STAT 461: Homework 5

Name: Kyle Salitrik | ***ID#:*** *997543474* | ***PSU ID:*** *kps168*

October 1, 2018

Problem 1

a)

$$
\begin{aligned}
&Y_{it} = \mu + \tau_i + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma^2) \\
&\quad i = A, B, C \\
&\quad t = 1, \ldots, r_i; \quad r_A = r_B = r_C = 2
\end{aligned}
$$

b)

$$
\begin{aligned}
&\overline{Y}_{B\cdot} = \frac{5-1}{2} = 4 \\
&\overline{Y}_{B\cdot} \sim N\left(4, \frac{\sigma^2}{2}\right)
\end{aligned}
$$

c)

$$
\begin{aligned}
\overline{Y}_{A\cdot} &= \frac{-14-4}{2} = -9 \\
\overline{Y}_{B\cdot} &= \frac{5-1}{2} = 2 \\
\overline{Y}_{C\cdot} &= \frac{-2+6}{2} = 2 \\
\overline{Y}_{\cdot\cdot} &= \frac{-14-4+5-1-2+6}{6} = \frac{-5}{3} \\
\text{SSE} &= \sum_{i=1}^{v} \sum_{t=1}^{r_i} \left(\overline{Y}_{it} - \overline{Y}_{i\cdot}\right)^2 \\
&= (-14+9)^2 + (-4+9)^2 + (5-2)^2 + (-1-2)^2 + (-2-2)^2 + (6-2)^2 \\
&= 100 \\
\text{SST} &= \sum_{i=1}^{v} r_i \left(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot}\right)^2 \\
&= 2\left(-9 - \frac{-5}{3}\right)^2 + 2\left(2 - \frac{-5}{3}\right)^2 + 2\left(2 - \frac{-5}{3}\right)^2 \\
&= \frac{484}{3} \approx 161.\overline{333} \\
\text{SSTOT} &= \sum_{i=1}^{v} \sum_{t=1}^{r_i} \left(\overline{Y}_{it} - \overline{Y}_{i\cdot}\right)^2 \\
&= \left(-14 - \frac{-5}{3}\right)^2 + \left(-4 - \frac{-5}{3}\right)^2 + \left(5 - \frac{-5}{3}\right)^2 + \left(-1 - \frac{-5}{3}\right)^2 + \left(-2 - \frac{-5}{3}\right)^2 + \left(6 - \frac{-5}{3}\right)^2 \\
&= \frac{784}{3} \approx 261.\overline{333}
\end{aligned}
$$

d)

$$
\widehat{\sigma^2} \approx \frac{\text{SSE}}{n-v} = \frac{756}{6-3} = 252
$$

e)

$$
\widehat{\Delta}_{AC} = \overline{Y}_{A\cdot} - \overline{Y}_{C\cdot} = -9 - 2 = -11
$$

f)

$$\overline{Y}_{A\cdot} - \overline{Y}_{C\cdot} \sim N(\mu + \tau_A, \sigma^2) + N(-\mu - \tau_C, \sigma^2) = N(0, 2\sigma^2)$$

g)

$$\frac{\overline{Y}_{A\cdot} - \overline{Y}_{C\cdot}}{K(\sigma)} \sim N(0, 2\left(\frac{1}{K(\sigma)}\right)^2 \sigma^2)$$
$$K(\sigma) = \sqrt{2}\sigma$$

h)

$$\frac{\text{SSE}}{\sigma^2} \sim \chi^2_{n-v}$$

i)

$$\frac{\left(\overline{Y}_{A\cdot} - \overline{Y}_{C\cdot}\right)^2 / (K(\sigma))^2}{\text{SSE}/\left[(n-v)\sigma^2\right]} \sim F_{1,(n-v)}$$

j)

Under the null hypothesis of $H_0 : \tau_A = \tau_B = \tau_C$, we use the test statistic $T^* = \frac{\text{SST}/(v-1)}{\text{SSE}/(n-v)}$ where $T^* \sim F_{(v-1),(n-v)}$.

