

学籍番号:

公衆衛生学

疫学演習 2019-6-5 & 2019-6-12

氏名:

問題1: 両群間計量データの平均値を比較する

200名の認知症患者を募集し, 認識能力テスト(cognitive test, COG), 及び脳萎縮の進行度 (brain atrophy, 脳体積の平均年間減少率, 単位は%) の検査を全員に行った. COG, 及び脳萎縮のデータは大きいほど認知症の進行度がより進んでいる. また, この200名の参加者から採取した血液検体を利用して, ある遺伝子の変異の有無を検査した. このデータは以下の表でまとめた:

変数	遺伝変異あり (n = 50)		遺伝変異なし (n = 150)	
	平均値 (mean)	標準偏差 (standard deviation)	平均値 (mean)	標準偏差 (standard deviation)
認識能力テスト, COG	69.2	9.0	60.2	9.0
脳萎縮度, atrophy, %/year	0.67	0.21	0.23	0.10

1. 帰無仮説を「遺伝子変異ありと変異なし両群の間に, COGの平均値は等しい」とする. 上記のデータ及び適宜な方法を使って検定せよ. 検定の結果を分かりやすく説明せよ. なお, 分散が等しいと仮定できる場合, 以下の式で両群の共通標準偏差が計算できる:

$$S = \sqrt{\frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}} \quad (1)$$

- ・ S_A : A群の標準偏差;
- ・ n_A : A群の人数;
- ・ S_B : B群の標準偏差;
- ・ n_B : B群の人数;
- ・ S : A群及びB群の共通標準偏差;
- ・ $n_A + n_B - 2$: 分散が等しい時の自由度.

また, EZR で t 値, 自由度 (degree of freedom) を使って P 値を計算する時, 以下のコマンドを利用してください:

```
2*pt(t value, degree of freedom, lower=FALSE)
```

答え

両群の標準偏差は 9.0 と推定され、分散が等しいと仮定できるから、Student の t 検定を行う:

$$T = \frac{\bar{X}_A - \bar{X}_B}{S\sqrt{1/n_A + 1/n_B}}$$

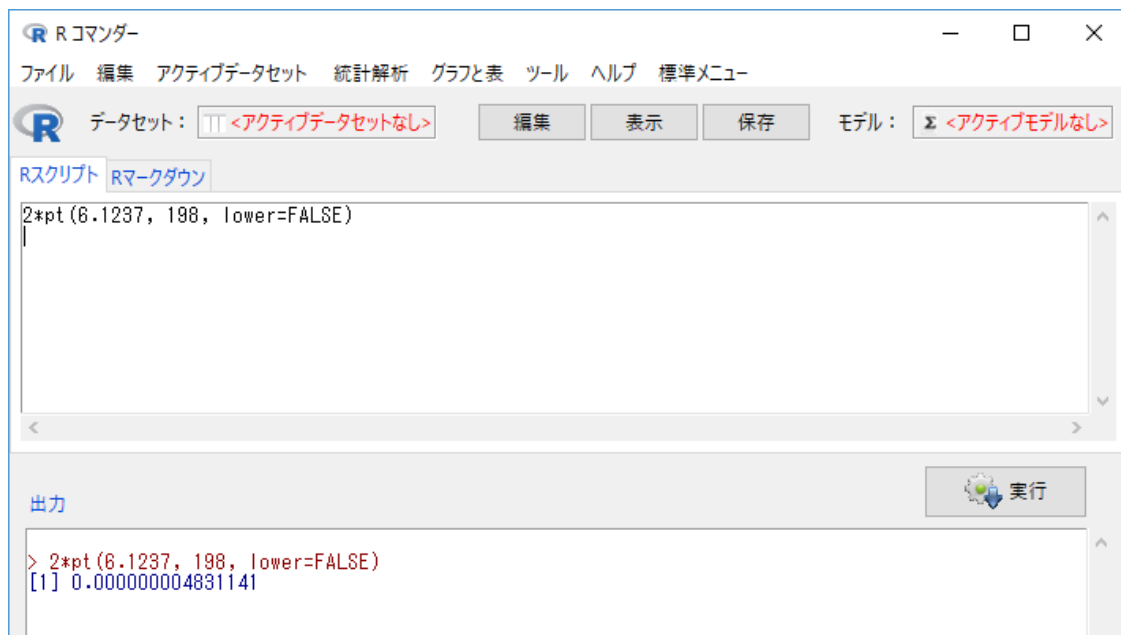
公式(1)により、共通分散/標準偏差を推定する:

