Discussion4

Jing Lyu

02/01/2023

For this discussion, we will use PlantGrowth dataset. It contains weights of plants produced under two distinct treatment conditions and a control condition.

One-way ANOVA model in cell-mean form:

$$Y_{i,j} = \mu_i + \epsilon_{i,j}, \ j = 1, \dots, n_i, i = 1, \dots, 3$$

One-way ANOVA model in factor effect form:

$$Y_{i,j} = \mu + \alpha_i + \epsilon_{i,j}, \ j = 1, \dots, n_i, i = 1, \dots, 3$$

where $\{\alpha_i\}$ satisfies that $\sum_{i=1}^{3} n_i \alpha_i = 0$ and $\{\epsilon_{i,j}\}$ are i.i.d. $N(0, \sigma^2)$.

Suppose α_i represent the effect from the three conditions, which are control ($i = 1$), treatment 1 ($i = 2$) and treatment 2 ($i = 3$).

Simultaneous inference

If multiple null hypotheses are tested simultaneously, the probability of falsely rejecting at least one of them increases beyond the pre-specified significance level. Therefore, we need to use simultaneous inference procedures to adjust for multiplicity and control the overall type I error rate.

- Bonferroni method
- Tukey-Kramer method
- Scheffe method

1.Perform pairwise t-tests with Bonferroni's correction:

```
pairwise.t.test(PlantGrowth$weight, PlantGrowth$group, p.adjust.method="bonferroni")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  PlantGrowth$weight and PlantGrowth$group
##
##      ctrl  trt1
## trt1 0.583 -
## trt2 0.263 0.013
##
## P value adjustment method: bonferroni
```

The adjusted p-values for the mean differences in weights between control and treatment 1 is 0.583, between control and treatment 2 is 0.263, between treatment 1 and treatment 2 is 0.013. We can see that the only significant difference is treatment 1 and treatment 2.

2.Multiple Tukey pairwise comparisons:

Only works for pairwise comparisons.

```r
res.aov <- aov(weight ~ group, data = PlantGrowth)
TukeyHSD(res.aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = weight ~ group, data = PlantGrowth)
##
## $group
##             diff        lwr       upr     p adj
## trt1-ctrl -0.371 -1.0622161 0.3202161 0.3908711
## trt2-ctrl  0.494 -0.1972161 1.1852161 0.1979960
## trt2-trt1  0.865  0.1737839 1.5562161 0.0120064
```

lwr, upr: the lower and upper-end points of the confidence interval at 95% (default)

p adj: p-value after multiple comparisons adjustment.

Only the difference between treatment 1 and treatment 2 is significant, as shown by the output, with an adjusted p-value of 0.012.

3.Scheffe method for pairwise and otherwise comparisons:

```r
#install.packages('DescTools')
library(DescTools)
ScheffeTest(res.aov)
```

```
##
##   Posthoc multiple comparisons of means: Scheffe Test
##     95% family-wise confidence level
##
## $group
##             diff     lwr.ci    upr.ci   pval
## trt1-ctrl -0.371 -1.0930531 0.3510531 0.4241
## trt2-ctrl  0.494 -0.2280531 1.2160531 0.2265
## trt2-trt1  0.865  0.1429469 1.5870531 0.0163 *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Suppose we want to test the following hypotheses simultaneously:

(1) Control vs Treatment 1

$$H_0 : \mu_1 = \mu_2 \quad vs \quad H_A : \mu_1 \neq \mu_2$$

   Contrast 1: $L_1 = \mu_1 - \mu_2 = \sum_{i=1}^{3} c_i \mu_i$ with $c_1 = 1, c_2 = -1, c_3 = 0$.

(2) Control vs Treatment 2

$$H_0 : \mu_1 = \mu_3 \quad vs \quad H_A : \mu_1 \neq \mu_3$$

   Contrast 2: $L_2 = \mu_1 - \mu_3 = \sum_{i=1}^{3} c_i \mu_i$ with $c_1 = 1, c_2 = 0, c_3 = -1$.

(3) Control vs the average of Treatment 1 and Treatment 2

$$H_0 : \mu_1 = \frac{\mu_2 + \mu_3}{2} \quad vs \quad H_A : \mu_1 \neq \frac{\mu_2 + \mu_3}{2}$$

   Contrast 3: $L_3 = \mu_1 - \frac{\mu_2 + \mu_3}{2} = \sum_{i=1}^{3} c_i \mu_i$ with $c_1 = 1, c_2 = -1/2, c_3 = -1/2$.

