Lab 4 - Multiple Comparisons

Solutions

Goals

The goal in this lab is to practice finding familywise confidence intervals for multiple comparisons.

Loading packages

Here are some packages with functionality you may need for this lab. Run this code chunk now.

```
library(readr)
library(ggplot2)
library(gridExtra)
library(mosaic)
library(dplyr)
library(gmodels)
options(pillar.sigfig = 10)  # print 10 significant digits in summarize output
longevity <- read_csv("http://www.evanlray.com/data/sleuth3/ex0501_longevity.csv") %>%
    mutate(Diet = factor(Diet, levels = c("NP", "N/N85", "N/R50", "R/R50", "lopro",
        "N/R40")))
nrow(longevity)
```

```
## [1] 349
```

Find the multiplier that would be used for Bonferroni 95% familywise intervals for 5 comparisons based on this data set.

```
bonferroni_mult <- qt(1-0.05/(2*5), df=nrow(longevity)-6)
bonferroni_mult
```

```
## [1] 2.590239
```

Find the multiplier that would be used for Scheffe 95% familywise intervals for 5 comparisons based on this data set.

```
scheffe_mult <- sqrt((6-1)*qf(1-0.05, df1=6-1, df2=nrow(longevity)-6))
scheffe_mult
```

```
## [1] 3.346868
```

Find familywise 95% confidence intervals

The code below finds individual 95% confidence intervals for the 5 differences in group means the researchers planned for this study. For each, add calculations of appropriately adjusted Bonferroni and Scheffe familywise intervals.

```r
## Fit linear model
lm_longevity <- lm(Lifetime ~ Diet, data=longevity)
```

(a) $H_0 : \mu_2 = \mu_3$ vs $H_A : \mu_2 \neq \mu_3$. Are the population mean lifetimes the same for the N/N85 and N/R50 groups?

```r
## Individual 95% confidence interval:
ind_mu2_mu3 <- fit.contrast(lm_longevity, "Diet", c(0, 1, -1, 0, 0, 0), conf.int = 0.95)
print("Individual 95% confidence interval:")
```

```
## [1] "Individual 95% confidence interval:"
```

```r
ind_mu2_mu3
```

```
##                          Estimate Std. Error   t value     Pr(>|t|)
## Diet c=( 0 1 -1 0 0 0 ) -9.605955   1.187682 -8.087982 1.057397e-14
##                          lower CI  upper CI
## Diet c=( 0 1 -1 0 0 0 ) -11.94201 -7.269897
## attr(,"class")
## [1] "fit_contrast"
```

```r
## Bonferroni 95% familywise interval:
print("Bonferroni 95% familywise interval:")
```

```
## [1] "Bonferroni 95% familywise interval:"
```

```r
fit.contrast(lm_longevity, "Diet", c(0, 1, -1, 0, 0, 0), conf.int = 1 - 0.05/5)
```

```
##                          Estimate Std. Error   t value     Pr(>|t|)
## Diet c=( 0 1 -1 0 0 0 ) -9.605955   1.187682 -8.087982 1.057397e-14
##                          lower CI  upper CI
## Diet c=( 0 1 -1 0 0 0 ) -12.68234 -6.529574
## attr(,"class")
## [1] "fit_contrast"
```

```r
## OR
c(ind_mu2_mu3[1, 1] - bonferroni_mult * ind_mu2_mu3[1, 2], ind_mu2_mu3[1, 1] +
    bonferroni_mult * ind_mu2_mu3[1, 2])
```

```
## [1] -12.682336  -6.529574
```

```r
## Scheffe 95% familywise interval:
print("Scheffe 95% familywise interval:")
```

```
## [1] "Scheffe 95% familywise interval:"
```

```
c(ind_mu2_mu3[1, 1] - scheffe_mult * ind_mu2_mu3[1, 2], ind_mu2_mu3[1, 1] +
    scheffe_mult * ind_mu2_mu3[1, 2])
```

```
## [1] -13.580971  -5.630939
```

No, there is evidence that the mean lifetimes are not the same for the **N/N85** and **N/R50** groups.

