Stat 581, Problem Set #4 Solutions

input factor = fluid type (1,2,3,4), response = insulating life

1. (a)

(i) $F_0 = 3.047$, $p = .0525$

(ii) The p-value quantifies a measure of evidence beyond a determination of statistical significance.

(b) $\Gamma_1 = \mu_1 - \mu_2$, $\Gamma_2 = \mu_3 - \mu_4$, $\Gamma_3 = (\mu_1 + \mu_2) - (\mu_3 + \mu_4)$

(i) $SS_{c_1} = 1.47$, $SS_{c_2} = 13.65$, $SS_{c_3} = 15.04$

For a set of $(a-1)$ orthogonal contrasts, $SS_{tr} = SS_{c_1} + \ldots + SS_{c_{a-1}}$. That is, treatment sum of squares can be decomposed into specific effects.

(ii) $F_{0_1} = 0.445$, $F_{0_2} = 4.138$, $F_{0_3} = 4.559$

($P_{0_1} = .5121$) ($P_{0_2} = .0554$) ($P_{0_3} = .0453$)

(iii) The experiment finds some evidence of a within manufacturer B effect and a between manufacturers A and B effect.

within B effect : fluid type 3 leads to greater lifetimes than fluid type 4

between A,B effect : manufacturer B leads to greater lifetimes than manuf. A.

(2.) input factor = cotton content (15%, 20, 25, 30, 35)
response = fiber strength

(a) see R output for pairwise comparison p-values

Fisher grouping		Tukey grouping	
30%	C	30%	D
25%	B	25%	C D
20%	B	20%	BC
35%	A	35%	AB
15%	A	15%	A

(b) (i) Fisher LSD controls the probability of a type I error for each pairwise comparison.

Tukey HSD controls the overall probability of a type I error across all pairwise comparisons.

(ii) $M_{HSD} = \frac{q_{.05,a,N-a}}{\sqrt{2}} \sqrt{\frac{2 \cdot MSE}{n}} = 5.37$

$\alpha_{ij} = P_0\left(|t_{N-a}| > \frac{q_{.05}}{\sqrt{2}}\right) = .007$ (comparison-wise error probability)

(iii) $M_{LSD} = t_{.025,N-a} \sqrt{\frac{2 \cdot MSE}{n}} = 3.75$

$\alpha = P_0\left(Q_{a,N-a} > \sqrt{2}\, t_{.025}\right) = .264$ (experiment-wise error probability)

(c) When multiple decisions are made in the presence of uncertainty, a measure of belief / evidence is necessary to avoid contradiction.

```
> library("readxl")
>
> setwd("C:/Users/aneath/iCloudDrive/Lexar/stat581 fall2021")
>
> hw4.data = read_excel("handout2data.xlsx")
> str(hw4.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 ...
>
> fluid = as.factor(na.omit(hw4.data$fluid))
> life = na.omit(hw4.data$life)
>
> aov.mod = aov(life~fluid)
> summary(aov.mod)
            Df Sum Sq Mean Sq F value Pr(>F)
fluid        3  30.17   10.05   3.047 0.0525 .
Residuals   20  65.99    3.30
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
> contrasts(fluid) = cbind( c(1,-1,0,0),
+                           c(0,0,1,-1),
+                           c(1,1,-1,-1))
>
> fluid
 [1] 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 3 3 3 4 4 4 4 4 4
attr(,"contrasts")
  [,1] [,2] [,3]
1    1    0    1
2   -1    0    1
3    0    1   -1
4    0   -1   -1
Levels: 1 2 3 4
>
> contr.mod = aov(life~fluid)
> summary(contr.mod,split = list(fluid=list("A"=1,"B"=2,"A-B"=3)))
             Df Sum Sq Mean Sq F value Pr(>F)
fluid         3  30.16   10.05   3.047 0.0525 .
  fluid: A    1   1.47    1.47   0.445 0.5121
  fluid: B    1  13.65   13.65   4.138 0.0554 .
  fluid: A-B  1  15.04   15.04   4.559 0.0453 *
Residuals    20  65.99    3.30
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

```
> library("multcomp")
```

```
> contr = glht(aov.mod, linfct = mcp( fluid = rbind( c(1,-1,0,0),
+                                                    c(0,0,1,-1),
+                                                    c(1,1,-1,-1)) ))
>
> summary(contr,test = adjusted("none"))

         Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts


Fit: aov(formula = life ~ fluid)

Linear Hypotheses:
       Estimate Std. Error t value Pr(>|t|)
1 == 0    0.700      1.049   0.667   0.5121
2 == 0    2.133      1.049   2.034   0.0554 .
3 == 0   -3.167      1.483  -2.135   0.0453 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Adjusted p values reported -- none method)

>
>
> percent = as.factor(na.omit(hw4.data$percent))
> strength = na.omit(hw4.data$strength)
>
> aov2.mod = aov(strength~percent)
> summary(aov2.mod)
            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
>
> comparisons.mod = glht(aov2.mod, linfct = mcp( percent = "Tukey"))
>
> summary(comparisons.mod,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)
```

```
> summary(comparisons.mod)

         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  0.03851 *
25 - 15 == 0    7.800      1.796   4.344  0.00261 **
30 - 15 == 0   11.800      1.796   6.572  < 0.001 ***
35 - 15 == 0    1.000      1.796   0.557  0.97977
25 - 20 == 0    2.200      1.796   1.225  0.73727
30 - 20 == 0    6.200      1.796   3.453  0.01885 *
35 - 20 == 0   -4.600      1.796  -2.562  0.11631
30 - 25 == 0    4.000      1.796   2.228  0.21016
35 - 25 == 0   -6.800      1.796  -3.787  0.00901 **
35 - 30 == 0  -10.800      1.796  -6.015  < 0.001 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
(Adjusted p values reported -- single-step method)

>
> cld(summary(comparisons.mod,test=univariate()))
 15  20  25  30  35
"a" "b" "b" "c" "a"
> cld(summary(comparisons.mod))
  15   20   25   30   35
 "a" "bc" "cd"  "d" "ab"
>
>
> a = 5
> n = 5
> df = a*(n-1)
>
> mse = 8.06
>
> m_lsd = qt(.025,df,lower.tail = FALSE)*sqrt(2*mse/n)
> m_lsd
[1] 3.745452
>
> m_tukey = qtukey(.05,a,df,lower.tail = FALSE)*sqrt(mse/n)
> m_tukey
[1] 5.372958
>
> 2*pt(qtukey(.05,a,df,lower.tail = FALSE)/sqrt(2),df,lower.tail = FALSE)
[1] 0.007198365
>
> ptukey(qt(.025,df,lower.tail=FALSE)*sqrt(2),a,df,lower.tail = FALSE)
[1] 0.2643089
```