$$\begin{aligned} \therefore S &= \sqrt{\frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2}} \\ \therefore S &= \sqrt{\frac{(50 - 1)9^2 + (150 - 1)9^2}{50 + 150 - 2}} = 9 \\ \Rightarrow T &= \frac{\bar{X}_A - \bar{X}_B}{S\sqrt{1/n_A + 1/n_B}} \\ &= \frac{69.2 - 60.2}{9 \times \sqrt{1/50 + 1/150}} \\ &= \frac{9}{9 \times 0.1633} = 6.1237 \end{aligned}$$

自由度 (degree of freedom)は $n_A + n_B - 2 = 198$, P値の計算は EZR を利用する:

```
2*pt(6.1237, 198, lower=FALSE)
```

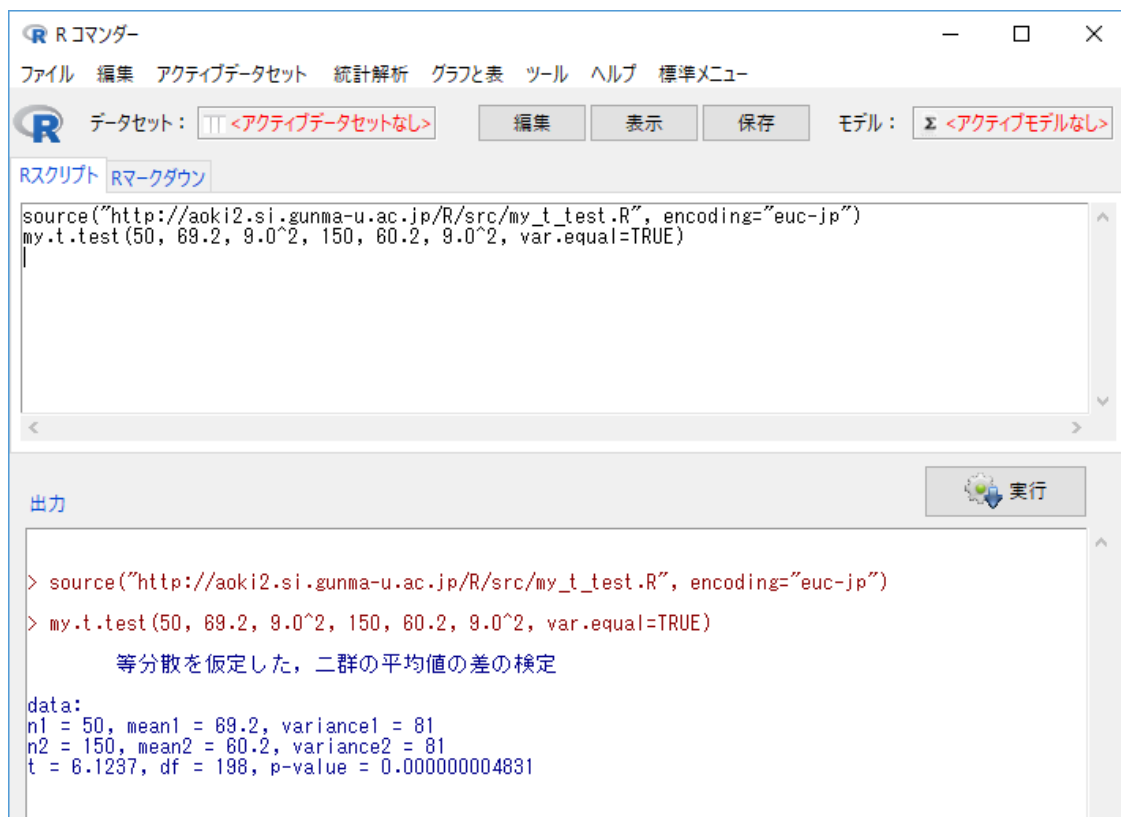
```
## [1] 4.831141e-09
```



以下のコードをRスクリプトに入力して,実行をクリックしてください.自分の検定結果とは一致するかを確認してください.

```
source("http://aoki2.si.gunma-u.ac.jp/R/src/my_t_test.R", encoding="euc-jp")
my.t.test(50, 69.2, 9.0^2, 150, 60.2, 9.0^2, var.equal=TRUE)
```

```
##
##
##
## data:
## n1 = 50, mean1 = 69.2, variance1 = 81
## n2 = 150, mean2 = 60.2, variance2 = 81
## t = 6.1237, df = 198, p-value = 4.831e-09
```



