

第五次作业

1 要求：

1. 作业步骤参考下面的文件；
2. 第一题假定服从正态，不要求做正态性检验；均数间多重比较的方法选择一种即可；
3. 第二题不要求做正态检验、方差齐性检验和多重比较；
4. 本次作业要求附上代码（R和SPSS均可）

2 第一题

研究运动是否可以强健骨骼，进行小鼠实验，将30只小鼠随机分配至3个不同处理组，每组10只小鼠，对照组小鼠常规活动，实验1组使小鼠每天跳跃30cm高台10次，实验2组使小鼠每天跳跃 60cm高台10次，8周后，检测小鼠骨密度(mg/cm^3)，数据如下。

请描述和表达数据特征，并判断是否满足方差分析的应用条件？

请比较各组小鼠的骨密度是否存在差别？

g	bd
1	611
1	621
1	614
1	593
1	653
1	600
1	554
1	603
1	569
1	593
2	635
2	605
2	638
2	594
2	599
2	632
2	631
2	588
2	607
2	596
3	650
3	622
3	626
3	628
3	635
3	622
3	643
3	674
3	643
3	650

2.1 第一问

数据特征描述：

此题数据为单因素完全随机设计，共计30个样本，依照运动程度划分为3个处理水平，每一水平各有10个样本。

根据方差分析的应用条件，该样本符合下述应用条件：

1. 各样本是相互独立的随机样本：题目中告知为同质小鼠随机分配至不同处理组。

2. 各样本来自正态分布总体：要求中存在正态假设。
3. 各样本方差相等，即方差齐性：见下文。

综上，该数据满足方差分析的应用条件。

2.1.1 方差齐性检验

采用Levene检验

设立检验假设：

$$\begin{aligned} H_0 : \sigma_1^2 &= \sigma_2^2 = \sigma_3^2 \\ H_1 : \exists i, j \in 1, 2, 3, i \neq j, s.t. \sigma_i^2 &\neq \sigma_j^2 \\ \alpha &= 0.1 \end{aligned} \quad (1)$$

```
# Part1
data <- read.csv("Homework5/Part1.csv")
summary(data)

g <- factor(rep(1:3, c(10, 10, 10)))

# 1.1 数据特征描述
summary(data$bd)

# 进行方差分析

g1 <- subset(data$bd, data$g == 1)
g2 <- subset(data$bd, data$g == 2)
g3 <- subset(data$bd, data$g == 3)

z1 <- abs(g1 - mean(g1))
z2 <- abs(g2 - mean(g2))
z3 <- abs(g3 - mean(g3))
z <- c(z1, z2, z3)
output <- anova(lm(z ~ g))

output$"F value"
output$"Pr(>F)"
```

得到输出结果

```
output$"F value"=0.7331455
output$"Pr(>F)"=0.4897155
```

由于 $Pr = 0.4897 > \alpha = 0.1$ ，故接受 H_0 ，认为方差齐。

2.2 第二问

依题意，此为单因素完全随机设计，故采用单因素方差分析。

由此提出检验假设：

$$\begin{aligned} H_0 : \mu_1 = \mu_2 = \mu_3 \\ H_1 : \exists i, j \in 1, 2, 3, i \neq j, s.t. \mu_i \neq \mu_j, \\ \alpha = 0.05 \end{aligned} \tag{2}$$

进行方差分析：

```
output <- anova(lm(data$bd ~ g))
output
```

得到下表：

方差分析表					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
g	2	7691.5	3845.7	8.3178	0.001533 **
Residuals	27	12483.5	462.4		
Total	29	20175			

可见， $Pr = 0.001533 < \alpha = 0.05$ ，从而拒绝 H_0 ，接受 H_1 ，也即认为至少存在两组，其间均数不等。

进而，我们采用Bonferroni法，来进行均数间多重比较。

```
mean(g1) - mean(g2)
t_12 <- (mean(g1) - mean(g2)) / sqrt(462.4 * (1 / 10 + 1 / 10))
t_12
pt_12 <- 2 * pt(t_12, 27)
pt_12
adjust_pt_12 <- pt_12 * 3
adjust_pt_12

mean(g1) - mean(g3)
t_13 <- (mean(g1) - mean(g3)) / sqrt(462.4 * (1 / 10 + 1 / 10))
t_13
pt_13 <- 2 * pt(t_13, 27)
pt_13
adjust_pt_13 <- pt_13 * 3
adjust_pt_13
```

```
mean(g2) - mean(g3)
t_23 <- (mean(g2) - mean(g3)) / sqrt(462.4 * (1 / 10 + 1 / 10))
t_23
pt_23 <- 2 * pt(t_23, 27)
pt_23
adjust_pt_23 <- pt_23 * 3
adjust_pt_23
```

