

# 検定入門



平均値の差の検定・分散比の検定

～主な関数と出力結果について～

# 平均値の差の検定：成績データ

## □ 成績データ

- 2つの高校A、Bにおいて、3年生の数学の学力に差があるかどうかを調べるため、A高校から8人、B高校から7人を無作為に選んで、実力テストを行ったところ、次のような結果を得た。

A 高校	70	67	81	92	78	62	85	73
B 高校	66	75	48	58	80	57	50	

A高校の平均点と、B高校の平均点に有意な差があるのでしょうか？

# 2標本 t 検定

---

## □ データに対する仮定

### ■ 正規分布を仮定できる

□ さらに等分散を仮定できる ⇒ t 検定

□ 等分散を仮定できない ⇒ Welchの検定

### ■ 正規分布を仮定できないが、分布の「形」は同じとする

□ Wilcoxon検定

## □ データに対する仮定のチェック

### ■ 等分散性の検定 ⇒ F検定

# 手計算であれば...

---

- 手計算であれば、検定統計量  $t$  の計算のために...
  - 等分散性を仮定しても問題ないか...
  - 各群の平均、分散を推定して...
  - 検定統計量を求めて...
  - 教科書等の巻末から棄却限界点を探してきて...
  - 検定統計量の値と比較して...
  - 信頼区間も求めたい...

F 検定 : `var.test()`

t 検定 : `t.test()`

# 「R」で 2標本 t 検定 : 関数の説明

---

## □ 「R」で t 検定を行うための関数

- `t.test(x, y, conf.level=1 -  $\alpha$ , var.equal=T)`

- `x, y` : データ

- `conf.level` : 信頼水準の設定

- `var.equal` : 等分散性を仮定するか

  - T: 仮定する → 2標本 t 検定

  - F: 仮定しない → Welchの検定

- その他の設定については、

  - JIN'S PAGE : <http://www1.doshisha.ac.jp/~mjin/R/>

  - の「Rと検定」をご覧ください

# 「R」で 2標本 t 検定: 例

- 次のプログラムを実行してみましょう

```
A <- c(70, 67, 81, 92, 78, 62, 85, 73)
```

```
B <- c(66, 75, 48, 58, 80, 57, 50)
```

```
t.test(x=A, y=B, conf.level=0.95, var.equal=TRUE)
```

