数理統計学　前回の解答

男性の上左側第 1 大臼歯の長さの平均は、これまでの多数の標本の計測により $\mu = 19.6$mm であることがわかっている。同じ位置の女性の臼歯について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)
> y <- round(rnorm(10,18.4,1.2),digits=1)
```

このとき、女性の臼歯の長さは男性の長さと異なるかどうか、有意水準５％で検定せよ。
（ただし、臼歯の長さは母分散未知の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。

女性の臼歯の母平均を μ_w とする。この値が男性の臼歯の母平均 $\mu = 19.6$mm と同じかどうかを証明したいので、帰無仮説と対立仮説は下記のようになる。

帰無仮説：女性で臼歯の長さは男性の臼歯と同じ長さである　　H_0: $\mu = 19.6$

対立仮説：女性で臼歯の長さは男性の臼歯と異なる長さである　　H_0: $\mu \neq 19.6$

（２）標本の抽出

R を使ってデータを作成すると

```
> set.seed(1825200)
> y <- round(rnorm(10,18.4,1.2),digits=1)
> y
 [1] 18.7 17.0 19.7 18.9 18.5 19.6 17.3 17.4 17.8 19.3
```

の１０個のデータが得られる。このデータから標本平均と標本分散を計算すると

```
> c(mean(y),var(y),sd(y))
 [1] 18.4200000  0.9795556  0.9897250
```

であることがわかる。

（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。

帰無仮説を真と仮定するので、母平均を $\mu = 19.6$ とし、さらに母分散が未知なので t で変換すると、

$$t = \frac{18.42 - 19.6}{\frac{0.9897}{\sqrt{10}}} = -3.7703\cdots$$

である。また、$\mathrm{P}(T < -3.131)$ の値は自由度 $10 - 1 = 9$ の t 分布に従うので、`pt(-3.770,9)` より $0.002208\cdots$ $(= 0.221\%)$ である。

（４）判定

この問題は有意水準５％の両側検定なので、（３）で求めた値が２．５％よりも小さい、もしくは９７．５％より大きい場合棄却する。実際の確率が $0.00221\cdots < 0.025$ なので、帰無仮説を棄却する。よって、対立仮説である $\mu \neq 19.6$ が正しいと言える。第 3 者にもわかりやすいように

「女性の臼歯の長さは男性の臼歯と異なる長さである」

と書いても良い。

ちなみに棄却域を使う場合は、`qt(1-0.05/2,9)` の値が 2.262 であることを使う。実際の値は $|t| > 2.262$ なので、帰無仮説を棄却する。

もし帰無仮説が棄却されない場合、帰無仮説と対立仮説のどちらが正しいか判定できないので、判断を保留することになる。

《t.test を使う場合》

対立仮説が "$\neq$" なので、alternative="two.side"を使う。

```
> t.test(y,mu=19.6,alternative="two.side")

