7.2 One-Way ANOVA: Inference for Factor Level Means

So far we have discussed the overall F test for comparing the means of the factor levels. However, when we reject the H_0 the only conlcusion we can make is that *at least one of the means is different*, and there is no way that we can identify which mean(s) is different and how much different. So, how are we going to answer the question of which diet leads[18] to a lower blood coagulation time?

To do so, we need to focus our analysis on the factor level means and their relative relationships, using various individual and family hypothesis tests.

7.2.1 Individual Hypotheses Tests/ Confidence Intervals

The quantities we may want to use are:

1. Single factor level means, i.e. μ_i
2. Differences between two factor level means, i.e. $D = \mu_i = \mu_{i'}$
3. Contrasts of factor level means, i.e. $L = \sum_i c_i \mu_i, \sum_i c_i = 0$
4. Linear Combinations of factor level means, i.e. $L = \sum_i c_i \mu_i$

For each one of these cases, we will present the estimator, the distribution of the estimator and we will construct appropriate hypothesis tests and confidence intervals.

Inference for a Single Factor Level Mean

After fitting the ANOVA model using Least Squares, we found that the LS estimators for μ_i are

$$\hat{\mu}_i = \bar{y}_{i.}$$

Under the model assumptions of normality and constant variance, we have that the distribution of $\hat{\mu}_i$ is **Normal** with *mean* and *variance* given by

$$E(\hat{\mu}_i) = \mu_i, \quad Var(\hat{\mu}_i) = \frac{\sigma^2}{n_i}$$

Of course, σ^2 can be estimated via its usual estimator of $RSS/(n-r)$, leading to the estimated variance of $\hat{\mu}_i$ being equal to

$$s^2_{\bar{y}_{i\cdot}} = \frac{1}{n_i} \cdot \frac{RSS}{n-r}$$

Combining the information above, we have

Testing & Confidence Interval for μ_i

Hypothesis Test:

$$H_0 : \mu_i = 0$$
$$H_\alpha : \mu_i \neq 0$$

Test Statistic and Distribution:

$$\frac{\bar{y}_{i\cdot}}{\sqrt{\frac{1}{n_i} \cdot \frac{RSS}{n-r}}} \text{ is distributed as } T_{n-r}$$

Equivalently, we can construct an *(1-α)% Confidence Interval*

$$\mu_i \in \bar{y}_{i\cdot} \pm T_{n-r}(\alpha/2) \sqrt{\frac{1}{n_i} \cdot \frac{RSS}{n-r}}$$

In order to obtain confidence intervals for the factor level means in `R`, we fit the means model and we use the `confint` command.

Inference for the Difference between two Factor Level Means

The **difference** between two factor level means (pairwise} comparison}) is defined as

$$D = \mu_i - \mu_{i'}$$

D is a difference of two means, hence its LS estimator is easily shown to be equal to:

$$\hat{D} = \bar{y}_{i.} - \bar{y}_{i'.}$$

Under the model assumptions, we can show that the distribution of $\hat{D}$ is also **Normal** with mean and variance:

$$E(\hat{D}) = \mu_i - \mu_{i'},\ Var(\hat{D}) = \sigma^2 \ \frac{1}{n_i} + \frac{1}{n_{i'}}$$

with an estimated variance given by

$$s^2_{\hat{D}} = \frac{RSS}{n-r} \ \frac{1}{n_i} + \frac{1}{n_{i'}}$$

Combining the information above, we have

Testing & Confidence Interval for D

Hypothesis Test:

$$\begin{array}{l} H_0 : \mu_i = \mu_{i'} \\ H_\alpha : \mu_i \neq \mu_{i'} \end{array} \Leftrightarrow \quad \begin{array}{l} H_0 : \mu_i - \mu_{i'} := D = 0 \\ H_\alpha : \mu_i - \mu_{i'} := D \neq 0 \end{array}$$

Test Statistic and Distribution:

$$\frac{\hat{D}}{\sqrt{\frac{RSS}{n-r} \ \frac{1}{n_i} + \frac{1}{n_{i'}}}} \text{ is distributed as } T_{n-r}$$

Equivalently, we can construct an *(1-α)% Confidence Interval*

$$D \in \hat{D} + T_{n-r}(\alpha/2) \ \sqrt{\frac{RSS}{n-r} \ \frac{1}{n_i} + \frac{1}{n_{i'}}}$$

