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 again use Tukey.

```
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="pairwise", adjust="tukey"),
infer=c(T,F), level=0.98)
##  contrast  estimate   SE df lower.CL upper.CL
##  L 1 - M 1   20.556 5.16 48    3.419     37.7
##  L 1 - H 1   20.000 5.16 48    2.864     37.1
##  L 1 - L 2   16.333 5.16 48   -0.803     33.5
##  L 1 - M 2   15.778 5.16 48   -1.358     32.9
##  L 1 - H 2   25.778 5.16 48    8.642     42.9
##  M 1 - H 1   -0.556 5.16 48  -17.692     16.6
##  M 1 - L 2   -4.222 5.16 48  -21.358     12.9
##  M 1 - M 2   -4.778 5.16 48  -21.914     12.4
##  M 1 - H 2    5.222 5.16 48  -11.914     22.4
##  H 1 - L 2   -3.667 5.16 48  -20.803     13.5
##  H 1 - M 2   -4.222 5.16 48  -21.358     12.9
##  H 1 - H 2    5.778 5.16 48  -11.358     22.9
##  L 2 - M 2   -0.556 5.16 48  -17.692     16.6
##  L 2 - H 2    9.444 5.16 48   -7.692     26.6
##  M 2 - H 2   10.000 5.16 48   -7.136     27.1
##
## Confidence level used: 0.98
## Conf-level adjustment: tukey method for comparing a family of 6 estimates
```

$\mu_{1L} - \mu_{2L}$: (-0.803,33.5)

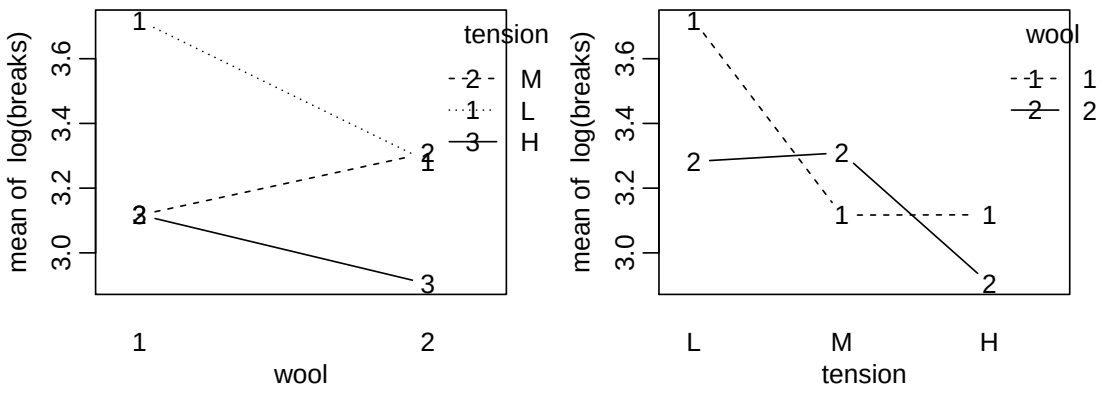
$\mu_{1M} - \mu_{2M}$: (-21.914,12.4)

$\mu_{1H} - \mu_{2H}$: (-11.358,22.9)

We can conclude that the

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)

ii)

Q5 — 1 point

Q6 — 3 points

Bonferroni.