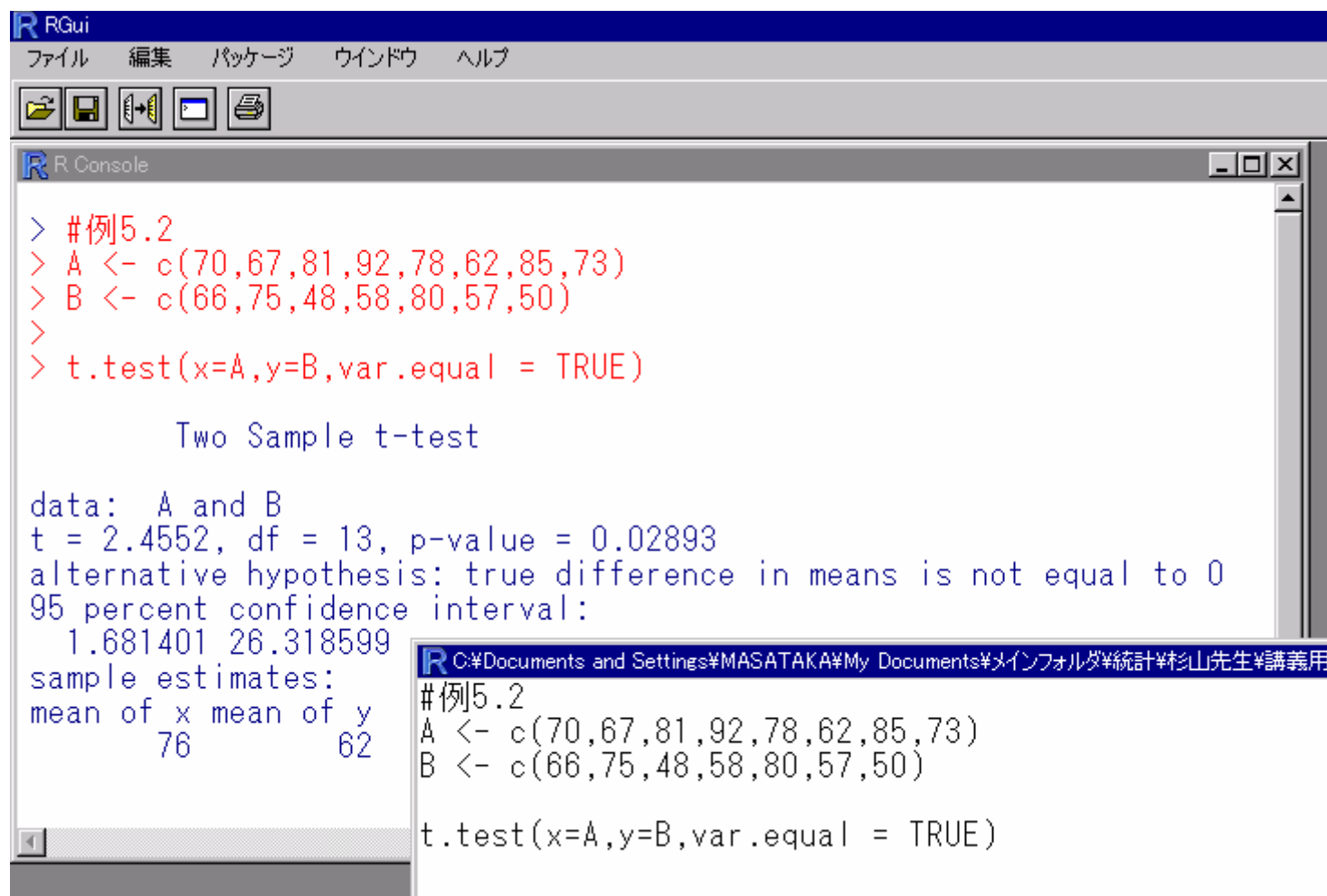
第1群のデータ

第2群のデータ

信頼水準を0.95に設定

等分散性の仮定

# 出力結果：t 検定



```
R RGui
ファイル 編集 パッケージ ウィンドウ ヘルプ

R Console
> #例5.2
> A <- c(70,67,81,92,78,62,85,73)
> B <- c(66,75,48,58,80,57,50)
>
> t.test(x=A,y=B,var.equal = TRUE)

Two Sample t-test

data: A and B
t = 2.4552, df = 13, p-value = 0.02893
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.681401 26.318599
sample estimates:
mean of x mean of y
      76      62

R C:\Documents and Settings\MASATAKA\My Documents\メインフォルダ\統計\杉山先生\講義用
#例5.2
A <- c(70,67,81,92,78,62,85,73)
B <- c(66,75,48,58,80,57,50)

t.test(x=A,y=B,var.equal = TRUE)
```

# 出力結果の説明1: t 検定

The image shows a screenshot of the R GUI interface. The main window is titled "R Console" and contains the following code and output:

```
> #
> A <- c(70,67,81,92,78,62,85,73)
> B <- c(66,75,48,58,80,57,50)
>
> t.test(x=A,y=B,var.equal = TRUE)
```

Annotations in yellow boxes:

- "2標本 t 検定" (Two-sample t-test) points to the `t.test` function call.
- "対立仮説: '2群の平均の差は0ではない'" (Alternative hypothesis: 'The difference in means of two groups is not equal to 0') points to the `var.equal = TRUE` argument.
- "各群の標本平均" (Sample means of each group) points to the output values 76 and 62.

The output text in the console is:

```
Two Sample t-test

data:  A and B
t = 2.4552, df = 13, p-value = 0.02893
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.681401 26.318599
sample estimates:
mean of x mean of y
      76      62
```

At the bottom of the console, the following code is visible:

```
#例5.2
A <- c(70,67,81,92,78,62,85,73)
B <- c(66,75,48,58,80,57,50)
t.test(x=
```



# 出力結果の説明2: t 検定

The image shows a screenshot of the R GUI with the R Console window open. The console displays the following code and output:

```
> #例5.2
> A <- c(70, 67, 81, 92, 78, 62, 85, 73)
> B <- c(58, 80, 57, 50)
> t.test(x=A, y=B, var.equal = TRUE)
```

The output is:

```
Two Sample t-test
data: A and B
t = 2.4552, df = 13, p-value = 0.02893
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.681401 26.318599
sample estimates:
mean of x mean of y
      76         62
```

Annotations in yellow boxes point to specific parts of the output:

- 検定統計量** (Test Statistic) points to `t = 2.4552`.
- P-値** (P-value) points to `p-value = 0.02893`.
- 自由度** (Degrees of Freedom) points to `df = 13`.

The text "Two Sample t-test" is also visible above the output.

# 出力結果の説明3: t 検定

```
> #例5.2
> A <- c(70,67,81,92,78,62,85,73)
> B <- c(66,75,48,58,80,57,50)
>
> t.test(x=A,y=B,var.equal = TRUE)

Two Sample t-test

data:  A and B
t = 2.4552, df = 13, p-value = 0.02893
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.681401 26.318599
sample estimates:
mean of x mean of y
      76      62

#例5.2
A <- c(70,67,81,92,78,62,85,73)
B <- c(66,75,48,58,80,57,50)

t.test(x=A,y=B,var.equal = TRUE)
```

# 分散比の検定：成績データ

## □ 成績データ

- 2つの高校A、Bにおいて、3年生の数学の学力に差があるかどうかを調べるため、A高校から8人、B高校から7人を無作為に選んで、実力テストを行ったところ、次のような結果を得た。