```
contrast.matrix = matrix(c(1,-1,0,1,0,-1,1,-1/2,-1/2),ncol = 3)
contrast.matrix
```

```
##      [,1] [,2] [,3]
## [1,]    1    1  1.0
## [2,]   -1    0 -0.5
## [3,]    0   -1 -0.5
```

```
ScheffeTest(res.aov, contrasts = contrast.matrix)
```

```
##
##    Posthoc multiple comparisons of means: Scheffe Test
##      95% family-wise confidence level
##
## $group
##                   diff      lwr.ci    upr.ci    pval
## ctrl-trt1        0.3710 -0.3510531 1.0930531 0.4241
## ctrl-trt2       -0.4940 -1.2160531 0.2280531 0.2265
## ctrl-trt1,trt2  -0.0615 -0.6868163 0.5638163 0.9681
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Choosing which method?

- All pair comparison, (near) balanced design, then Tukey method is the best.

- If testing m hypothesis, m is small: Bonferroni

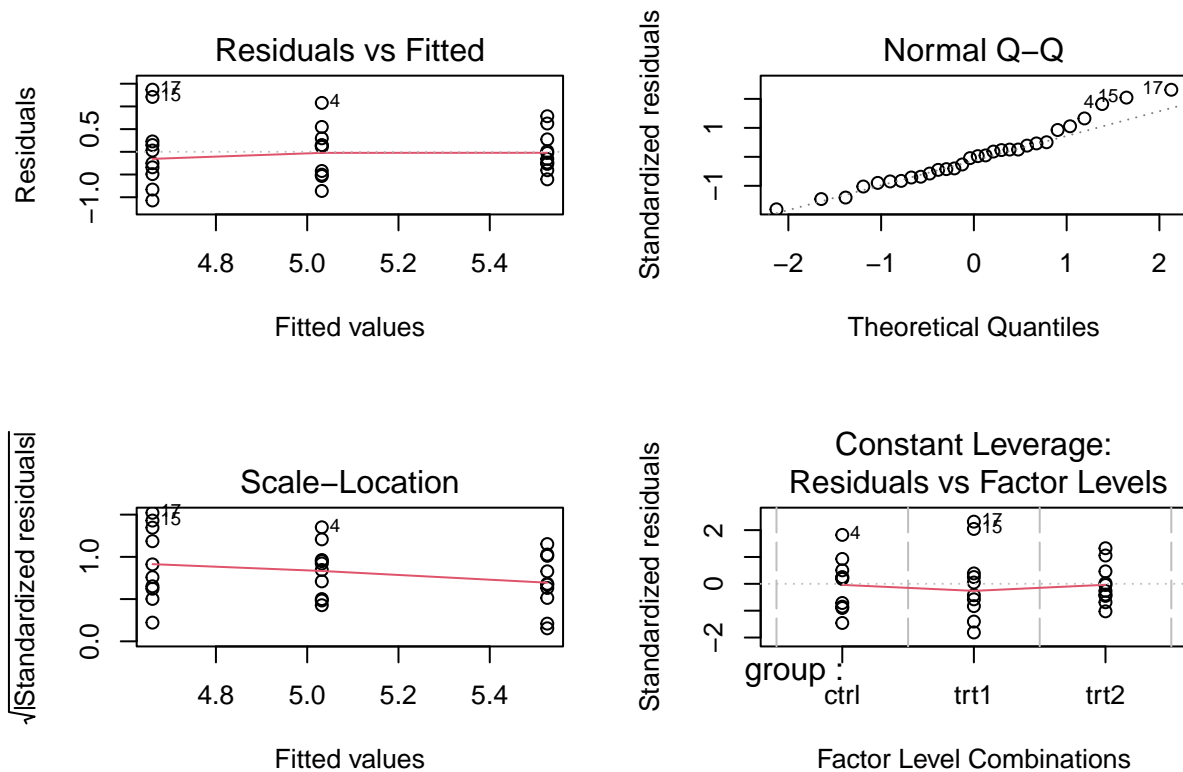- If testing m hypothesis, m is large: Scheffe method.