Inference for a Contrasts among Factor Level Means

A contrast is a comparison involving *two or more level means*, and is defined as:

$$L = \sum_{i=1}^{r} c_i \mu_i, \quad \text{where} \sum_{i=1}^{r} c_i = 0$$

Since L is a linear combination of factor level means, the LS estimator of L is given by

$$\hat{L} = \sum_{i=1}^{r} c_i \bar{y}_{i.}$$

Under the model assumptions, we can show that the distribution of $\hat{L}$ is also **Normal** with mean and variance:

$$E(\hat{L}) = \sum_{i=1}^{r} c_i \mu_i, \; Var(\hat{L}) = \sigma^2 \sum_{i=1}^{r} \frac{c_i^2}{n_i}$$

The estimated variance of $\hat{L}$ is

$$s^2_{\hat{L}} = \frac{RSS}{n-r} \cdot \sum_{i=1}^{r} \frac{c_i^2}{n_i}$$

Combining the information above, we have

Testing & Confidence Interval for L

Hypothesis Test:

$$H_0 : L = 0$$
$$H_\alpha : L \neq 0$$

Test Statistic and Distribution:

$$\frac{\hat{L}}{\sqrt{\frac{RSS}{n-r} \cdot \sum_{i=1}^{r} \frac{c_i^2}{n_i}}} \text{ is distributed as } T_{n-r}$$

Equivalently, we can construct an *(1-α)% Confidence Interval*

$$L \in \hat{L} + T_{n-r}(\alpha/2)\sqrt{\frac{RSS}{n-r}\cdot\sum_{i=1}^{r}\frac{c_i^2}{n_i}}$$

Inference for a Linear Combination of Factor Level Means

The quantity that we refer to here is

$$L = \sum_{i=1}^{r} c_i\mu_i, \quad \text{no restrictions on } c_i's$$

Point estimator and distribution for this L is exactly the same as before. What changes is the distribution of the test statistic $\hat{L}/s_{\hat{L}}$. Specifically:

Testing & Confidence Interval for L

Hypothesis Test:

$$H_0 : L = 0$$
$$H_\alpha : L \neq 0$$

Test Statistic and Distribution:

$$\left(\frac{\hat{L}}{\sqrt{\frac{RSS}{n-r}\cdot\sum_{i=1}^{r}\frac{c_i^2}{n_i}}}\right)^2 \quad \text{is distributed as } F_{1,n-r}$$

Equivalently, we can construct an *(1-α)% Confidence Interval*

$$L^2 \in \hat{L}^2 + F_{1,n-r}\sqrt{\frac{RSS}{n-r}\cdot\sum_{i=1}^{r}\frac{c_i^2}{n_i}}$$

7.2.2 Family Hypotheses Tests/ Confidence Intervals

The confidence coefficient $1-\alpha$ for the estimation procedures described is a statement confidence coefficient and applies only to a particular estimate, not to a series of estimates. Similarly the specified Type I error rate α applies only to a particular test and not to a series of tests. Therefore, there is a need for **family** hypothesis tests and confidence intervals.

Remark: When we perform a family hypothesis test or when we construct a confidence interval for a mean/difference/contrast, the point estimators, the variance estimators and the test statistics remain the same. What changes is the **distribution** of the test statistic and as a result the “multiplier” inn the confidence interval.

Below, we discuss the most popular distributions used when performing family testing in ANOVA:

Bonferroni Correction

The Bonferroni Correction is easily applied in practice and it works well when the family of interest is a particular set of pairwise comparisons, contrasts, or linear combinations that is specified by the user. However, these joinnt intervals become very wide (and as a consequence uninformative) when we have too many simultaneous tests.

The idea is the same as the one we have seen before in the context of regression: Suppose m is the number of statements in the family. In order to control the family wise error rate to be α, we need to reduce the error rate for each individual comparison to be α/m. That is we need to increase the significance level from $(1-\alpha)$ to $(1-\alpha/m)$.

In `R`, we can obtain the p-values for the Bonferroni corrections for pairwise differences using the `pairwise.t.test` command.

Blood Coagulation Experiment

```
pairwise.t.test(coagulation$coag, coagulation$diet, p.adjust.method = "bonferroni
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  coagulation$coag and coagulation$diet
##
##   A       B       C
## B 0.02282 -       -
## C 0.00108 0.95266 -
## D 1.00000 0.00518 0.00014
##
## P value adjustment method: bonferroni
```

Based on the Bonferroni, we find that Diets A and C, Diets B and D and Diets C and D are statistically different leading to different coagulation times. On the other hand, Diets A and D are statistically "the same". Also, Diets B and C also lead to similar coagulation times.

