STAT 222 Spring 2022 HW8

Matthew Zhao

```
warpbreaks = read.table("http://www.stat.uchicago.edu/~yibi/s222/warpbreaks.txt", h=T)
warpbreaks$wool = as.factor(warpbreaks$wool)
warpbreaks$tension = factor(warpbreaks$tension, labels=c("L","M","H"))
lm1 = lm(breaks ~ wool*tension, data=warpbreaks)
```

Q1 — 5 points

We can use Tukey's HSD to control the FWER, where the HSD is given as $\frac{q_{g,dfE,\alpha}}{\sqrt{2}} \times \sqrt{\text{MSE}(\frac{1}{r} + \frac{1}{r})}$

```
anova(lm1)
## Analysis of Variance Table
##
## Response: breaks
##              Df Sum Sq Mean Sq F value   Pr(>F)
## wool          1    451   450.7   3.765 0.058213 .
## tension       2   2034  1017.1   8.498 0.000693 ***
## wool:tension  2   1003   501.4   4.189 0.021044 *
## Residuals    48   5745   119.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Calculating HSD:

```
mse<-119.7
qval<-qtukey(1-0.02,6,48)/sqrt(2)
qval * sqrt(mse*(2/9))
## [1] 17.137
```

```
sort(mean(breaks~wool+tension,data=warpbreaks))
##     2.H     1.M     1.H     2.L     2.M     1.L
## 18.7778 24.0000 24.5556 28.2222 28.7778 44.5556
```

```
    2.H      1.M       1.H       2.L       2.M       1.L
 18.7778   24.0000   24.5556   28.2222   28.7778   44.5556
 -------------------------------------------------
                                ----------------------------
```

Q2 — 5 points

We can again use Tukey's HSD

```
anova(lm1)
## Analysis of Variance Table
##
## Response: breaks
##               Df Sum Sq Mean Sq F value   Pr(>F)
## wool           1    451   450.7   3.765 0.058213 .
## tension        2   2034  1017.1   8.498 0.000693 ***
## wool:tension   2   1003   501.4   4.189 0.021044 *
## Residuals     48   5745   119.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Calculating HSD:

```
mse<-119.7
qval<-qtukey(1-0.02,3,48)/sqrt(2)
qval * sqrt(mse*(2/(9*2)))
## [1] 10.192
```

```
sort(mean(breaks~tension,data=warpbreaks))
##       H       M       L
## 21.6667 26.3889 36.3889
```

```
     H          M         L
 21.6667    26.3889   36.3889
-------------------
          -------------------
```

Q3 — 5 points

We use Bonferroni.

```
anova(lm1)
## Analysis of Variance Table
##
## Response: breaks
##               Df Sum Sq Mean Sq F value   Pr(>F)
## wool           1    451   450.7   3.765 0.058213 .
## tension        2   2034  1017.1   8.498 0.000693 ***
## wool:tension   2   1003   501.4   4.189 0.021044 *
## Residuals     48   5745   119.7
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
lm1emmean <- emmeans(lm1,~tension:wool)
summary(contrast(lm1emmean, method=list(Ls = c(1,0,0,-1,0,0),Ms = c(0,1,0,0,-1,0),Hs = c(0,0
infer=c(T,F), level=0.98)
##  contrast estimate   SE df lower.CL upper.CL
##  Ls          16.33 5.16 48     1.71    30.96
##  Ms          -4.78 5.16 48   -19.40     9.85
##  Hs           5.78 5.16 48    -8.85    20.40
##
## Confidence level used: 0.98
## Conf-level adjustment: bonferroni method for 3 estimates
```

$\mu_{1L} - \mu_{2L}$: (1.71,30.96)

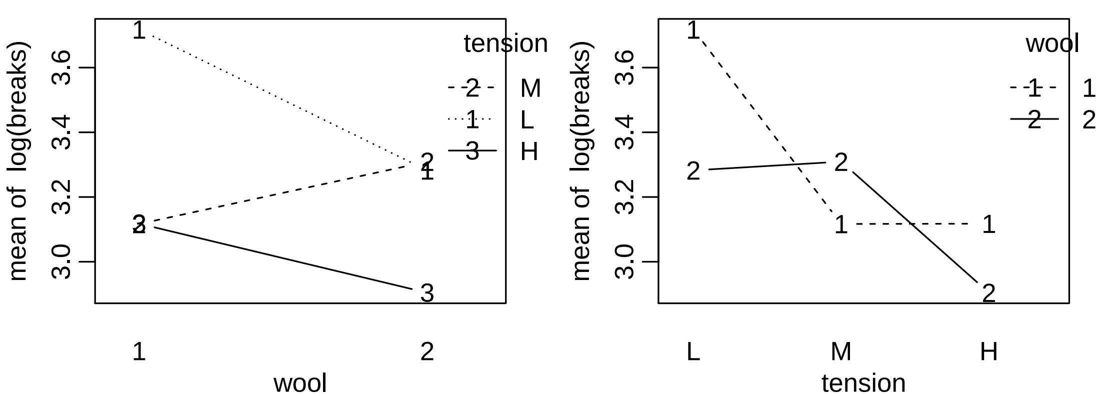
$\mu_{1M} - \mu_{2M}$: (-19.40,9.85)

$\mu_{1H} - \mu_{2H}$: (-8.85,20.40)

We can conclude that there could potentially be an interaction since the CIs for the effect of changing wool type change with tension level. This means that there could be an interaction. However, since the CIs all overlap, with the first and third CIs having significant overlap, we cannot be sure that there is an interaction.

