

確率と統計

中山クラス
第10週

本日の内容

第5章

5.4 t分布を用いた検定(第9週のスライド)

5.5 相関係数の検定

練習問題(1), (2)

第3回レポート作成

5.5 相関係数の検定

帰無仮説:「母集団において相関が0である」と設定

→ 無相関検定

母集団相関係数(母相関: ρ)の検定

標本相関係数 r を次のように変換

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

t は帰無仮説のもとで自由度 $df = n - 2$ の t 分布に従う。

→ t 分布を用いて相関係数の検定を行うことができる。

相関係数の検定(例題)

指導法データ(表2.1, p.38)

「統計テスト1」と「統計テスト2」の得点の相関係数を検定

(1) 帰無仮説と対立仮説(ρ :母相関)

帰無仮説(H_0): 母相関は0である($\rho = 0$)

帰無仮説(H_1): 母相関は0でない($\rho \neq 0$) 両側検定

(2) 検定統計量の選択(r :標本相関)

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

(3) 有意水準の決定

$$\alpha = 0.05 (5\%)$$

(4) 検定統計量の実現値

```
> 統計テスト1<-toukei[,1]
```

```
> 統計テスト2<-toukei[,2]
```

```
> 標本相関<-cor(統計テスト1,統計テスト2)
```

```
> 標本相関
```

```
[1] 0.749659
```

```
> サンプルサイズ<-length(統計テスト1)
```

```
> サンプルサイズ
```

```
[1] 20
```

```
> t分子<-標本相関*sqrt(サンプルサイズ-2)
```

```
> t分子
```

```
[1] 3.180534
```

```
> t分母<-sqrt(1-標本相関^2)
```

```
> t分母
```

```
[1] 0.6618243
```

```
> t統計量<-t分子/t分母
```

```
> t統計量
```

```
[1] 4.805707
```

(5) 帰無仮説の棄却or採択の決定

検定統計量 t は帰無仮説($\rho = 0$)のもとで、自由度 $df = n - 2 = 20 - 2 = 18$ の t 分布に従う。

> qt(0.025, 18) #自由度18の t 分布で下側確率が
0.025となる t の値

[1] -2.100922

> qt(0.975, 18) #自由度18の t 分布で下側確率が
0.975となる t の値

[1] 2.100922

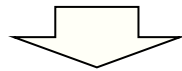
> qt(0.025, 18, lower.tail=FALSE) #自由度18の t 分布で
上側確率が0.025となる t の値