Bonferroni法进行均数间多重比较表					
	Mean Diff	Mean SE	t	Pr(>t)	Adjust P
1 & 2	-11.4	462.4	0.11058	0.2461708	0.7385124
1 & 3	-38.2	462.4	-3.972276	0.0004762257	0.001428677
2 & 3	-26.8	462.4	-2.786833	0.009622585	0.02886775

```
pairwise.t.test(data$bd, g, p.adjust.method = "bonf")
```

得到下表：

均数间多重比较表		
	1	2
2	0.7384	-
3	0.0014	0.0289

从而得出结论，认为实验1组和对照组间没有显著差异，而实验2组和对照组间、实验2组和实验1组间均具有显著差异。

3 第二题

拟对3个降血脂中药复方制剂与标准降血脂药（安要明）的疗效进行比较。取品种相同和健康的雄性家兔16只，按体重相近的原则配成区组，每个区组4只家兔，共4个区组，将区组内的4只家兔随机分配至4种药物干预组。动物均饲以同样高脂饮食，各组每天分别灌胃服用相应的药物，45天后观察冠状动脉根部动脉粥样硬化斑块大小(cm^3)，实验数据如题表3-7所示。请比较4种药物降脂疗效。

bgc
0.000
0.283
0.114
0.094
0.009
0.196
0.146
0.131
0.003
0.217
0.158
0.065
0.001
0.236
0.159
0.087

依据题意，此为随机区组设计的方差分析，共有4个区组，每组随机分配4种不同处理。

建立对于处理因素的检验假设：

$$\begin{aligned}
 H_0 &: \text{不同药物的降脂疗效相同} \\
 H_1 &: \text{至少两种药物的降脂疗效不同} \\
 \alpha &= 0.05
 \end{aligned}
 \tag{3}$$

建立对于区组因素的检验假设：

$$\begin{aligned}
 H_0 &: \text{区组因素对降脂疗效的影响相同} \\
 H_1 &: \text{区组因素对降脂疗效的影响不完全相同} \\
 \alpha &= 0.05
 \end{aligned}
 \tag{4}$$

进行双因素方差分析（Two-way ANOVA）。

```

# Part2
bgc <- read.csv("Homework5/Part2.csv")$bgc
data <- matrix(bgc, 4, 4, byrow = TRUE)
data

m_treat <- rep(apply(data, 2, mean), 4)
m_treat
m_block <- rep(apply(data, 1, mean), c(4, 4, 4, 4))

```

```

m_block
m_bgc ← mean(bgc)

ss_treat ← sum((m_treat - m_bgc)^2)
ss_treat
ss_block ← sum((m_block - m_bgc)^2)
ss_block
ss_residuals ← sum((bgc - m_treat - m_block + m_bgc)^2)
ss_residuals
ss_total ← sum((bgc - m_bgc)^2)
ss_total
ms_treat ← ss_treat / 3
ms_treat
ms_block ← ss_block / 3
ms_block
ms_residuals ← ss_residuals / 9
ms_residuals
f_treat ← ms_treat / ms_residuals
f_treat
f_block ← ms_block / ms_residuals
f_block
pr_treat ← 1 - pf(ms_treat / ms_residuals, 3, 9)
pr_treat
pr_block ← 1 - pf(ms_block / ms_residuals, 3, 9)
pr_block

# 简便做法
output ← data.frame(
  Y = bgc,
  Treat = factor(rep(1:4, times = 4)),
  Block = factor(rep(1:4, c(4, 4, 4, 4)))
)
output
output.aov ← aov(Y ~ Treat + Block, data = output)
output.aov
summary(output.aov)

```

得到下表

方差分析表

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treat	3	0.11058	0.03686	44.719	9.87e-06 ***
Block	3	0.00035	0.00012	0.141	0.933
Residuals	9	0.00742	0.00082		
Total	15	0.11834			

从而得出结论：

对于药物因素， $Pr(> F) = 9.87e - 06 < \alpha = 0.05$ ，拒绝 H_0 ，接受 H_1 ，认为至少两种药物的降脂疗效不同。

对于体重分组因素， $Pr(> F) = 0.9995667 > \alpha = 0.05$ ，接受 H_0 ，认为区组因素对降脂疗效的影响相同。