手計算の結果とは一致していると確認できる.

この検定結果は「両群のCOG平均値が等しい」という帰無仮説を棄却するために非常に強い証拠を提供したと言える.

There is strong evidence against the null hypothesis that the means of COG are the same in the two groups.

2. この患者データから、遺伝子変異ありとなしの群の間に脳萎縮度 (atrophy) の比較を 1. と同じ方法で検定してもよいか?どの検定方法を使えば 1. と同じ検定方法を使えるかどうかを判断できるを説明せよ.実際にこの検定方法を行ってください.

なお,EZR で F 値,両群の分散,両群それぞれの自由度 (df) を使って P 値を計算する時に,以下のコマンドを利用してください:

```
2*pf(F value, df in group 1, df in group 2, lower=FALSE)
```

答え

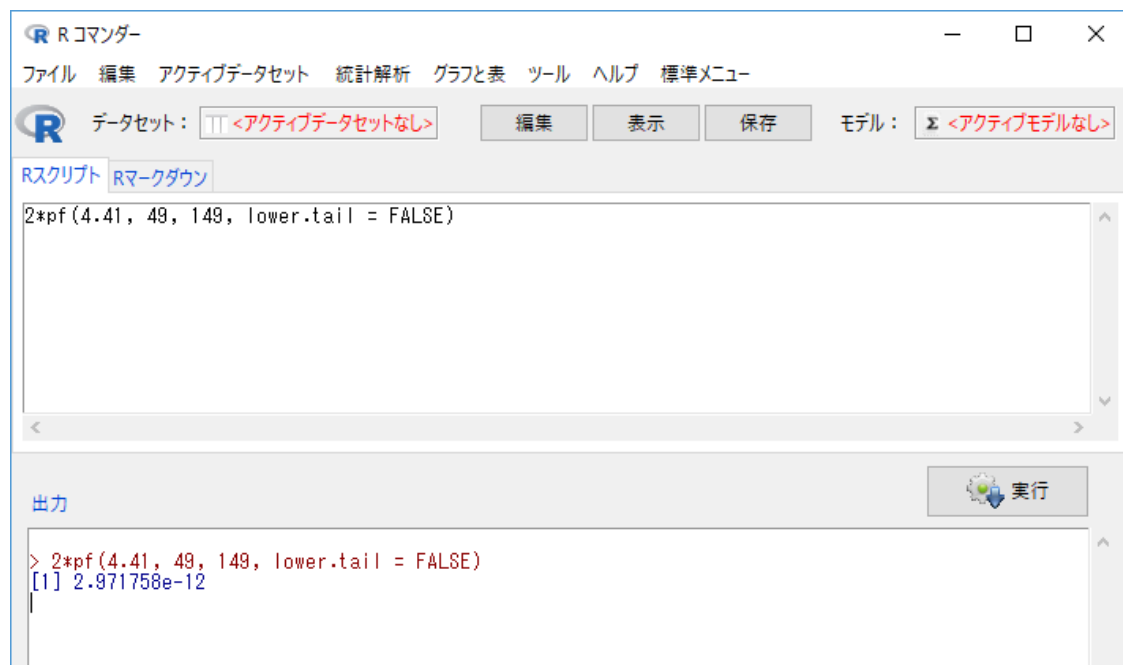
テーブルから両群の標準偏差はそれぞれ 0.21, 0.10 だと推定され,分散 (variance) が等しいという前提が満たされていない.1.の検定方法を使う時には,両群の分散が等しいという前提条件が必須だから,同じ Student t 検定を行うことができない.「両群の分散」が等しいという帰無仮説を検定するには F 検定を利用する:

$$\begin{aligned} F &= \frac{S_A^2}{S_B^2} \\ &= \frac{0.21^2}{0.10^2} \\ &= 4.41 \end{aligned}$$

