Stat 581, Problem Set #2 Solutions

(1.) $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ (cotton blend has no effect on strength)

$H_A: \mu_i \neq \mu_j$ for some pair (i,j) (cotton blend does have an effect on strength)

(2.) If there are differences in the factor level means, further investigation is required to determine where the differences occur.

(3.) see Box Plot (4.) see output for summary statistics

(5.) $Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, n \end{cases}$, $\sum_{i=1}^{a} \tau_i = 0$

estimates: $\hat{\mu} = \bar{y}_{..}$, $\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$

$\hat{\mu} = 15.04$, $\hat{\tau}_1 = -5.24$, $\hat{\tau}_2 = 0.36$, $\hat{\tau}_3 = 2.56$, $\hat{\tau}_4 = 6.56$, $\hat{\tau}_5 = -4.24$

(6.) (a) $F_0 = 14.76$, p-value = .000

The experiment finds that cotton blend has an effect on strength.

(b) $t_0^{(4,3)} = 2.228$, $p = .04$ The experiment finds that 30% blend is stronger than 25% blend

(c) $t_0^{(3,2)} = 1.225$, $p = .23$ The experiment finds that 25% blend has the same strength as 20% blend

(d) $\underline{15, 35}$ (A) $\underline{20, 25}$ (B) $\underline{30}$ (C)

The experiment finds that 30% blend is the strongest fiber, followed by 20%, 25% (no difference found) and 15%, 35% (no difference found).

7.

(1) The scientific context of the effect being tested.

(2) The complexity of the effect being tested.

(3) The size of effect.

(4) The quality of the experimental design.

```
> library("readxl")

> library("multcomp")
>
> setwd("C:/Users/aneath/iCloudDrive/Lexar/stat581 fall2021")
> hw2.data = read_excel("handout2data.xlsx")
> str(hw2.data)
Classes ‘tbl_df’, ‘tbl’ and 'data.frame':     25 obs. of  11 variables:
 $ strength: num  7 7 15 11 9 12 17 12 18 18 ...
 $ percent : num  15 15 15 15 15 20 20 20 20 20 ...
 $ 20g     : num  24 28 37 30 NA NA NA NA NA NA ...
 $ 30g     : num  37 44 31 35 NA NA NA NA NA NA ...
 $ 40g     : num  42 47 52 38 NA NA NA NA NA NA ...
 $ life    : num  17.6 18.9 16.3 17.4 20.1 21.6 16.9 15.3 18.6 17.1 ...
 $ fluid   : num  1 1 1 1 1 1 2 2 2 2 ...
 $ rate    : num  575 542 530 539 570 565 593 590 579 610 ...
 $ rf power: num  160 160 160 160 160 180 180 180 180 180 ...
 $ brand   : chr  "acme" "acme" "acme" "acme" ...
 $ wear    : num  2.1 2.4 2.5 2.3 2.2 2 1.9 2.1 2.2 2.4 ...
>
> percent = as.factor(na.omit(hw2.data$percent))
> strength = na.omit(hw2.data$strength)
>
> boxplot(strength~percent)
>
> means = by(strength,percent,mean)
> variances = by(strength,percent,var)
> sample.size = by(strength,percent,length)
>
> cbind(means,variances,sample.size)
   means variances sample.size
15   9.8      11.2           5
20  15.4       9.8           5
25  17.6       4.3           5
30  21.6       6.8           5
35  10.8       8.2           5
>
> contrasts(percent) = contr.sum
> hw2.model = aov(strength~percent)
>
> dummy.coef(hw2.model)
Full coefficients are

(Intercept):    15.04
percent:           15    20    25    30    35
                -5.24  0.36  2.56  6.56 -4.24
> summary(hw2.model)
            Df Sum Sq Mean Sq F value   Pr(>F)
percent      4  475.8  118.94   14.76 9.13e-06 ***
Residuals   20  161.2    8.06
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>



> hw2.lsd = glht(hw2.model, linfct = mcp( percent = "Tukey"))
> summary(hw2.lsd,test=univariate())
```

```
        Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts


Fit: aov(formula = strength ~ percent)

Linear Hypotheses:
             Estimate Std. Error t value Pr(>|t|)
20 - 15 == 0    5.600      1.796   3.119 0.005409 **
25 - 15 == 0    7.800      1.796   4.344 0.000315 ***
30 - 15 == 0   11.800      1.796   6.572 2.11e-06 ***
35 - 15 == 0    1.000      1.796   0.557 0.583753
25 - 20 == 0    2.200      1.796   1.225 0.234715
30 - 20 == 0    6.200      1.796   3.453 0.002514 **
35 - 20 == 0   -4.600      1.796  -2.562 0.018595 *
30 - 25 == 0    4.000      1.796   2.228 0.037541 *
35 - 25 == 0   -6.800      1.796  -3.787 0.001157 **
35 - 30 == 0  -10.800      1.796  -6.015 7.01e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Univariate p values reported)

>
> cld(summary(hw2.lsd,test=univariate()))
 15  20  25  30  35
"a" "b" "b" "c" "a"
>
> pairwise.t.test(strength, percent, p.adjust.method = "none")

        Pairwise comparisons using t tests with pooled SD

data:  strength and percent

   15      20      25      30
20 0.00541 -       -       -
25 0.00031 0.23471 -       -
30 2.1e-06 0.00251 0.03754 -
35 0.58375 0.01859 0.00116 7.0e-06

P value adjustment method: none
```