

# rd-5. 二標本の t 検定 (ウエルチの検定)

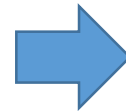
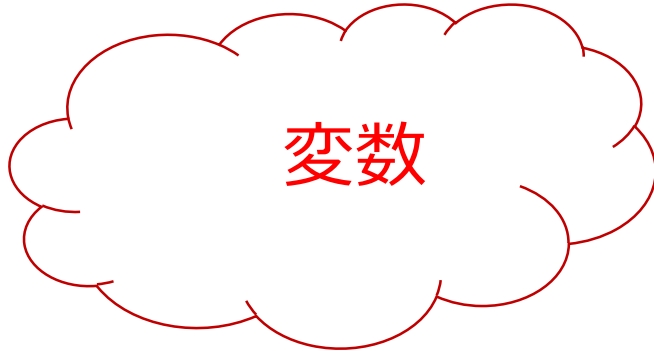
(Rシステムでデータサイエンス演習)

<https://www.kkaneko.jp/cc/rd/index.html>

金子邦彦

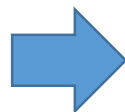
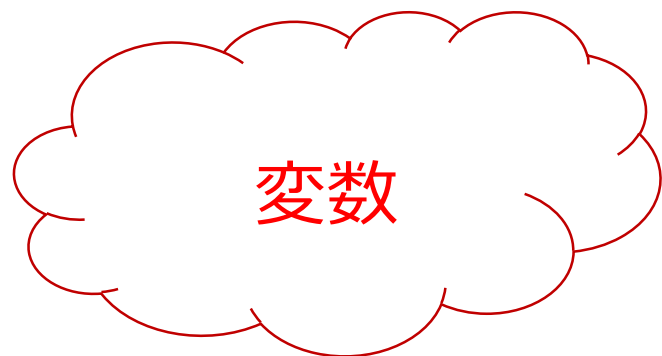


母集団



標本は小数

# 標本の例



128	118
104	110
124	96
85	85
120	109

標本 1   標本 2

平均

112.2   103.6

※各標本は違う

# p 値



Database Lab.

値が変化する何か

<変数>



128  
104  
124  
85  
120

標本 1

100  
106  
89  
89  
105

標本 2

値が変化する何か

<変数>



「この2つの変数の母平均は等しい」と  
思って大丈夫か？

**母平均が等しい** 2つの変数から  
標本 1, 標本 2 が得られる確率は  
0.1541 (p値)  
→ 十分ありえる

# Rでの二標本のt検定（ウェルチの検定）



## ◆ 二標本のt検定（ウェルチの検定）

**t.test( <標本 1 >, <標本 2 >, var.equal=F)**

```
> t.test( c(128, 104, 124, 85, 120), c(100, 106, 89, 89, 105), var.equal=F )
```

```
      welch Two Sample t-test
```

```
data:  c(128, 104, 124, 85, 120) and c(100, 106, 89, 89, 105)
```

```
t = 1.6434, df = 5.6914, p-value = 0.1541
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-7.325827 36.125827
```

```
sample estimates:
```

```
mean of x mean of y
```

```
112.2      97.8
```

p値が表示される

```
t.test( c(128, 104, 124, 85, 120), c(100, 106, 89, 89, 105),  
var.equal=F )
```

# p 値



値が変化する何か  
→

<変数>

128  
104  
124  
85  
120

標本 1

180  
191  
189  
131  
130  
150

↑  
標本 3

値が変化する何か

<変数>

母平均が等しい 2つの変数から  
標本 1 と標本 3 が得られる確率は  
0.006908 (p値)

→ いくらなんでも、偶然とは思えない

# R での二標本の t 検定 (ウェルチの検定)



```
> t.test( c(128, 104, 124, 85, 120), c(180, 190, 189, 131, 130, 150), var.equal=F )
```

```
Welch Two Sample t-test
```

```
data: c(128, 104, 124, 85, 120) and c(180, 190, 189, 131, 130, 150)
```

```
t = -3.5422, df = 8.4866, p-value = 0.006908
```

```
alternative hypothesis: true difference in means is not equal to 0
```

```
95 percent confidence interval:
```

```
-81.35114 -17.58219
```

```
sample estimates:
```

```
mean of x mean of y
```

```
112.2000 161.6667
```

p値が表示される

```
t.test( c(128, 104, 124, 85, 120), c(180, 190, 189, 131, 130, 150), var.equal=F )
```

# 二標本の t 検定 (ウェルチの検定)



Database Lab.

```
> t.test( c(128, 104, 124, 85, 120), c(100, 106, 89, 89, 105), var.equal=F )  
  
welch Two Sample t-test  
  
data: c(128, 104, 124, 85, 120) and c(100, 106, 89, 89, 105)  
t = 1.6434, df = 5.6914, p-value = 0.1541  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
-7.325827 36.125827  
sample estimates:  
mean of x mean of y  
112.2 97.8
```

```
> t.test( c(128, 104, 124, 85, 120), c(180, 190, 189, 131, 130, 150), var.equal=F )  
  
welch Two sample t-test  
  
data: c(128, 104, 124, 85, 120) and c(180, 190, 189, 131, 130, 150)  
t = -3.5422, df = 8.4866, p-value = 0.006908  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
-81.35114 -17.58219  
sample estimates:  
mean of x mean of y  
112.2000 161.6667
```

- p 値 < 0.05 が, 判断の分かれ目の 目安



# 演習問題



1. サイズが5以上の数値データを, 2個準備しなさい

データ1	<数値データを記入>
データ2	<数値データを記入>

2. 1のデータは何ですか? 簡単な説明文を書きなさい

<簡単な説明文>

3. 1のデータについて t検定 (ウエルチの検定) を行い, その p 値を書きなさい

<p値>