

Rによる検定と 推定の実践

内田 治

西澤 英子

Rによる統計的検定と推定 p.35~63

60191055 山崎 大輝

例:1つの平均値

従来の英語のテストの平均点 50点

新しい指導メソッドのもとで英語を学んだ学生を無作為に20人選びテストを受けてもらったところ、結果が以下ようになった。

66,45,41,62,55,62,48,46,70,42

68,76,81,48,47,42,64,52,43,50

検定の手順

①仮説を立てる

帰無仮説 $H_0 : \mu = 50$ (母平均は50である)

対立仮説 $H_1 : \mu > 50$ (母平均は50より大きい)

テスト結果の母平均が高くなったかどうか検定している

検定の手順

②有意水準の設定

有意水準 $\alpha = 0.05$

③使用する統計的手法の決定

今回は「1サンプルのt検定」を実施する。

※Rでは関数`t.test()`で実施することが出来る。

Rの操作

①Rへのデータの読み込み

```
x <- c(66,54,...,52,43,50)
```

②基本統計量の計算

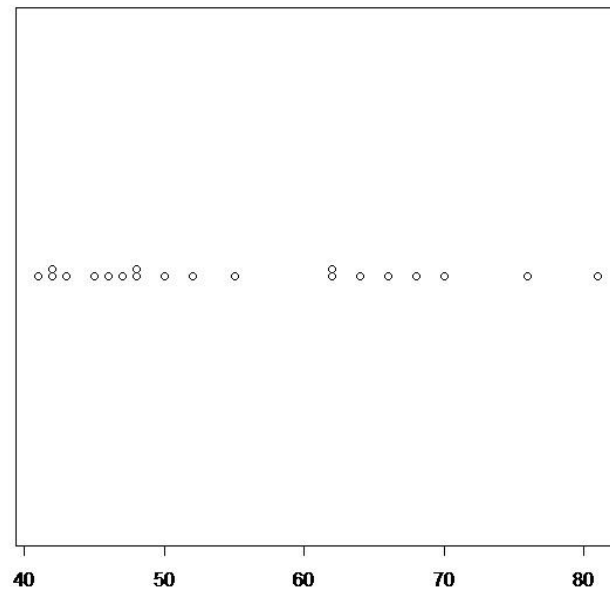
```
mean(x);sd(x);
```

※mean()は平均値,sd()は標準偏差を計算する

Rの操作

③グラフの作成

```
> stripchart(x, method="stack", pch=1)
```



Rの操作

④t.test()の実施

`t.test(x,mu = 50,alternative="greater")`

mu 比較したい母平均の値を設定

alternative 選択した対立仮説の方向を設定

上片側検定 : `alternative="greater"`

下片側検定 : `alternative="less"`

両側検定 : `alternative="two.sided"`

検定の方角について

例

帰無仮説 $H_0 : \mu_A = \mu_B$

対立仮説 $H_1 : \mu_A > \mu_B$

のとき以下のように変形

帰無仮説 $H_0 : \mu_A - \mu_B = 0$

対立仮説 $H_1 : \mu_A - \mu_B > 0$

「 $\mu_A - \mu_B$ greater than 0」という表現になるので

Alternative = “greater”と指定する

検定の結果

One Sample t-test

```
data:  x
t = 1.9674, df = 19, p-value = 0.03196
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 50.65395      Inf
sample estimates:
mean of x
 55.4
```

結果の読み方

One Sample t-test

```
data:  x
t = 1.9674, df = 19, p-value = 0.03196
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 50.65395      Inf
sample estimates:
mean of x
 55.4
```

$p\text{-value} = 0.03196 < \text{有意水準} = 0.05$

→帰無仮説は棄却、母平均は50より大きいことが分かる

信頼区間の算出

One Sample t-test

```
data:  x
t = 1.9674, df = 19, p-value = 0.03196
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 50.65395      Inf
sample estimates:
mean of x
 55.4
```

母平均の95%信頼区間を算出している

例:2つの平均値 (等分散仮定あり)

生活習慣の違いによって2つのグループに分ける。

それぞれのグループのコレステロール値は下記の表のようになる。

A	182	188	186	198	205	206	210	191	200	186	203	168	202	196	218	184	219	204	228	211
B	178	184	182	194	201	202	206	187	196	182	199	164	198	192	214	180	215	200	224	207

母分散の値は不明

A, Bグループは同じ分散

検定の手順

①仮説を立てる

帰無仮説： $\mu_A = \mu_B$ (A,Bグループの母平均は等しい)

対立仮説： $\mu_A \neq \mu_B$ (A,Bグループの母平均は等しくない)

②有意水準の設定

有意水準 $\alpha = 0.05$

③使用する統計的手法の決定

「2つのサンプルのt検定」を実行する。

※Rでは関数`t.test()`で実施することが出来る。

Rの操作

t.test()の実施

```
> t.test(data3_2$A,data3_2$B,  
         paired=F,    ←データが対応付けられているか  
         alternative = “two.sided”,  
         var.equal = T ←2つの群の分散が等しいか  
         )
```

結果の読み方

$p\text{-value} = 0.3894 > \text{有意水準} = 0.05$

この結果より、帰無仮説 H_0 は棄却されず、2グループのコレステロール値の母平均に差があるとはいえないと判断する。

95%信頼区間の結果は

下限 = -5.299871

上限 = 13.299871

0を含んでいるので、検定結果と一致する。

例:2つの平均値 (等分散仮定なし)