In general:

	DF	Sum Sq	Mean Sq	F-Value
Treatment	v-1	SST	$\text{SST}/(v-1)$	$\frac{\text{SST}/(v-1)}{\text{SSE}/(n-v)}$
Error	n-v	SSE	$\text{SSE}/(n-v)$	NA
Total	n-1	SSTOT	NA	NA

For our case:

	DF	Sum Sq	Mean Sq	F-Value
Treatment	2	$\frac{484}{3}$	$\frac{242}{3}$	$\frac{121}{50}$
Error	3	100	$\frac{100}{3}$	
Total	5	$\frac{784}{3}$	NA	NA

k)

Based on the below output from R, we can conclude that there is no significant difference in the response of the three populations. In this case, H_0 should not be rejected.

```
Analysis of Variance Table

Response: delta
          Df Sum Sq Mean Sq F value Pr(>F)
meds       2 161.33  80.667    2.42 0.2367
```

l)

Examining the pairwise comparisons for the variables, we can see that none of the p-values are significant for any of the contrasts. This confirms that there is no significant difference between the treatments of the 3 populations.

$H_0 : \tau_A = \tau_B \quad H_A : \tau_A \neq \tau_B$
$H_0 : \tau_A = \tau_C \quad H_A : \tau_A \neq \tau_C$
$H_0 : \tau_B = \tau_C \quad H_A : \tau_B \neq \tau_C$

$$T^* = \frac{\sqrt{\frac{r_i + r_j}{r_i r_j}}\left(\overline{Y}_i - \overline{Y}_j\right)}{\sqrt{SSE/(n-v)}}$$

$$T^* \sim t_{n-v} = t_3$$

```
 contrast           estimate         SE  df  t.ratio  p.value
A - B          -1.100000e+01  5.773503   3   -1.905   0.2817
A - Control    -1.100000e+01  5.773503   3   -1.905   0.2817
B - Control    -1.776357e-15  5.773503   3    0.000   1.0000
```

PROBLEM 2

$H_0 : \tau_{reg} = \tau_{deo} \quad H_A : \tau_{reg} \neq \tau_{deo}$
$H_0 : \tau_{reg} = \tau_{moi} \quad H_A : \tau_{reg} \neq \tau_{moi}$
$H_0 : \tau_{deo} = \tau_{moi} \quad H_A : \tau_{deo} \neq \tau_{moi}$

$$T^* = \frac{\sqrt{\frac{r_i + r_j}{r_i r_j}}\left(\overline{Y}_i - \overline{Y}_j\right)}{\sqrt{SSE/(n-v)}}$$

$$T^* \sim t_{n-v} = t_9$$

Examining the ANOVA table, one can see that there is a significant difference between at least one pair of treatment. Therefore, we reject H_0 and must examine further to determine which treatments are significant.

Looking at the pairwise comparisons, we can observe that the deodorant soap lost less than the regular soap, and that the moisturizing soap also lost less than the regular soap. However, there was no significant difference between the moisturizing and deodorant soaps.

```
Analysis of Variance Table

Response: loss
          Df  Sum Sq Mean Sq  F value     Pr(>F)
type       2 16.1220  8.0610   104.45 5.914e-07 ***
Residuals  9  0.6946  0.0772
---
Signif. codes:  0 Ś***Š 0.001 Ś**Š 0.01 Ś*Š 0.05 Ś.Š 0.1 Ś Š 1

--------------------------------------------------------------------

 contrast  estimate       SE df t.ratio p.value
deo - moi    0.7075 0.196437  9   3.602  0.0143
deo - reg    2.7350 0.196437  9  13.923  <.0001
moi - reg    2.0275 0.196437  9  10.321  <.0001
```