        One Sample t-test

data:  y
t = -3.7702, df = 9, p-value = 0.004415
alternative hypothesis: true mean is not equal to 19.6
                          ⋮
```

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ことがでる。

結果一覧

学籍番号	$\bar{y}$	s	t	確率	判定
1825200	18.42	0.9897	-3.770	0.00221	棄却する
1523051	18.45	0.8554	-4.251	0.00107	棄却する
1723016	18.65	1.069	-2.810	0.01018	棄却する
1823019	18.49	1.132	-3.101	0.00635	棄却する
1823038	18.70	0.9707	-2.932	0.00835	棄却する
1823058	18.37	1.115	-3.490	0.00342	棄却する
1823067	18.53	0.9358	-3.616	0.00280	棄却する
1823068	18.99	1.225	-1.574	0.07491	棄却できない
1823070	18.21	1.337	-3.288	0.00471	棄却する
1823072	18.41	1.772	-2.124	0.03131	棄却できない
1823076	18.23	0.9393	-4.612	0.00063	棄却する
1823091	18.02	1.183	-4.223	0.00111	棄却する
1823100	17.91	1.485	-3.599	0.00288	棄却する
1823102	19.26	1.063	-1.012	0.16904	棄却できない
1823109	18.27	1.490	-2.823	0.00998	棄却する
1823131	18.84	1.082	-2.220	0.02676	棄却できない

数理統計学 片側検定と検出力

例．燃費の差 (片側検定)

ガソリンに添加剤を加えた場合、燃費が良くなるかどうかを調べたい。そこで、ガソリンに添加剤を加えた後で走行実験を行った結果、次のようであった。

15.3　13.9　14.7　14.0　15.1　　（単位は km/ℓ）

通常のガソリンでの燃費が 13.5 km/ℓ としたとき、添加剤に効果があったかどうか有意水準１％で検定せよ。ただし、燃費の分布は正規分布とみなしてよいとする。

１）仮説を立てる

帰無仮説 H_0：燃費は変わらない　$\mu = 13.5$　　対立仮説 H_1：燃費が良くなった　$\mu > 13.5$

(正確な帰無仮説は燃費は悪くなったか変わらない《$\mu \leqq 13.5$》)

２）母集団から標本を抽出する。（省略）

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。（b) のみ）

b) 母分散が未知なので、$t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$ で変換すると

$$t = \frac{14.6 - 13.5}{\frac{0.6325}{\sqrt{5}}} = 3.8888$$

このとき、t は自由度 $5 - 1 = 4$ の t 分布に従うので、t 以下になる確率は

`pt(3.8888,4)`$= 0.9911457$

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。

b) ３）で得られた確率が 0.9911457 つまり約９９．１％と９９％以上の大きな値なので、帰無仮説を棄却する。つまり対立仮説が正しいので、燃費は良くなったと判定できる。

ちなみに棄却域を使う場合は、`qt(1-0.01,4)` の値が 3.746947 であることを使い、$t > 3.746947$ なので、帰無仮説を棄却することが出来る。

・`t.test` を使った場合

`t.test(x,mu=帰無仮説の母平均,alternative="two.side"` or `"greater"` or `"less")`

$\mu \neq \mu_0$：`"two.side"`　$\mu > \mu_0$：`"greater"`　$\mu < \mu_0$：`"less"`　を使う

上の例の場合

```
> x <- c(15.3,13.9,14.7,14.0,15.1)
> t.test(x,mu=13.5,alternative="greater")