(b) $H_0 : \mu_3 = \mu_4$ **vs** $H_A : \mu_3 \neq \mu_4$**. Are the population mean lifetimes the same for the N/R50 and R/R50 groups?**

```
## Individual 95% confidence interval:
ind_mu3_mu4 <- fit.contrast(lm_longevity, "Diet", c(0,0,1,-1,0,0), conf.int = 0.95)
print("Individual 95% confidence interval:")
```

```
## [1] "Individual 95% confidence interval:"
```

```
ind_mu3_mu4
```

```
##                           Estimate Std. Error   t value  Pr(>|t|) lower CI
## Diet c=( 0 0 1 -1 0 0 ) -0.5885312    1.19355 -0.493093 0.6222624 -2.93613
##                          upper CI
## Diet c=( 0 0 1 -1 0 0 ) 1.759068
## attr(,"class")
## [1] "fit_contrast"
```

```
## Bonferroni 95% familywise interval:
print("Bonferroni 95% familywise interval:")
```

```
## [1] "Bonferroni 95% familywise interval:"
```

```
fit.contrast(lm_longevity, "Diet", c(0,0,1,-1,0,0), conf.int = 1-0.05/5)
```

```
##                           Estimate Std. Error   t value  Pr(>|t|)
## Diet c=( 0 0 1 -1 0 0 ) -0.5885312    1.19355 -0.493093 0.6222624
##                           lower CI upper CI
## Diet c=( 0 0 1 -1 0 0 ) -3.680111 2.503048
## attr(,"class")
## [1] "fit_contrast"
```

```
## Scheffe 95% familywise interval:
print("Scheffe 95% familywise interval:")
```

```
## [1] "Scheffe 95% familywise interval:"
```

```
c(ind_mu3_mu4[1,1]-scheffe_mult*ind_mu3_mu4[1,2], ind_mu3_mu4[1,1]+scheffe_mult*ind_mu3_mu4[1,2])
```

```
## [1] -4.583186  3.406123
```

Yes, the population mean lifetimes are the same for the **N/R50** and **R/R50** groups.

(c) $H_0 : \mu_3 = \mu_6$ vs $H_A : \mu_3 \neq \mu_6$. Are the population mean lifetimes the same for the N/R50 and N/R40 groups?

```
## Individual 95% confidence interval:
ind_mu3_mu6 <- fit.contrast(lm_longevity, "Diet", c(0,0,1,0,0,-1), conf.int = 0.95)
print("Individual 95% confidence interval:")
```

```
## [1] "Individual 95% confidence interval:"
```

```
ind_mu3_mu6
```

```
##                          Estimate Std. Error   t value   Pr(>|t|)
## Diet c=( 0 0 1 0 0 -1 ) -2.819484   1.171097 -2.407558 0.01658711
##                          lower CI   upper CI
## Diet c=( 0 0 1 0 0 -1 ) -5.122919 -0.5160481
## attr(,"class")
## [1] "fit_contrast"
```

```
## Bonferroni 95% familywise interval:
print("Bonferroni 95% familywise interval:")
```

```
## [1] "Bonferroni 95% familywise interval:"
```

```
fit.contrast(lm_longevity, "Diet", c(0,0,1,0,0,-1), conf.int = 1-0.05/5)
```

```
##                          Estimate Std. Error   t value   Pr(>|t|)
## Diet c=( 0 0 1 0 0 -1 ) -2.819484   1.171097 -2.407558 0.01658711
##                          lower CI  upper CI
## Diet c=( 0 0 1 0 0 -1 ) -5.852904 0.2139368
## attr(,"class")
## [1] "fit_contrast"
```

```
## Scheffe 95% familywise interval:
print("Scheffe 95% familywise interval:")
```

```
## [1] "Scheffe 95% familywise interval:"
```

```
c(ind_mu3_mu6[1,1]-scheffe_mult*ind_mu3_mu6[1,2], ind_mu3_mu6[1,1]+scheffe_mult*ind_mu3_mu6[1,2])
```

```
## [1] -6.738990  1.100023
```

Yes, the population mean lifetimes are the same for the **N/R50** and **N/R40** groups.