自由度 (degree of freedom)はそれぞれ $n_A - 1 = 49$; $n_B - 2 = 149$,P値の計算は EZR を利用する:

```
2*pf(4.41, 49, 149, lower.tail = FALSE)
```

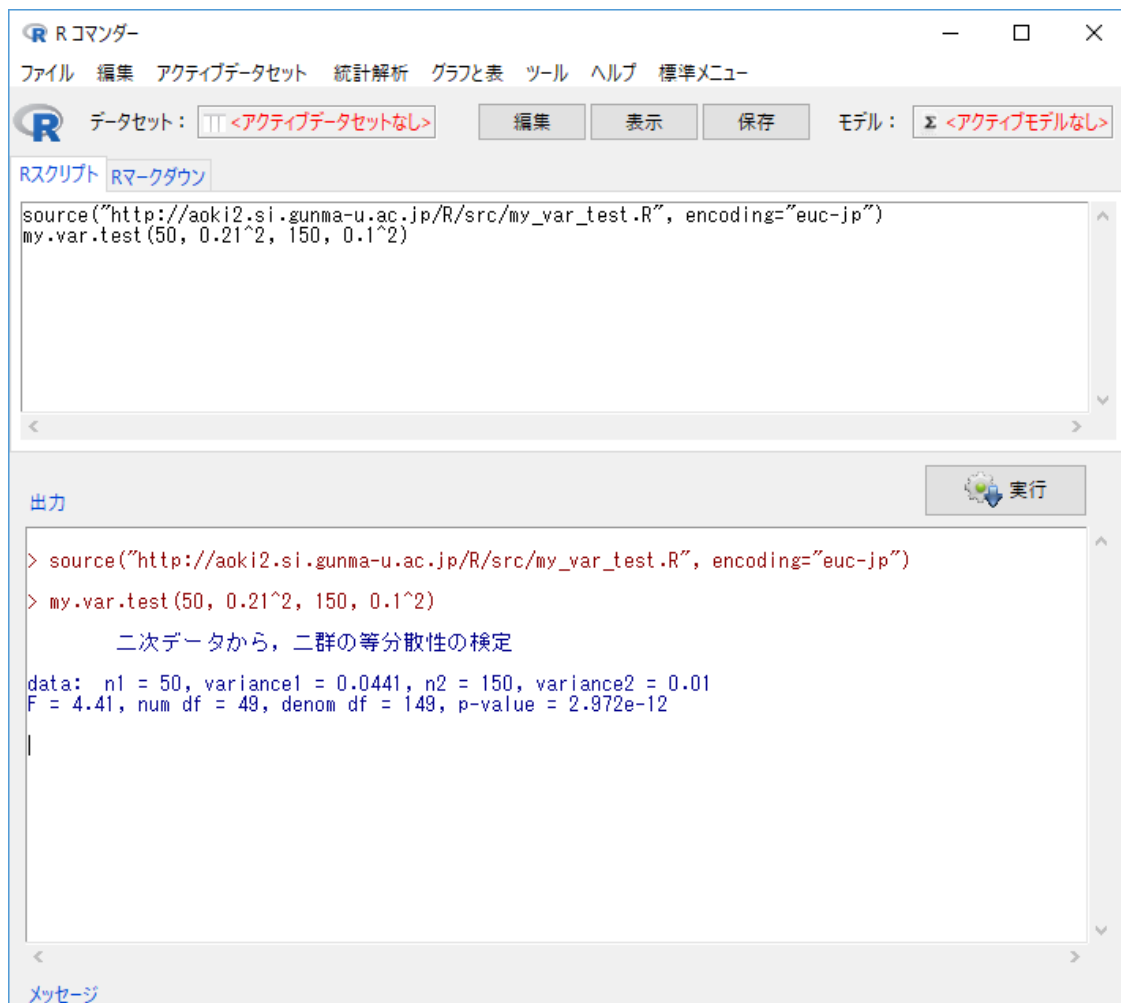
```
## [1] 2.971758e-12
```



以下のコードをRスクリプトに入力して,実行をクリックしてください.自分の検定結果とは一致するかを確認してください.

```
source("http://aoki2.si.gunma-u.ac.jp/R/src/my_var_test.R", encoding="euc-jp")
my.var.test(50, 0.21^2, 150, 0.1^2)
```

```
##
##
##
## data:  n1 = 50, variance1 = 0.0441, n2 = 150, variance2 = 0.01
## F = 4.41, num df = 49, denom df = 149, p-value = 2.972e-12
```



手計算の結果とは一致していると確認できる.

この検定結果は「両群の脳萎縮度の分散が等しい」という帰無仮説を棄却するために非常に強い証拠を提供したと言える.

There is strong evidence against the null hypothesis that the variances of atrophy are the same in the two groups.

3. 2.の結果を踏まえて、帰無仮設「両群の脳萎縮度の平均値が等しい」を検定せよ。なお、両群の分散が等しいという前提が満たされていない時に、自由度(df)の計算式は以下となる:

