STAT 6910: HW 6

David Angeles

```
## Warning: package 'emmeans' was built under R version 3.4.4
## NOTE: As of emmeans versions > 1.2.3,
##       The 'cld' function will be deprecated in favor of 'CLD'.
##       You may use 'cld' only if you have package:multcomp attached.
```

Problem 1

Under what circumstances should the two-way main effects model (6.2.3) be used rather than the two-way complete model (6.2.2)? [1 sentence may suffice.] Discuss the interpretation of main effects in each model. [You can focus on comparing the interpretations of, e.g., $\alpha_1 - \alpha_2$ vs. $\alpha_1^\star - \alpha_2^\star$ in these models.]

Answer:

The two-way main effects model (6.2.3) should be used rather than the two-way complete model (6.2.2) when an experimenter has sufficient knowledge about the two treatment factors being studied to state with reasonable certainty that the factors do not interact.

Using the complete two-way model, one cannot compare directly the effects of the different levels of the factors, say level 1 of A, α_1, with level 2 of A, α_2. Instead, one must compare the effects of the different levels of A averaged over the effects of B. So, $\alpha_1 - \alpha_2$ is not estimable, but $\alpha_1^\star - \alpha_2^\star = \bar{\tau}_{1\bullet} - \bar{\tau}_{2\bullet}$ is estimable. On the other hand, using the two-way main effects model one can compare the effect of changing from level 1 to 2 of factor A, $\alpha_1 - \alpha_2$, since factor A doesn't need to be averaged over factor B because this model assumes there is no interaction between factor A and B.

Problem 2

Solution on the very back.

Problem 7 part (a)

The data shown in Table 6.22 are a subset of the data given by Anderson and McLean (1974) and show the strength of a weld in a steel bar. Two factors of interest were gage bar setting (the distance the weld die travels during the automatic weld cycle) and time of welding (total time of the automatic weld cycle). Assume that the levels of both factors were selected to be equally spaced.

```r
weld.data = read.table("~/Desktop/Stats 6910/HW_4_and_5/weld.strength.txt",
                       header = TRUE)
weld.data$trtmt <- as.factor(weld.data$trtmt)
```

(a) Using the cell-means model (6.2.1) for these data, test the hypothesis that there is no difference in the effects of the treatment combinations on weld strength against the alternative hypothesis that at least two treatment combinations have different effects.

```r
cell_model <- aov(strength ~ trtmt , data = weld.data)
```

Let τ_{ij} be the deviation from the overall mean weld strength, μ, where factor A, the gage bar setting, has level i and factor B, the time of welding, has level j. We can test the effects of Factor A and B on welding strength using an α-level of $\alpha = 0.05$, and the following test:

The Null Hypothesis of No Effect on Weld Strength
$H_0 : \tau_{ij} = 0 \;\forall\, i = 1, 2, ..., a$ and $\forall\, j = 1, 2, ..., b$
vs. The Alternative Hypothesis of Some Effect on Weld Strength $H_A : \tau_{ij} \neq \tau_{hk}$ for some $i \neq h$ and $j \neq k$

```r
anova(cell_model)
```

```
## Analysis of Variance Table
##
## Response: strength
##           Df Sum Sq Mean Sq F value    Pr(>F)
## trtmt     14 1261.2  90.086  8.2395 0.0001092 ***
## Residuals 15  164.0  10.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the analysis of variance table we obtain a p-value equal to 0.0001, which is less that the α-level of $\alpha = 0.05$. Therefore, we can reject the null hypothesis of no effect on welding strength in favor of the alternative hypothesis of some effect on welding strength.

Problem 8

For the experiment described in Exercise 7, use the two-way complete model instead of the equivalent cell means model.

(a) Test the hypothesis of no interaction between gage bar setting and time of weld and state your conclusion.

Let τ_{ij} be the deviation from the overall mean weld strength, μ, where factor A, the gage bar setting, has level i and factor B, the time of welding, has level j. We can test the effects of Factor A and B on welding strength using an α-level of $\alpha = 0.05$, and the following test:

The Null Hypothesis of No Interaction Between A and B
$H_0 : (\alpha\beta)_{ij} - (\alpha\beta)_{iq} - (\alpha\beta)_{sj} + (\alpha\beta)_{sq} = 0 \;\forall\, i < s$ and $\forall\, j < j$

vs. The Alternative Hypothesis of Some Effect on Weld Strength
$H_A : (\alpha\beta)_{ij} - (\alpha\beta)_{iq} \neq (\alpha\beta)_{sj} - (\alpha\beta)_{sq}$ for some $i \neq s$ and $j \neq q$