Tukey's Paired Comparison Procedures

Tukey's Paired Comparison test are best when the family of interest is the set of **all pairwise comparisons** of factor level means, i.e. it consists of estimates of **all** pairs $D = \mu_i - \mu_{i'}$.

A confidence interval is given by

$$D \in \hat{D} + \frac{q(\alpha/2;\ r,\ n-r)}{\sqrt{2}}\ s_{\hat{D}},$$

where $q(\alpha/2;\ r,\ n-r)$ refers to the $\alpha/2$ upper quantile of the studentized range for r means and $n-r$ degrees of freedom. The coverage probability is exact when the sample sizes in each group are identical and is approximate otherwise.

To obtain Tukey family CIs for all pairwise comparisons in `R`, we use the `TukeyHSD` command.

Blood Coagulation Experiment

Tukey Simultaneous 95% CI for all mean differences is given by:

```
TukeyHSD(aov(coag~diet, coagulation), data=coagulation)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = coag ~ diet, data = coagulation)
##
## $diet
##     diff         lwr       upr     p adj
## B-A    5   0.7245544  9.275446 0.0183283
## C-A    7   2.7245544 11.275446 0.0009577
## D-A    0  -4.0560438  4.056044 1.0000000
## C-B    2  -1.8240748  5.824075 0.4766005
## D-B   -5  -8.5770944 -1.422906 0.0044114
## D-C   -7 -10.5770944 -3.422906 0.0001268
```

The conclusion basedonn Tuley's tests is the same as the one with Bonferroni. Using this output and the actual computed differences, with a 95% family confidence level, we also conclude that:

- Diet B leads to higher coagulation time when compared to Diet A
- Diet C leads to higher coagulation time when compared to Diet A
- Diet D leads to lower coagulation time when compared to Diet B
- Diet D leads to lower coagulation time when compared to Diet C

Scheffe's Method for Contrasts

The Scheffe's Method performs better when the family of interest is a set of contrasts among the factor level means:

$$L = \quad c_i\mu_i, \text{ where } \quad c_i = 0$$

A confidence interval is given by

$$L \in \hat{L} + (r-1)\ F_{r-1,n-r}(\alpha)\ s_{\hat{L}}$$

To obtain Scheff'e family CIs for all pairwise comparisons in `R`, we use the `ScheffeTest` in the `DescTools` library.

Blood Coagulation Experiment

If we want Scheffe's method for all pairwise comparisons, we have

```
library(DescTools)
```

```
##
## Attaching package: 'DescTools'
```

```
## The following objects are masked from 'package:caret':
##
##     MAE, RMSE
```

```
ScheffeTest(aov(coag~diet, coagulation))
```

```
##
##   Posthoc multiple comparisons of means: Scheffe Test
##     95% family-wise confidence level
##
## $diet
##     diff     lwr.ci    upr.ci    pval
## B-A    5   0.342883  9.657117 0.03233 *
## C-A    7   2.342883 11.657117 0.00210 **
## D-A    0  -4.418129  4.418129 1.00000
## C-B    2  -2.165452  6.165452 0.55494
## D-B   -5  -8.896424 -1.103576 0.00876 **
## D-C   -7 -10.896424 -3.103576 0.00031 ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Using a 5% significance level, the results are the same as with the other methods.

If we want to obtain intervals for specific contrasts, such as

$$L_1 = \mu_A - \frac{1}{2}\mu_B - \frac{1}{2}\mu_C$$

and

$$L_2 = \mu_B - \frac{1}{2}\mu_C - \frac{1}{2}\mu_D$$

then we can specify this in the `contrasts` argument as follows:

```
ScheffeTest(aov(coag~diet, coagulation), contrasts=matrix(c(1,-0.5,-0.5,0,
                                                            0,1,-0.5,-0.5), r
```

```
##
##   Posthoc multiple comparisons of means: Scheffe Test
##     95% family-wise confidence level
##
## $diet
##        diff     lwr.ci    upr.ci   pval
## A-B,C -6.0 -10.165452 -1.834548 0.0032 **
## B-C,D  1.5  -2.031434  5.031434 0.6482
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```