

釘Ａ	$5.2,\ 4.9,\ 4.7,\ \cdots,\ 5.8$	$\bar{x} = 4.92,\ \ s_x^2 = 0.1617$
釘Ｂ	$5.6,\ 6.1,\ 6.0,\ \cdots,\ 6.8$	$\bar{y} = 5.96,\ \ s_y^2 = 0.4040$

共通の母分散の推定値は

$$\begin{aligned} s^2 &= \frac{1}{(n-1)+(m-1)}\left\{(n-1)s_x^2 + (m-1)s_y^2\right\} \\ &= \frac{1}{14+14}\{14 \times 0.1617 + 14 \times 0.4040\} = 0.28285 \end{aligned}$$

母分散が未知なので、区間推定と同様の変換を使って、

$$t = \frac{(\bar{x}-\bar{y}) - (\mu_x - \mu_y)}{s\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

が自由度 $(n-1)+(m-1)$ の t 分布になることを利用する。実際に t の値を求めるには帰無仮説が真であることから $\mu_x - \mu_y = 0$ として

$$t = \frac{(4.92-5.96)-(0)}{\sqrt{0.2829}\sqrt{\frac{1}{15}+\frac{1}{15}}} = -5.355$$

棄却域を求めるのであれば、自由度 28 の t 分布の両側１％点なので

`qt(1-0.01/2,28)`$\Rightarrow$ 2.763262 である。

確率を求めるのであれば、自由度 28 の t 分布における $\Pr\{T < -5.355\}$ の値なので

`pt(-5.355,28)`$\Rightarrow$ 0.000005258 である（約１９万分の１）。

４）判定

棄却域を使う場合は、$|t| > 2.763$ と明らかに棄却域に入るので帰無仮説を棄却する。

確率を使う場合は、０．５％以下か９９．５％以上のとき棄却すればよいので、明らかに棄却する。

いずれの方法でも釘の長さは異なるといえる。

≪ t.test の使い方≫

平均の検定問題で、母分散が未知の場合についてはt.testという関数を使うとtの値や確率を自動的に計算してくれます。

《1標本の場合》

t.test(データ, mu = 帰無仮説, alternative = "対立仮説")

《2標本の場合 ($H_0 : \mu_x = \mu_y$)》

t.test(データ**X**, データ**Y**, alternative = "対立仮説" , var.equal = T)

いずれの場合も対立仮説は

two.side…両側検定 ($\mu \neq \mu_0$)　less…片側検定 ($\mu < \mu_0$)　greater…片側検定 ($\mu > \mu_0$)

である。2標本の場合で等分散が仮定できない場合は、var.equal = F とすればよい。

《教科書の例の場合》

```
> x <- c(5.2,4.9,4.7,4.6,5.0,5.4,4.6,5.0,4.4,4.9,5.4,4.8,4.8,4.3,5.8)
> y <- c(5.6,6.1,6.0,5.9,5.0,5.2,6.6,4.9,6.3,5.7,6.5,5.5,6.9,6.4,6.8)
> c(mean(x),var(x),sd(x))
 [1] 4.9200000 0.1617143 0.4021371
> c(mean(y),var(y),sd(y))
 [1] 5.9600000 0.4040000 0.6356099
> t.test(x,y,alternative="two.side",var.equal = T)

        Two Sample t-test

data:  x and y
t = -5.3553, df = 28, p-value = 1.051e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -1.4378041 -0.6421959
sample estimates:
mean of x mean of y
     4.92      5.96
```

t.testの結果で注目するところはtの値 (t=), 自由度 (df=), 確率 (p-value=) である。特に確率については、直接有意水準の値と比較できるようになっているので、確率の値が有意水準以下であれば、帰無仮説を棄却することができる。

例題の結果はtの値が-5.3553で、Tは自由度28のt分布にしたがっていて、帰無仮説が真の場合このようなデータが得られる確率が$1.051 \times 10^{-5} = 0.00001051$ (約9万5千分の1) と、有意水準である1%より明らかに小さな値なので、帰無仮説を棄却できる。よって、2種類の釘の長さは異なるといえる。

2019年度 数理統計学 D科　小テスト解答用紙　2019.12.20

・平均に関する仮説検定

　Ｘ社，Ｙ社の同容量のバッテリーにモーターを繋ぎ、動作時間を調べた。下記の結果から両社のバッテリーに違いがあるかどうか有意水準１％で検定せよ（母分散は共通)。

Ｘ社	13.5	15.2	13.4	14.7	13.4	13.2
Ｙ社						

（単位：日）

なお、Ｙ社のデータは下記のプログラムで作成すること。

```
> set.seed(学籍番号)
> y <- round(rnorm(5,11.9,0.876),digits=1)
```

途中式の値を書き、t.test は使わないで解答すること。

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２３日（月）１７時頃まで

１つ目のデータ y[1] の値の表《確認用》

学籍番号	y[1]	学籍番号	y[1]	学籍番号	y[1]	学籍番号	y[1]
1825200	12.1	1823038	12.9	1823070	10.6	1823100	11.5
1523051	12.4	1823058	11.2	1823072	13.7	1823102	11.7
1723016	11.1	1823067	12.6	1823076	12.2	1823109	13.4
1823019	12.7	1823068	12.6	1823091	12.2	1823131	13.7