$$df = \frac{(S_A^2/n_A + S_B^2/n_B)^2}{(S_A^2/n_A)^2/(n_A - 1) + (S_B^2/n_B)^2/(n_B - 1)} \quad (2)$$

また, EZR で t 値, 自由度 (df) を使って P 値を計算する時, 以下のコマンドを利用してください:

```
2*pt(t value, df, lower=FALSE)
```

答え

2.の検定結果から、「両群の脳萎縮度の分散が等しい」という帰無仮説を棄却されたため, Welch の t 検定を採用する。

$$\begin{aligned} \Rightarrow T &= \frac{\bar{X}_A - \bar{X}_B}{\sqrt{S_A^2/n_A + S_B^2/n_B}} \\ &= \frac{0.67 - 0.23}{\sqrt{0.21^2/50 + 0.10^2/150}} \\ &= \frac{0.44}{\sqrt{0.0009486667}} = 14.28551 \end{aligned}$$

自由度は公式(2)により計算できる:

$$\begin{aligned} df &= \frac{(S_A^2/n_A + S_B^2/n_B)^2}{(S_A^2/n_A)^2/(n_A - 1) + (S_B^2/n_B)^2/(n_B - 1)} \\ &= \frac{(0.21^2/50 + 0.10^2/150)^2}{(0.21^2/50)^2/(50 - 1) + (0.10^2/150)^2/(150 - 1)} \\ &= 58.58105 \end{aligned}$$

