

Rによる 統計的検定と推定

-第5章 多重比較-

システム工学部
情報通信システム学科
60191060 蘆田 誠人

目次

- 多重比較
- 一元配置分散分析
- 多重比較
- テューキーのHSD法
- ボンフェローニ法
- ホルム法
- ダネット法

多重比較とは

- 3つ以上の比較対象があるときに, 2つずつ取り上げて差の有無を検定する方法
- 分散分析で有意差があった場合にどの群とどの群に有意差があるか調べる場合に使用されることが多い

一元配置分散分析の例(1)

今、プラスチック成型部品の素材として、4つの素材を用意する

どの素材が強度に優れているか実験する

一元配置実験のデータ

A1	A2	A3	A4
18.5	17.2	26.1	32.5
16.4	23.9	21.3	32.1
23.4	18.7	23.6	25.9
19.8	23.1	29.7	25.5

一元配置分散分析の例(2)

帰無仮説 $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

対立仮説 H_1 : 少なくとも1つの母平均が他の母平均と異なる

一元配置分散分析の例(3)

① データの入力とデータフレームの作成

```
> x <- read.table("R-sample.txt",header=T,sep=",")  
> A1 <- x[, "A1"]  
> A2 <- x[, "A2"]  
> A3 <- x[, "A3"]  
> A4 <- x[, "A4"]  
> MYDATA <- data.frame(A=factor(c(rep("A1",4),rep("A2",4),rep("A3",4),  
rep("A4",4))),y=c(A1,A2,A3,A4))
```

一元配置分散分析の例(4)

②データの確認

> MYDATA

A y

1 A1 18.5

2 A1 16.4

15 A4 25.9

16 A4 25.5

一元配置分散分析の例(5)

③分散分析の実施

```
> AN1 <- aov(y~A,data=MYDATA)
```

```
> summary(AN1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	3	226.1	75.35	6.426	0.00766 **
Residuals	12	140.7	11.73		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

多重比較の方法(1)

- 検定統計量を調整して全体の有意水準を維持する方法
 - テューキーのHSD法
- p値或いは有意水準を調整して全体の有意水準を維持する方法
 - ボンフェローニ法, ホルム法

多重比較の方法(2)

- テューキーのHSD法

母平均に関してグループ間の全比較を同時に検定

1: 母集団の分布は正規分布

2: すべての群を通して母分散は等しい

- ボンフェローニ法

p値または有意水準を調整してt検定を複数回行う

それぞれの検定の有意水準を低く設定する

例) 10回の検定 → 有意水準を $\frac{0.05}{10} = 0.005$ に変更

多重比較の方法(3)

- ダネット法

ある特定の1つのグループと他のグループの母平均を比較
するときに用いる

- 1: 検定対象となる実験群及び対照群が正規分布
- 2: 母分散が互いに等しい

テューキーのHSD法(1)

```
> x <- read.table("R-sample.txt",header=T,sep=",")
> A1 <- x[, "A1"]
> A2 <- x[, "A2"]
> A3 <- x[, "A3"]
> A4 <- x[, "A4"]
> MYDATA <- data.frame(A=factor(c(rep("A1",4),rep("A2",4),
+ rep("A3",4),rep("A4",4))),y=c(A1,A2,A3,A4))
> TukeyHSD(aov(y~A,data=MYDATA))
```

テューキーのHSD法(2)

Tukey multiple comparisons of means
95% family-wise confidence level

Fit: aov(formula = y ~ A, data = MYDATA)

\$A

A4はA1及びA2と有意差がある

	diff	lwr	upr	p adj
A2-A1	1.200	-5.988545	8.388545	0.9585528
A3-A1	5.650	-1.538545	12.838545	0.1446383
A4-A1	9.475	2.286455	16.663545	0.0096075
A3-A2	4.450	-2.738545	11.638545	0.3036827
A4-A2	8.275	1.086455	15.463545	0.0228708
A4-A3	3.825	-3.363545	11.013545	0.4249118

ボンフェローニ法

```
> DATAY <- c(A1,A2,A3,A4)
> DATAX <- c(rep("A1",4),rep("A2",4),rep("A3",4),rep("A4",4))
> pairwise.t.test(DATAY,DATAX,p.adj="bonferroni")
```

Pairwise comparisons using t tests with pooled SD

data: DATAY and DATAX

	A1	A2	A3
A2	1.000	-	-
A3	0.227	0.546	-
A4	0.012	0.031	0.841

A4はA1及びA2と有意差がある

P value adjustment method: bonferroni

ホルム法

```
> pairwise.t.test(DATAY, DATAX, p.adj="holm")
```

Pairwise comparisons using t tests with pooled SD

data: DATAY and DATAX

	A1	A2	A3
A2	0.629	-	-
A3	0.151	0.273	-
A4	0.017	0.026	0.280

A4はA1及びA2と有意差がある

P value adjustment method: holm

ダネット法

```
> MYDATA <-  
data.frame(A=factor(c(rep("A1",4),rep("A2",4),rep("A3",4),rep("A4",4))),y=c(A1,  
A2,A3,A4))  
  
> res1 <- aov(y~A),data=MYDATA)  
> library(multcomp)  
> res2 <- glht(res1,linfct=mcp(A="Dunnett"))  
> summary(res2)
```

エラーが出たときは
パッケージをインストール

まとめ

多重比較の方法の方法を紹介してきた

- ✓ テューキーのHSB法
- ✓ ボンフェローニ法
- ✓ ホルム法
- ✓ ダネット法