血圧の測定方法によって2つのグループに分ける。
それぞれのグループの最低血圧は下記の表のようになる。

A	80	75	66	79	63	75	78	68	82	69	82	79	84	73	84	70	79	81	81	80
B	76	76	64	78	69	87	89	86	83	76	76	77	73	90	91	71	82	81	88	72

母分散の値は不明

A, Bグループの分散は等しくない

検定の手順

①仮説を立てる

帰無仮説： $\mu_A = \mu_B$ (A,Bグループの母平均は等しい)

対立仮説： $\mu_A \neq \mu_B$ (A,Bグループの母平均は等しくない)

②有意水準の設定

有意水準 $\alpha = 0.05$

③使用する統計的手法の決定

「ウェルチの検定」を実行する。

※Rでは関数`t.test()`で実施することが出来る。

Rの操作

t.test()の実施

```
> t.test(data3_3$A,data3_3$B,  
         paired=F,      ←データが対応付けられているか  
         alternative = “two.sided”,  
         var.equal=F ←2つの群の分散が等しいか  
         )
```

結果の読み方

$p\text{-value} = 0.2312 > \text{有意水準} = 0.05$

この結果より、帰無仮説 H_0 は棄却されず、2グループの血圧の母平均に差があるとはいえないと判断する。

95%信頼区間の結果は

下限 = -7.195586

上限 = 1.795586

0を含んでいるので、検定結果と一致する。

例:2つの平均値 (対応あり)

体脂肪を減らすといわれている飲料水の摂取・未摂取によって2つのグループに分ける。

それぞれのグループの摂取前後の腹囲は下記の表のようになる。

A	97.4	90.7	68.1	94.4	84.5	89	92.5	74.5	91.3	76.3	65.9	89.6	84.1	72.4	96	63.3	105.3	64
B	98.8	88.4	66.7	90.2	84.4	90.5	93.8	73.9	90.9	77.4	65.7	88.3	78.9	71.5	96.8	62.4	99.3	64.2

A：当該飲料摂取前

B：当該飲料を毎日摂取し、4週間経過

検定の手順

①仮説を立てる

帰無仮説： $\mu_A = \mu_B$ (摂取前後で変わらない)

対立仮説： $\mu_A > \mu_B$ (摂取後のほうがサイズが減少)

②有意水準の設定

有意水準 $\alpha = 0.05$

③使用する統計的手法の決定

「対応のあるt検定」を実行する。

※Rでは関数`t.test()`で実施することが出来る。

Rの操作

t.test()の実施

```
> t.test(data3_4$A,data3_4$B,  
         paired=T, ←データが対応付けられているか  
         alternative = "two.sided",  
         )
```

結果の読み方

$$p\text{-value} = 0.04219 < \text{有意水準} = 0.05$$

この結果より

帰無仮説 H_0 は棄却

飲料水を摂取することにより腹囲サイズの母平均は減少したと判断する。

例:2つの分散

A, Bグループに数学のテストを受けさせた。

それぞれのグループのテスト結果は下記の表のようになる。

A	82	57	66	75	89	72	67	62	49	74		
B	32	63	44	74	66	52	55	40	79	84	62	89

検定の手順

①仮説を立てる

帰無仮説： $\sigma_A^2 = \sigma_B^2$ (A, Bの母分散は等しい)

対立仮説： $\sigma_A^2 \neq \sigma_B^2$ (A, Bの母分散は等しくない)

②有意水準の設定

有意水準 $\alpha = 0.05$

③使用する統計的手法の決定

「2つの分散の比に関するF検定」を実行する。

※Rでは関数`var.test()`で実施することが出来る。

Rの操作

var.test()の実施

```
> var.test(data3_5$A,data3_5$B,  
            alternative = "two.sided",  
            )
```

結果の読み方

p-value = 0.2181 > 有意水準 = 0.05

この結果より、帰無仮説 H_0 は棄却されず、2グループの母分散は等しくないとはいえないと判断される。

95%信頼区間の結果は

下限 = 0.1204446

上限 = 1.6905750

1を含んでいるので、検定結果と一致する。

例:3つ以上の分散

3箇所の水源地(x)から収集した上水道水に含まれる塩素の量(y)を記録したもの

3つの水源地の塩素の量は等しいと言えるか

x	y
1	1.75
1	0.63
1	0.44
1	1.74
1	1.27
1	1.2
1	1.05

x	y
2	1.12
2	0.77
2	0.3
2	1.16
2	0.29
2	0.11
2	0.06

x	y
3	0.64
3	0.47
3	0.42
3	0.74
3	0.56
3	1.03
3	0.21

検定の手順

①仮説を立てる

帰無仮説 : $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$ (全グループの母分散は等しい)

対立仮説 : (少なくとも1グループの母分散は等しくない)

②有意水準の設定

有意水準 $\alpha = 0.05$

③使用する統計的手法の決定

「バートレットによる分散の同一性の検定」を実行する。

※Rでは関数**bartlett.test()**で実施することが出来る。

Rの操作

bartlett.test()の実施

```
> bartlett.test(y ~ x,data=data3_7)
```

結果の読み方

Bartlett test of homogeneity of variances

data : y by x

Bartlett's K-squared = 2.4343, df = 2, p-value = 0.2961

p-value = 0.2961 > 有意水準 = 0.05

この結果より、帰無仮説 H_0 は棄却されず、3グループの塩素量の母分散は等しくないとはいえないと判断される。