A 高校	70	67	81	92	78	62	85	73
B 高校	66	75	48	58	80	57	50	

等分散を仮定して t 検定を行いました。この仮定は妥当と言えるのでしょうか？

# 「R」で F 検定 : 関数の説明

---

## □ 「R」で F 検定を行うための関数

- `var.test(x, y, conf.level=1 -  $\alpha$ )`

- `x, y` : データ

- `conf.level` : 信頼水準の設定

- その他の設定については、

JIN'S PAGE : <http://www1.doshisha.ac.jp/~mjin/R/>

の「Rと検定」をご覧ください

# 「R」で F 検定: 例

---

- 次のプログラムを実行してみましょう

```
A <- c(70, 67, 81, 92, 78, 62, 85, 73)
```

```
B <- c(66, 75, 48, 58, 80, 57, 50)
```

```
var.test(x=A, y=B, conf.level=0.95)
```

第1群のデータ

第2群のデータ

信頼水準を0.95に設定

# 出力結果:F検定

```
RGui
ファイル 編集 パッケージ ウィンドウ ヘルプ

R Console
> A <- c(70, 67, 81, 92, 78, 62, 85, 73)
> B <- c(66, 75, 48, 58, 80, 57, 50)
>
> var.test(x=A, y=B, conf.level=0.95)

      F test to compare two variances

data:  A and B
F = 0.6626, num df = 7, denom df = 6, p-value = 0.5997
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1163381 3.3915838
sample estimates:
ratio of variances
      0.6626003

R 無題 - R Editor
A <- c(70, 67, 81, 92, 78, 62, 85, 73)
B <- c(66, 75, 48, 58, 80, 57, 50)

var.test(x=A, y=B, conf.level=0.95)
```

# 出力結果の説明1:F検定

The image shows a screenshot of the RGui interface. The main window is titled "RGui" and contains a menu bar (ファイル, 編集, パッケージ, ウィンドウ, ヘルプ) and a toolbar. Below the toolbar is the "R Console" window, which displays the following R code and output:

```
> var.test(x=A, y=B, conf.level=0.95)
data: A and B
F = 0.6626, num df = 7, denom df = 6, p-value = 0.5997
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1163381 3.3915838
sample estimates:
ratio of variances
 0.6626003
```

Annotations in yellow boxes explain parts of the output:

- 分散比較のためのF検定** (F-test for variance comparison) points to the `var.test` function call.
- 対立仮説:"分散比は1ではない"** (Alternative hypothesis: "variance ratio is not equal to 1") points to the text `alternative hypothesis: true ratio of variances is not equal to 1`.
- 標本分散から求めた比** (Ratio calculated from sample variances) points to the `sample estimates: ratio of variances` line.

A secondary window titled "R 無題 - R Editor" is visible at the bottom, showing the data used for the test:

```
A <- c(70, 67, 81, 92, 78, 62, 85, 73)
B <- c(66, ...)
```

# 出力結果の説明2:F検定

The image shows a screenshot of the R GUI with the R Console window open. The console displays the following code and output:

```
> A <- c(70, 67, 81, 92, 78, 62, 85, 73)
> B <- c(66, 75, 48, 58, 80, 57, 50)
var.test(A, y=B, conf.level=0.95)

F test to compare two variances

data: A and B
F = 0.6626, num df = 7, denom df = 6, p-value = 0.5997
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1163381 3.3915838
sample estimates:
ratio of variances
```

Annotations in yellow boxes point to specific parts of the output:

- 検定統計量** (Test Statistic) points to `F = 0.6626`.
- 分子の自由度** (Numerator Degrees of Freedom) points to `num df = 7`.
- 分母の自由度** (Denominator Degrees of Freedom) points to `denom df = 6`.
- P-値** (P-value) points to `p-value = 0.5997`.

A smaller window titled "無題 - R Editor" is visible in the foreground, showing the same code used in the console.



# 出力結果の説明3:F検定

```
RGui
ファイル 編集 パッケージ ウィンドウ ヘルプ

R Console
> A <- c(70, 67, 81, 92, 78, 62, 85, 73)
> B <- c(66, 75, 48, 58, 80, 57, 50)
>
> var.test(x=A, y=B, conf.level=0.95)

      F test to compare two variances

data:  A and B
F = 0.6626, num df = 7, denom df = 6, p-value = 0.5997
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval
 0.1163381 3.3915838
sample estimates:
ratio of variances
      0.6626003

R Editor
A <- c(70, 67, 81, 92, 78, 62, 85, 73)
B <- c(66, 75, 48, 58, 80, 57, 50)

var.test(x=A, y=B, conf.level=0.95)
```

95%信頼区間

95 percent confidence interval  
0.1163381 3.3915838

# 検定で用いる主な関数と引数

---

- 今回は、2標本問題で両側検定を行いましたが、引数を適当に設定することで、1標本問題や片側検定等を扱うこともできます。
- 他にも、適合度検定や相関係数の検定、多群の検定等も、同様に実行できますので参考URLの検定の項を参照してみてください。
  - R-Tips : 「統計解析篇」の検定に関する箇所
  - JIN'S PAGE : 「Rと検定」