P値の計算は EZR を利用する:

```
2*pt(14.28551, 58.58105, lower=FALSE)
```

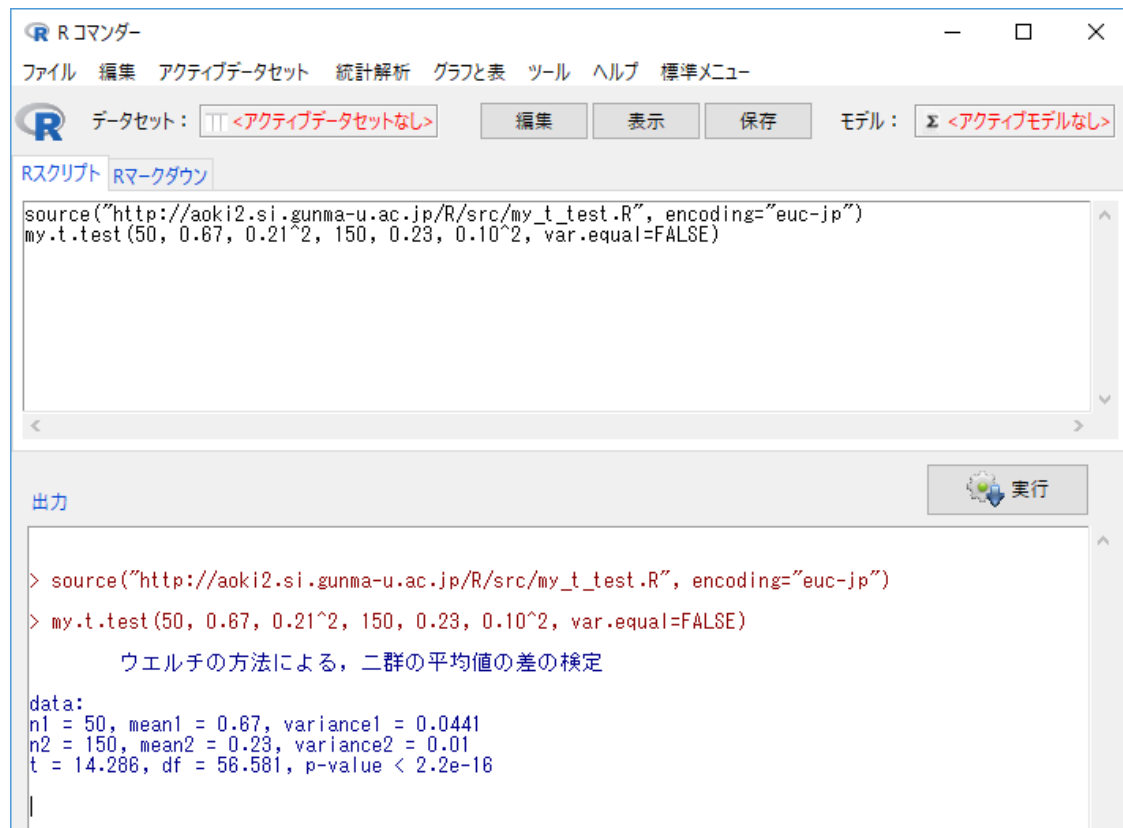
```
## [1] 9.601543e-21
```



以下のコードをRスクリプトに入力して,実行をクリックしてください.自分の検定結果とは一致するかを確認してください.

```
source("http://aoki2.si.gunma-u.ac.jp/R/src/my_t_test.R", encoding="euc-jp")
my.t.test(50, 0.67, 0.21^2, 150, 0.23, 0.10^2, var.equal=FALSE)
```

```
##
##
##
## data:
## n1 = 50, mean1 = 0.67, variance1 = 0.0441
## n2 = 150, mean2 = 0.23, variance2 = 0.01
## t = 14.286, df = 56.581, p-value < 2.2e-16
```



The screenshot shows the R Commander window. The menu bar includes 'ファイル', '編集', 'アクティブデータセット', '統計解析', 'グラフと表', 'ツール', 'ヘルプ', and '標準メニュー'. The 'データセット' field shows '<アクティブデータセットなし>' and the 'モデル' field shows '<アクティブモデルなし>'. The 'Rスクリプト' tab is active, displaying the following code:

```
source("http://aoki2.si.gunma-u.ac.jp/R/src/my_t_test.R", encoding="euc-jp")
my.t.test(50, 0.67, 0.21^2, 150, 0.23, 0.10^2, var.equal=FALSE)
```

The '出力' (Output) window shows the execution results:

```
> source("http://aoki2.si.gunma-u.ac.jp/R/src/my_t_test.R", encoding="euc-jp")
> my.t.test(50, 0.67, 0.21^2, 150, 0.23, 0.10^2, var.equal=FALSE)

ウェルチの方法による、二群の平均値の差の検定

data:
n1 = 50, mean1 = 0.67, variance1 = 0.0441
n2 = 150, mean2 = 0.23, variance2 = 0.01
t = 14.286, df = 56.581, p-value < 2.2e-16
```

手計算の結果とは一致していると確認できる。

この検定結果は「両群の脳萎縮度の平均値が等しい」という帰無仮説を棄却するために非常に強い証拠を提供したと言える。

There is strong evidence against the null hypothesis that the means of atrophy are the same in the two groups.

問題2

参考図書:

- 1.「Rによる保健医療データ解析演習」,中澤 港, (<http://minato.sip21c.org/msb/medstatbookx.pdf>)
- 2.「みんなの医療統計 12日間で基礎理論とEZRを完全マスター!」,新谷 歩.