        One Sample t-test

data:  x
t = 3.8891, df = 4, p-value = 0.008852
alternative hypothesis: true mean is greater than 13.5
95 percent confidence interval:
 13.99702      Inf
sample estimates:
mean of x
     14.6
```

`t.test` の場合 `p-value` の値を直接有意水準と比較すればよく、`p-value = 0.008852` と有意水準の１％より低いので、帰無仮説を棄却することができる。

つまり、対立仮説が正しいと判定するので燃費が良くなったと判定できる。

・両側検定／片側検定と検出力について

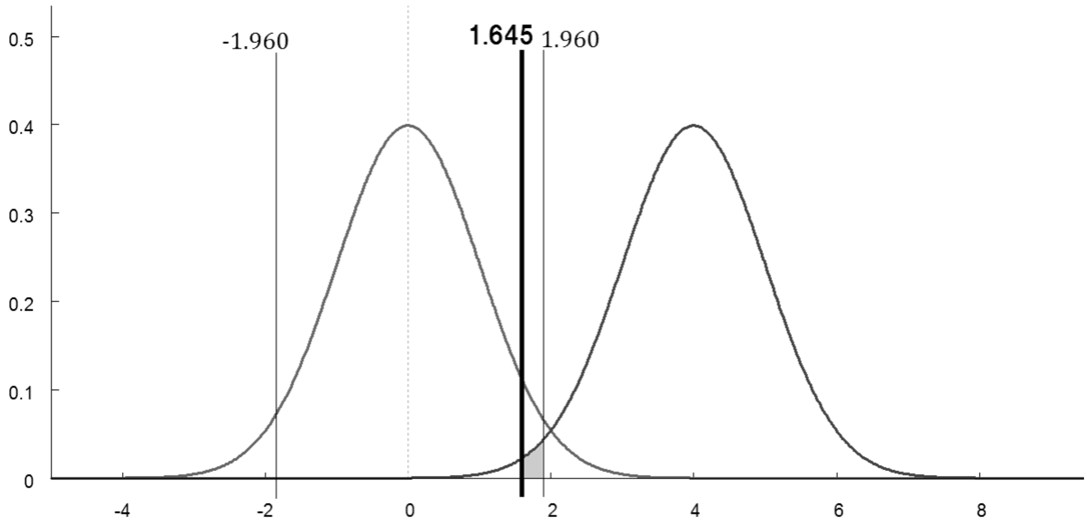
平均に関する検定を行う場合、対立仮説に応じて両側検定（$\mu \neq \mu_0$ の場合）と片側検定（$\mu > \mu_0$ もしくは $\mu < \mu_0$ の場合）を選んで検定を行う必要がある。検定法を間違えて行うと、「有意水準（第１種の過誤の確率）が守られない」や「検出力（１－第２種の確率＝対立仮説が正しいときに帰無仮説を棄却する確率）が低くなる」ことになり、いずれにしても良い検定とは言えない。

・両側検定（$H_1 : \mu \neq \mu_0$）を片側検定（$H_1 : \mu > \mu_0$）と間違えた場合

例えば母分散が既知で有意水準が５％の場合、z の値が $(\infty, -1.960]$ もしくは $[1.960, \infty)$ のとき帰無仮説を棄却し対立仮説が正しいと判定する。しかし間違って片側検定を行った場合、$[1.645, \infty)$ のときに帰無仮説を棄却するため、$[1.645, 1.960)$ の範囲に入ったときに帰無仮説が棄却されない範囲であるが棄却されてしまう（さらに、負の範囲では一切帰無仮説が棄却されない問題が発生する）。本来帰無仮説が棄却されない範囲で帰無仮説を棄却してしまうと、有意水準が保たれていないことになってしまう。

・片側検定（$H_1 : \mu > \mu_0$）を両側検定（$H_1 : \mu \neq \mu_0$）と間違えた場合

例えば母分散が既知で有意水準が５％の場合、z の値が $[1.645, \infty)$ のとき帰無仮説を棄却し対立仮説が正しいと判定する。これを間違って両側検定で行うと $[1.645, 1.960)$ の範囲に入ったときに帰無仮説が棄却される範囲であるが棄却されない。実際に帰無仮説が真（$H_0 : \mu = \mu_0$）と対立仮説が真（$H_1 : \mu = \mu_1 > \mu_0$）の場合の z の分布をグラフにすると下の図のようになる。



片側検定を両側検定で行った場合、灰色の部分の面積分だけ第２種の過誤の確率が上がるため、検出力が下がる（さらに、$(\infty, -1.960]$ の範囲も棄却される範囲となってしまうが、対立仮説 $H_1 : \mu > \mu_0$ が真の場合、上の図からも明らかなように z の値が $(\infty, -1.960]$ になることはほとんどない）。

2019年度 数理統計学 D科　小テスト解答用紙　2019.12.13

・平均に関する仮説検定（片側検定）

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）

<table>
<tr><td rowspan="2">2019年度神奈川工科大学
数理統計学
演習問題</td><td>学科</td><td>学年</td><td>組</td><td>学　籍　番　号</td><td>氏　　　名</td><td rowspan="2"></td></tr>
<tr><td></td><td></td><td></td><td></td><td></td></tr>
</table>

提出先：K3-3309号室前　20番のボックス　提出期限：１２月１６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	23.8	1823038	24.9	1823070	21.6	1823100	23.0
1523051	24.2	1823058	22.5	1823072	26.0	1823102	23.2
1723016	22.4	1823067	24.4	1823076	23.9	1823109	25.6
1823019	24.6	1823068	24.5	1823091	23.9	1823131	26.0