Q4 — 6 points

```
par(mai=c(.6,.6,.1,.3),mgp=c(2,.6,0))
with(warpbreaks, interaction.plot(wool, tension, log(breaks), type="b"))
with(warpbreaks, interaction.plot(tension, wool, log(breaks), type="b"))
```



i) C_1: $\sum_i c_{ij}$

j=L: $(1)\mu_{1L} + (-1)\mu_{2L} = 1 + (-1) = 0$

j=M: $(-1)\mu_{1M} + (1)\mu_{2M} = (-1) + 1 = 0$

$\sum_j c_{ij}$

i=1: $(1)\mu_{1L} + (-1)\mu_{1M} = 1 + (-1) = 0$

i=2: $(-1)\mu_{2L} + (1)\mu_{2M} = (-1) + 1 = 0$

C_2: $\sum_i c_{ij}$

j=H: $(1)\mu_{1H} + (-1)\mu_{2H} = 1 + (-1) = 0$

j=M: $(-1)\mu_{1M} + (1)\mu_{2M} = (-1) + 1 = 0$

$\sum_j c_{ij}$

i=1: $(1)\mu_{1H} + (-1)\mu_{1M} = 1 + (-1) = 0$

i=2: $(-1)\mu_{2H} + (1)\mu_{2M} = (-1) + 1 = 0$

C_3: $\sum_i c_{ij}$

j=L: $(0.5)\mu_{1L} + (-0.5)\mu_{2L} = 0.5 + (-0.5) = 0$

j=M: $(-1)\mu_{1M} + (1)\mu_{2M} = (-1) + 1 = 0$

j=H: $(-0.5)\mu_{1H} + (0.5)\mu_{2H} = (-0.5) + 0.5 = 0$

$\sum_j c_{ij}$

i=1: $(0.5)\mu_{1L} + (0.5)\mu_{1H}(-1)\mu_{1M} = 0.5 + 0.5 + 1 = 0$

i=2: $(-0.5)\mu_{2L} + (-0.5)\mu_{2H} + (1)\mu_{2M} = (-0.5) + (-0.5) + 1 = 0$

ii) We use Scheffe since we are performing contrasts after observing the data/based off of prior analysis. We compare $\frac{\hat{C}}{\text{SE}(\hat{C})}$ with $\sqrt{df_{\text{AB}} F_{df_{AB}, dfE, \alpha}}$, where $\hat{C} = \sum_{ij} c_{ij} \bar{y}_{ij}$.

```
anova(lm1)
## Analysis of Variance Table
##
## Response: breaks
##              Df Sum Sq Mean Sq F value   Pr(>F)
## wool          1    451   450.7   3.765 0.058213 .
## tension       2   2034  1017.1   8.498 0.000693 ***
## wool:tension  2   1003   501.4   4.189 0.021044 *
## Residuals    48   5745   119.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
mean(breaks~wool+tension,data=warpbreaks)
##     1.L     2.L     1.M     2.M     1.H     2.H
## 44.5556 28.2222 24.0000 28.7778 24.5556 18.7778
```

$\hat{C}_1 = 44.5556 - 28.2222 - (24.0000 - 28.7778) = 21.1112$

$\hat{C}_2 = 24.5556 - 18.7778 - (24.0000 - 28.7778) = 10.5556$

$\hat{C}_3 = \frac{\hat{C}_1 + \hat{C}_2}{2} = 15.8334$

$\text{SE}(\hat{C}) = \sqrt{\text{MSE} \times \sum_{ij} \frac{c_{ij}^2}{r}}$

```r
mse <- 119.7
se_c12 <- sqrt(mse*(4*(1/9)))
se_c3 <- sqrt(mse*(((4*((0.5^2)/9))) + (2*(1/9))))
sprintf("C_1 t-stat: %f",21.1112/se_c12)
## [1] "C_1 t-stat: 2.894390"
sprintf("C_2 t-stat: %f",10.5556/se_c12)
## [1] "C_2 t-stat: 1.447195"
sprintf("C_3 t-stat: %f",15.8334/se_c3)
## [1] "C_3 t-stat: 2.506616"

fcrit <- qf(0.04,(2-1)*(3-1),48,lower.tail = F)
scheffe <- sqrt((2-1)*(3-1)*fcrit)
sprintf("Scheffe Critical Value: %f",scheffe)
## [1] "Scheffe Critical Value: 2.624773"
```

At FWER = 0.04, only the first contrast is significant. We can conclude that the change in effect of wool type on breaks is nonzero from low to medium. However, we cannot conclude that there is a significant interaction effect since only 1 of 3 contrasts testing for interaction were significant.

Q5 — 1 point

The upper bound of the FWER for the entire analysis is no more than the sum of the FWER for each family so it is $0.02 + 0.02 + 0.02 + 0.04 = 0.1$.

Q6 — 3 points

Bonferroni: $\hat{C} \pm t_{dfE,\alpha/2/m} * \text{SE}(\hat{C})$

```r
anova(lm1)
## Analysis of Variance Table
##
## Response: breaks
##              Df Sum Sq Mean Sq F value   Pr(>F)
## wool          1    451   450.7   3.765 0.058213 .
## tension       2   2034  1017.1   8.498 0.000693 ***
## wool:tension  2   1003   501.4   4.189 0.021044 *
## Residuals    48   5745   119.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
lm1emmean <- emmeans(lm1,~tension:wool)
summary(contrast(lm1emmean, method=list(C_1 = c(1,-1,0,-1,1,0), C_2 = c(0,-1,1,0,1,-1),C_3 =
infer=c(F,T), level=0.96)
##  contrast estimate   SE df t.ratio p.value
##  C_1          21.1 7.29 48   2.895  0.0171
##  C_2          10.6 7.29 48   1.447  0.4630
##  C_3          15.8 6.32 48   2.507  0.0469
```

```
##
## P value adjustment: bonferroni method for 3 tests
```

We can conclude at FWER=0.04 level that only the first contrast is significant. This means that