Model Diagnostics

- Examine the assumption of homogeneity of variance for error terms.
- Examine the assumption of normality for error terms.
- Examine the assumption of independence for error terms.
- Outliers

```
par(mfrow=c(2,2))
plot(res.aov)
```

1.Outliers

In residuals vs. fitted values plot, outliers are discovered in points 17, 15, 4, which can have a significant impact on normality and homogeneity of variance. To meet the test assumptions, it can be beneficial to remove outliers.

2.Variance homogeneity

The homogeneity of variances can also be checked using Bartlett's or Levene's tests. Levene's test is less sensitive to deviations from normal distribution. The `leveneTest()` method is from the car package.

```
bartlett.test(weight ~ group, data = PlantGrowth)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  weight by group
## Bartlett's K-squared = 2.8786, df = 2, p-value = 0.2371
```

```
library(car)
leveneTest(weight ~ group, data = PlantGrowth)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  2  1.1192 0.3412
##       27
```

Both p-values from Bartlett's and Levene's tests are not less than the significance level of 0.05. This indicates that there is no evidence of significantly unequal variances of error terms.

As a result, we can infer that the variations in the different treatment groups are homogeneous.

3.Normality

Normal Q-Q plot can be used to verify the residuals are normally distributed. All points should be about in a straight line.

We can infer normality because all of the points lie roughly along this reference line.

Normality assumption can also be checked by Shapiro-Wilk normality test.

```
# extract the residuals
aov_residuals <- residuals(object = res.aov)
shapiro.test(x = aov_residuals )
```

```
##
##  Shapiro-Wilk normality test
##
## data:  aov_residuals
## W = 0.96607, p-value = 0.4379
```

The test result shows no evidence of normality violation.

Remedies for departures from model assumptions

The Kruskal-Wallis rank-sum test is a non-parametric alternative to one-way ANOVA that can be employed when the model assumptions are violated. It can be used to test whether samples originate from the same distribution.

The null hypothesis is that the medians of all groups are equal, and the alternative hypothesis is they are not all equal.

```
kruskal.test(weight ~ group, data = PlantGrowth)
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  weight by group
## Kruskal-Wallis chi-squared = 7.9882, df = 2, p-value = 0.01842
```

The result shows those groups have significantly different distributions of weights.

Other questions

Can you statistically conclude if there is one group where the mean weights is the highest? Use the most appropriate method to support your statement. Suppose $\alpha = 0.05$**.** You just need to care about whether the group with the highest weights has significantly higher weights than the group with the second highest weights.

Use the Tukey's method to construct simultaneous confidence intervals for all pairwise comparisons:

```
library(dplyr)
alpha=0.05
PlantGrowth%>%group_by(group)%>%summarise(Mean=mean(weight))
```

```
## # A tibble: 3 x 2
##   group  Mean
##   <fct> <dbl>
## 1 ctrl   5.03
## 2 trt1   4.66
## 3 trt2   5.53
```

```
# The largest weight is in treatment 2, and the second largest is in control.
T.ci=TukeyHSD(res.aov,conf.level = 1-alpha)
T.ci$group['trt2-ctrl',]
```

```
##        diff       lwr       upr       p adj
##   0.4940000 -0.1972161  1.1852161  0.1979960
```

Since the CI covers 0, we cannot reject the null hypothesis $H_0 : \mu_1 = \mu_3$. The group with the largest mean weights can be either control or treatment 2. We can't conclude that there is a group with the largest mean weights at the significance level 0.05.

For the above one-way ANOVA model in factor effect form, write down the likelihood and log-likelihood. Since $Y_{i,j} \sim_{iid} \mathcal{N}(\mu + \alpha_i, \sigma^2)$:

Likelihood:

$$l_2(\mu, \alpha_1, \alpha_2, \alpha_3, \sigma) = \prod_{i=1}^{r} \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(Y_{i,j} - \mu - \alpha_i)^2}{2\sigma^2} \right]$$

Log-likelihood:

$$L_2(\mu, \alpha_1, \alpha_2, \alpha_3, \sigma) = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \log \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(Y_{i,j} - \mu - \alpha_i)^2}{2\sigma^2} \right] \right\}$$

subject to the constraint that $\sum_i n_i \alpha_i = 0$.

Please try to simplify the above equations.