Problem 3

$H_0 : \tau_0 = \tau_1 \quad H_A : \tau_0 \neq \tau_1$

$H_0 : \tau_0 = \tau_2 \quad H_A : \tau_0 \neq \tau_2$

$H_0 : \tau_0 = \tau_3 \quad H_A : \tau_0 \neq \tau_3$

$H_0 : \tau_1 = \tau_2 \quad H_A : \tau_1 \neq \tau_2$

$H_0 : \tau_1 = \tau_3 \quad H_A : \tau_1 \neq \tau_3$

$H_0 : \tau_2 = \tau_3 \quad H_A : \tau_2 \neq \tau_3$

$$T^* = \frac{\sqrt{\frac{r_i + r_j}{r_i r_j}} \left(\overline{Y}_i - \overline{Y}_j \right)}{\sqrt{SSE / (n - v)}}$$

$$T^* \sim t_{n-v} = t_{28}$$

The p-value obtained from the ANOVA is extremely large, indicating that H_0 should not be rejected. Continuing to look at the pairwise comparisons, the p-values for each contrast are also very large, indicating that there is no difference between the treatments.

```
Analysis of Variance Table

Response: time
          Df   Sum Sq   Mean Sq F value Pr(>F)
presses    3 0.008047 0.0026824  0.2455 0.8638
Residuals 28 0.305953 0.0109269

-----------------------------------------------------------------------

 contrast     estimate         SE df t.ratio p.value
 0 - 1     0.036142857 0.05151381 28   0.702  0.8956
 0 - 2     0.013142857 0.05151381 28   0.255  0.9940
 0 - 3    -0.004857143 0.06120753 28  -0.079  0.9998
 1 - 2    -0.023000000 0.04674802 28  -0.492  0.9602
 1 - 3    -0.041000000 0.05725440 28  -0.716  0.8899
 2 - 3    -0.018000000 0.05725440 28  -0.314  0.9890
```

Code Appendix

```r
####################################################################
#### Setup
####################################################################
## Install and load libraries
# ipak function taken from: https://gist.github.com/stevenworthington/3178163
ipak <- function(pkg) {
  new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])]
  if (length(new.pkg))
    install.packages(new.pkg, dependencies = TRUE)
  sapply(pkg, require, character.only = TRUE)
}
packages <- c("ggplot2", "reshape2", "gridExtra", "TSA", "astsa", "orcutt",
              "nlme", "fGarch", "vars", "lsmeans")
ipak(packages)

####################################################################
#### Problem 1
####################################################################
meds = c(rep('A',2),rep('B',2),rep('Control',2))
delta = c(-14,-4,5,-1,-2,6)
bpData = data.frame(delta, meds)

# Calculate ANOVA table and Pairwise Comparison
bpModel = aov(delta ~ meds, data=bpData)
anova(bpModel)
bpLSM = lsmeans(bpModel, ~ meds)
contrast(bpLSM, method='pairwise')

####################################################################
#### Problem 2
####################################################################
# From HW1
loss = c(-.3,-.1,-.14,.4,2.63,2.61,2.41,3.15,1.86,2.03,2.26,1.82)
type = c(rep('reg',4), rep('deo',4), rep('moi',4))
losses = data.frame(loss, type)

# Calculate ANOVA table and Pairwise Comparison
lossModel = aov(loss~type, data=losses)
anova(lossModel)
lossLSM = lsmeans(lossModel, ~type)
contrast(lossLSM, method='pairwise')


####################################################################
#### Problem 3
####################################################################
# From HW4
time = c(38.14, 38.20, 38.31, 38.14, 38.29, 38.17, 38.20,
         38.28, 38.17, 38.08, 38.25, 38.18, 38.03, 37.95, 38.26, 38.30, 38.21,
         38.17, 38.13, 38.16, 38.30, 38.34, 38.34, 38.17, 38.18, 38.09, 38.06,
         38.14, 38.30, 38.21, 38.04, 38.37)
presses = c(rep('0',7), rep('1',10), rep('2',10), rep('3',5))
lights = data.frame(time, presses)

# Calculate ANOVA table and Pairwise Comparison
lightModel = aov(time~presses, data=lights)
anova(lightModel)
lightLSM = lsmeans(lightModel, ~presses)
contrast(lightLSM, method='pairwise')
```