(d) $H_0 : \mu_3 = \mu_5$ vs $H_A : \mu_3 \neq \mu_5$. Are the population mean lifetimes the same for the N/R50 and N/R50 lopro groups?

```
## Individual 95% confidence interval:
ind_mu3_mu5 <- fit.contrast(lm_longevity, "Diet", c(0,0,1,0,-1,0), conf.int = 0.95)
print("Individual 95% confidence interval:")
```

```
## [1] "Individual 95% confidence interval:"
```

```
ind_mu3_mu5
```

```
##                          Estimate Std. Error  t value   Pr(>|t|)  lower CI
## Diet c=( 0 0 1 0 -1 0 ) 2.611469    1.19355 2.187984 0.02934503 0.2638701
##                          upper CI
## Diet c=( 0 0 1 0 -1 0 ) 4.959068
## attr(,"class")
## [1] "fit_contrast"
```

```
## Bonferroni 95% familywise interval:
print("Bonferroni 95% familywise interval:")
```

```
## [1] "Bonferroni 95% familywise interval:"
```

```
fit.contrast(lm_longevity, "Diet", c(0,0,1,0,-1,0), conf.int = 1-0.05/5)
```

```
##                          Estimate Std. Error  t value   Pr(>|t|)   lower CI
## Diet c=( 0 0 1 0 -1 0 ) 2.611469    1.19355 2.187984 0.02934503 -0.4801107
##                          upper CI
## Diet c=( 0 0 1 0 -1 0 ) 5.703048
## attr(,"class")
## [1] "fit_contrast"
```

```
## Scheffe 95% familywise interval:
print("Scheffe 95% familywise interval:")
```

```
## [1] "Scheffe 95% familywise interval:"
```

```
c(ind_mu3_mu5[1,1]-scheffe_mult*ind_mu3_mu5[1,2], ind_mu3_mu5[1,1]+scheffe_mult*ind_mu3_mu5[1,2])
```

```
## [1] -1.383186  6.606123
```

Looking at just the individual 95% confidence interval, it appear that the mean lifetimes for the **N/R50** and **N/R50** lopro groups are different. Once we adjust for multiple comparisons, however, the population mean lifetimes appear to be the same for the **N/R50** and **N/R50** lopro groups.

(e) $H_0 : \mu_2 = \mu_1$ **vs** $H_A : \mu_2 \neq \mu_1$**. Are the population mean lifetimes the same for the N/N85 and NP groups?**

```
ind_mu2_mu1 <- fit.contrast(lm_longevity, "Diet", c(-1,1,0,0,0,0), conf.int = 0.95)
print("Individual 95% confidence interval:")
```

```
## [1] "Individual 95% confidence interval:"
```

```
ind_mu2_mu1
```

```
##                          Estimate Std. Error  t value     Pr(>|t|) lower CI
## Diet c=( -1 1 0 0 0 0 ) 5.289187   1.301006 4.065458 5.949477e-05 2.730232
##                          upper CI
## Diet c=( -1 1 0 0 0 0 ) 7.848142
## attr(,"class")
## [1] "fit_contrast"
```

```
## Bonferroni 95% familywise interval:
print("Bonferroni 95% familywise interval:")
```

```
## [1] "Bonferroni 95% familywise interval:"
```

```
fit.contrast(lm_longevity, "Diet", c(-1,1,0,0,0,0), conf.int = 1-0.05/5)
```

```
##                          Estimate Std. Error  t value     Pr(>|t|) lower CI
## Diet c=( -1 1 0 0 0 0 ) 5.289187   1.301006 4.065458 5.949477e-05  1.91927
##                          upper CI
## Diet c=( -1 1 0 0 0 0 ) 8.659104
## attr(,"class")
## [1] "fit_contrast"
```

```
## Scheffe 95% familywise interval:
print("Scheffe 95% familywise interval:")
```

```
## [1] "Scheffe 95% familywise interval:"
```

```
c(ind_mu2_mu1[1,1]-scheffe_mult*ind_mu2_mu1[1,2], ind_mu2_mu1[1,1]+scheffe_mult*ind_mu2_mu1[1,2])
```

```
## [1] 0.9348907 9.6434838
```

No, there is evidence that the mean lifetimes are not the same for the **N/N85** and **NP** groups.

What is the interpretation of the individual 95% confidence interval obtained in part (a) from the fit.contrast function?

We are 95% confident that the difference in population mean lifetimes for the **N/N85** and **N/R50** groups is between -11.942 and -7.270. For 95% of samples, an interval calculated in this way will contain the difference in population mean lifetimes for the **N/N85** and **N/R50** groups.

What is the interpretation of the Bonferroni intervals?

We are 95% confident that the difference in population mean lifetimes for the **N/N85** and **N/R50** groups is between -12.68 and -6.53, for the **N/R50** and **R/R50** groups is between -3.68 and 2.50, for the **N/R50** and **N/R40** groups is between -5.85 and 0.21, for the **N/R50** and **N/R50** lopro groups is between -0.48 and 5.70, and for the **N/N85** and **NP** groups is between 1.92 and 8.66. For 95% of samples, all 5 of the intervals constructed in this way will simultaneously contain the difference in means they are estimating.