HW2

Sam Mottahedi

July 8, 2017

Problem 1

Source	SS	DF	MS	F	p-value
Between treatments	64.42	3	21.4733333	8.98	8.6632306×10^{-4}
Within treatments	40.63	17	2.39		
Total	105.05	20			

Problem 2

```
y <- c(3129, 3000 ,2865 ,2890,3200, 3300 ,2975 ,
       3150,2800 ,2900 ,2985 ,3050,2600 ,2700,
       2600, 2765)
x <- rep(1:4, times=1, each=4)

fit <- aov(y ~ as.factor(x))
summary(fit)
```

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## as.factor(x)  3 489740  163247   12.73 0.000489 ***
## Residuals    12 153908   12826
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Source	SS	DF	MS	F	p-value
Between treatments	489740	3	163247	12.73	0.000489
Within treatments	153908	12	12826		
Total	6.43648×10^5	15			

$p - value < 0.05$ The mean tensile strength for the four different mixing techniques are different.

Problem 3

a)

```
tapply(df$V1, df$V2, mean)
```

```
##    1    2    3
## 29.4 29.6 28.0
```

```r
tapply(df$V1, df$V2, var)
```

```
##    1    2    3
## 13.3 10.8  5.0
```

```r
var(df$V1)
```

```
## [1] 8.857143
```

```r
describeBy(df$V1, df$V2)
```

```
##
##  Descriptive statistics by group
## group: 1
##    vars n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 5 29.4 3.65     28    29.4 2.97  26  35     9 0.51    -1.69 1.63
## ------------------------------------------------------------
## group: 2
##    vars n mean   sd median trimmed  mad min max range  skew kurtosis   se
## X1    1 5 29.6 3.29     29    29.6 2.97  25  34     9 -0.06    -1.58 1.47
## ------------------------------------------------------------
## group: 3
##    vars n mean   sd median trimmed  mad min max range skew kurtosis se
## X1    1 5   28 2.24     28      28 1.48  25  31     6    0    -1.69  1
```

b)

$$se_{blue} = 1.63$$

$$se_{green} = 1.47$$

$$se_{orange} = 1$$

c)

```r
summary( aov(V1 ~ V2, df))
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## V2           2    7.6     3.8   0.392  0.684
## Residuals   12  116.4     9.7
```

$$p - value = 0.684 > 0.05$$

We can't rejecting the null $(H_0 : \mu_1 = \cdots = \mu_4)$. The mean differences are not significant.

d)

```r
c1 <- c(1, 0, -1)
c2 <- c(0, 1, -2)
mat <- cbind(c1, c2)
contrasts(df$V2) <- mat

fit <- aov(V1 ~ V2, df)
summary.aov(fit, split=list(V2=list('orange vs blue'=1, 'orange vs green'=2)))
```

```
##                      Df Sum Sq Mean Sq F value Pr(>F)
## V2                    2    7.6     3.8   0.392  0.684
##   V2: orange vs blue  1    4.9     4.9   0.505  0.491
##   V2: orange vs green 1    2.7     2.7   0.278  0.607
## Residuals            12  116.4     9.7
```

since both p-values for "orange vs. blue" and "orange vs green" are greater than 0.05 we can't reject the null hypothesis $H_0 : \mu_{orange} = \mu_{green}$ and $H_0 : \mu_{orange} = \mu_{blue}$.

Problem 4

a)

Source	SS	DF	MS	F	p-value
Drug	56.10	3	18.7	7.2247263	0.0012804
Error	62.12	24	2.5883333		
Total	6.43648×10^5	27			

$7.2247263 \mid 0.0012804 < 0.05$

rejecting the null hypothesis $H_0 : \mu_1 = \cdots = \mu_4$, there is significant difference between bio-activity levels of the four drugs.

b)

Six.

c)

```
qt(1-0.05/2, 28-4)
```

```
## [1] 2.063899
```

d)

Tukey:
```
qtukey(0.95, 4, 24)/sqrt(2)
```

```
## [1] 2.758609
```

benferroni:
```
qt(1-0.95/(2*6), 24)
```

```
## [1] 1.456063
```

scheffe:
```
S2 = (4-1) * qf(0.05, 4-1, 28-4)
sqrt(S2)
```

```
## [1] 0.5893068
```

e)

```
u1 = 66.1; u2 = 65.75; u3= 62.63; u4 = 63.85

l1 = u1 - u2; l2= u1 - u3
l3 = u1 - u4; l4 = u2- u3
l5= u2-u4; l6=u4-u3

mse = 2.5883333


s2 = mse*(1/7 + 1/7)
```

tukey: 2.758609

```
for (l in c(l1, l2, l3, l4, l5, l6)){

    print(2 * l / sqrt(s2))

}
```

```
## [1] 0.8139954
## [1] 8.070183
## [1] 5.232828
## [1] 7.256188
## [1] 4.418832
## [1] 2.837355
```

significant pairs:

all except $\mu_1 - \mu_2$

Scheffe: 0.5893068

```
for (l in c(l1, l2, l3, l4, l5, l6)){

    print(l^2 / (3 * s2))

}
```

```
## [1] 0.05521571
## [1] 5.427321
## [1] 2.281874
## [1] 4.387688
## [1] 1.627173
## [1] 0.6708822
```

all except $\mu_1 - \mu_2$

Bonferroni: 1.456063

```
for (l in c(l1, l2, l3, l4, l5, l6)){

    print(l / sqrt(s2))

}
```

```
## [1] 0.4069977
## [1] 4.035092
## [1] 2.616414
## [1] 3.628094
## [1] 2.209416
## [1] 1.418678
```

all except $\mu_1 - \mu_2, \mu_4 - \mu_3$

f)

The benferroni is more conservative compared to other two. Scheffe is usually used when the sample size are different. since the sample sizes are equal the tukey method is more appropriate for this situations.