[1] 2.100922

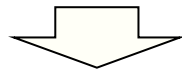
棄却域

$$t < -2.100922, \quad 2.100922 < t$$

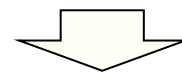
検定統計量の実現値 = $4.805707 > 2.100922$



帰無仮説は棄却される



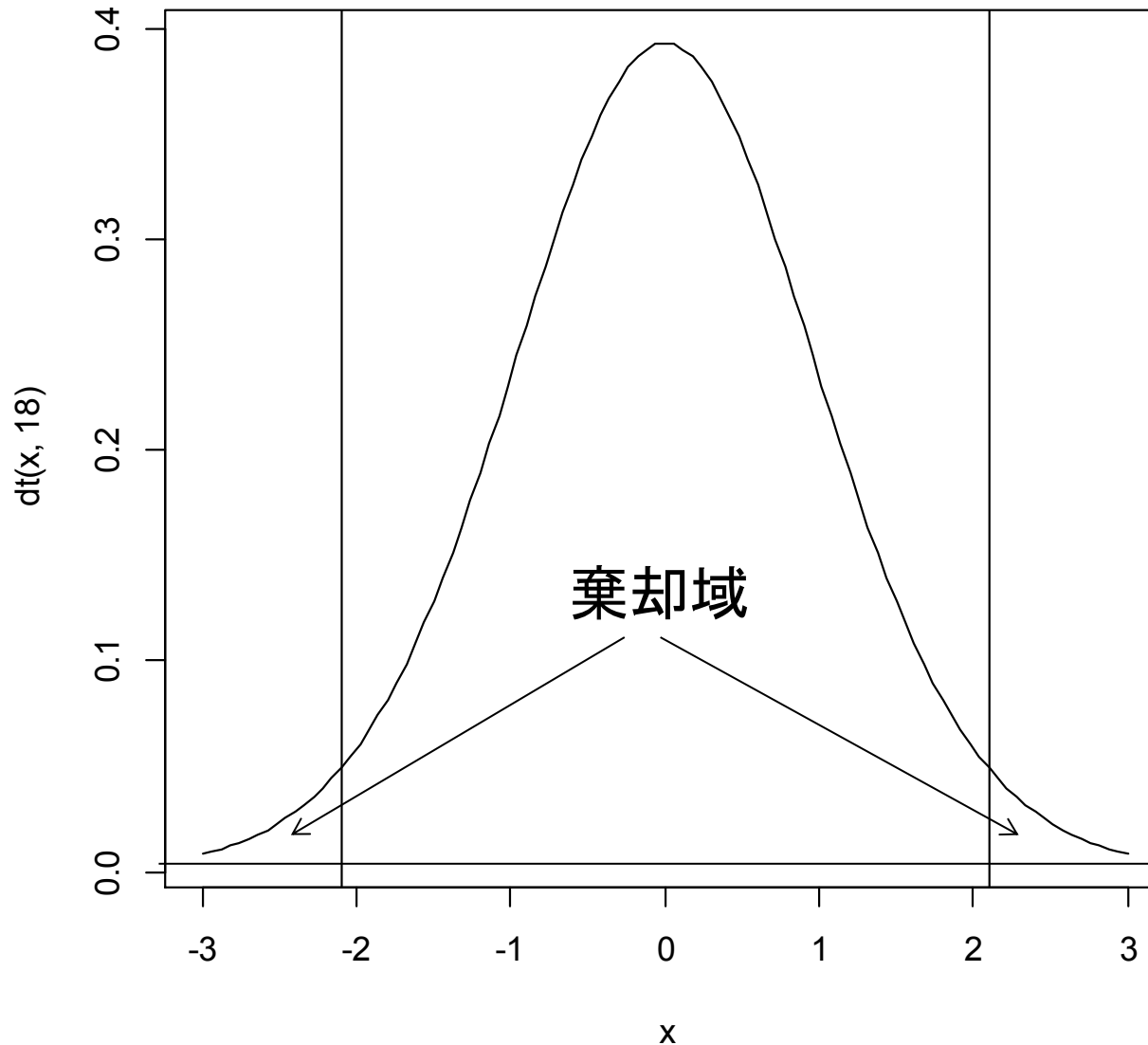
「統計テスト1」の得点と「統計テスト2」の得点の相関係数は5%水準で有意である.



相関がある


```
> curve(dt(x,18), -3,3)
> abline(v=qt(0.025,18))
> abline(v=qt(0.975,18))
```

自由度=18のt分布
有意水準=5%の棄却域
(下側=2.5%, 上側=2.5%)



P値による検定

```
> pt(4.805707, 18, lower.tail=FALSE) #自由度18のt分布で上側確率Prob(t>4.805707)を計算
```

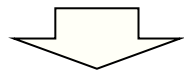
```
[1] 7.08114e-05
```

```
> 2*pt(4.805707, 18, lower.tail=FALSE) #両側検定のために2倍する
```

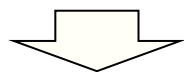
```
[1] 0.0001416228
```

$7.08114e-05 < 0.025$

$0.0001416228 < 0.05$ (有意水準)



t=4.805707となる確率は有意水準(5%)よりも低い



帰無仮説は棄却される

cor.testによる検定

```
> cor.test(統計テスト1, 統計テスト2)
```

Pearson's product-moment correlation

data: 統計テスト1 and 統計テスト2

t = 4.8057, df = 18, p-value = 0.0001416

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.4596086 0.8952048

sample estimates:

cor

0.749659

練習問題(1)

- (A) 教科書の119～122頁に記載されているt分布を用いる方法により検定せよ. t統計量に対する棄却域を求める方法と, p値を用いる方法を試みよ. 但し, 有意水準は5%とする.
- (B) t.test関数を用いて検定を行い, (A)の結果と比較せよ.

練習問題(2)

- (A) 教科書の124～127頁に記載されているt分布を用いる方法により検定せよ. t統計量に対する棄却域を求める方法と, p値を用いる方法を試みよ. 但し, 有意水準は5%とする.
- (B) cor.test関数を用いて検定を行い, (A)の結果と比較せよ.

レポート作成では下記の点に留意

- 帰無仮説と対立仮説を日本語で示せ.
- 検定統計量を文字と数式で示せ.
- 片側検定か両側検定かを説明せよ.
- 有意水準を示せ.
- 検定統計量の実現値と棄却域を示せ.
- P値を示せ.
- 帰無仮説を棄却するか採択するかを理由を付して述べよ.

来週の予定

◆第3回レポート解説

◆第5章

5.6 独立性の検定(カイ二乗検定)

5.7 サンプルサイズの検定結果への影響

◆第4回レポート課題の出題 & 作成

締め切り: 2014年1月10日(金) 17:00

◆練習問題(4), (5)