```
weld.data$gage <- as.factor(weld.data$gage)
weld.data$time <- as.factor(weld.data$time)
```

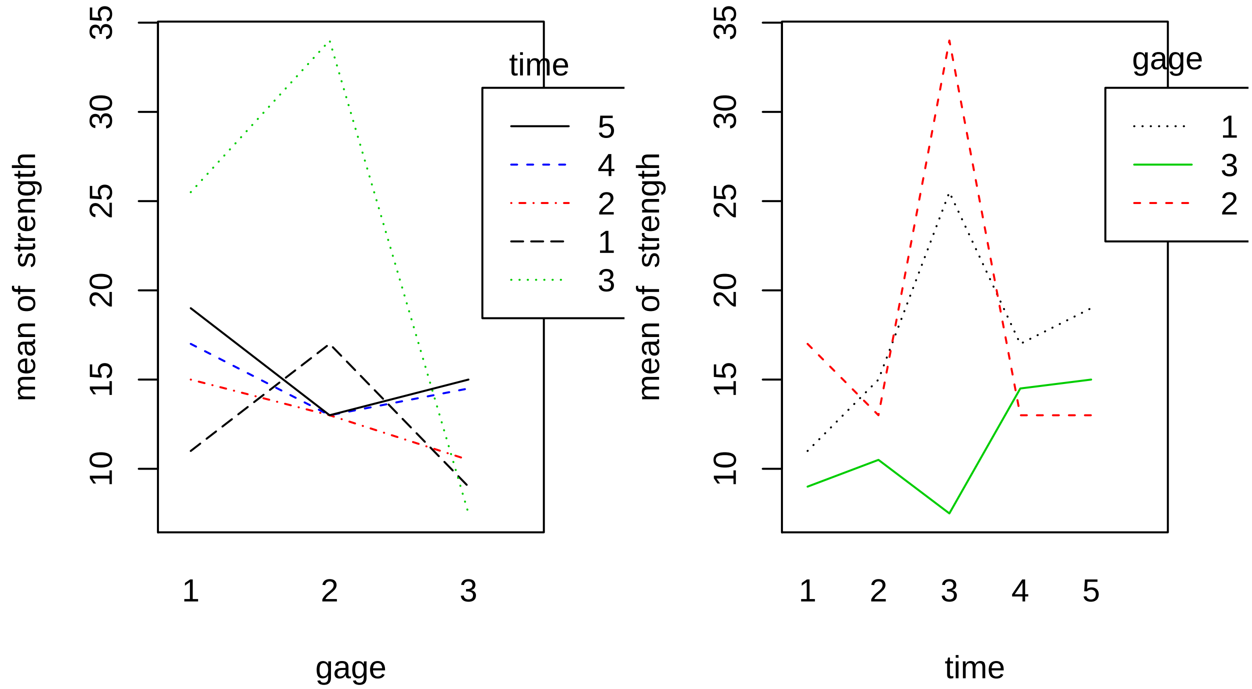
```
model_AB <- aov(strength ~ gage*time, data = weld.data)
anova(model_AB)
```

```
## Analysis of Variance Table
##
## Response: strength
##           Df Sum Sq Mean Sq F value    Pr(>F)
## gage       2 278.60 139.300 12.7409 0.0005838 ***
## time       4 385.53  96.383  8.8155 0.0007212 ***
## gage:time  8 597.07  74.633  6.8262 0.0007535 ***
## Residuals 15 164.00  10.933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the analysis of variance table we obtain a p-value equal to 0.0008, which is less that the α-level of $\alpha = 0.05$. Therefore, we can reject the null hypothesis of no interaction between between A and B in favor of the alternative hypothesis of some interaction between A and B.

(b) Draw an interaction plot for the two factors Gage bar setting and Time of welding. Does your interaction plot support the conclusion of your hypothesis test? Explain.

```
attach(weld.data)
par(mfrow = c(1, 2))
interaction.plot(x.factor = gage, trace.factor = time, response = strength,
col = c(1, 2, 3, 4), leg.bty = "o", leg.bg = "white")
interaction.plot(x.factor = time, trace.factor = gage, response = strength,
col = c(1, 2, 3), leg.bty = "o", leg.bg = "white")
```



From the interaction plots above, Mean of Strength vs. Time and Means of Strength vs. Gage, we can see that the lines aren't parallel and cross multiple times in both plots. This implies that there is some interaction between gage setting and welding time which supports the conclusion of part (a).

(c) In view of your answer to part (b), is it sensible to investigate the differences between the effects of gage bar setting? Why or why not? Indicate on your plot what would be compared.

It is not sensible because if we were to compare the differences between the gage bar settings we would have to compare them averages over welding time due to the fact that there is some interaction. The difference between gage bar setting level 2 and level 1 averaged over welding time is 1.5. However, if we look at the plot, Mean Strength vs. Gage, we see that the welding time at level 3 produces a much higher mean strength when the gage bar setting is at level 2. However, this is not apparent when we investigate the differences between the effects of gage bar averaged over welding time.

```
gage_avg <- tapply(weld.data$strength, weld.data$gage , mean); gage_avg
```

```
##    1    2    3
## 17.5 18.0 11.3
```

Problem 21

The experiment was run in order to examine the amount of time taken to boil a given amount of water on the four different burners of her stove, and with 0, 2, 4, or 6 teaspoons of salt added to the water. Thus the experiment had two treatment factors with four levels each.

The experimenter ran the experiment as a completely randomized design by taking $r = 3$ observations on each of the 16 treatment combinations in a random order. The data are shown in Table 6.26. The experimenter believed that there would be no interaction between the two factors.

```
boiling.data = read.table("~/Desktop/Stats 6910/HW_4_and_5/water.boiling.txt",
                          header = TRUE)
# make salt and burner factors and add column of two digit treatment combination
boiling.data <- within(boiling.data,{
   f_salt = factor(salt); f_burner = factor(burner)
   trtmt= factor(10*burner + salt)})
# Display boiling.data
head(boiling.data, 10)
```

```
##    salt burner time order trtmt f_burner f_salt
## 1     0      1    7     7    10        1      0
## 2     0      1    8    21    10        1      0
## 3     0      1    7    30    10        1      0
## 4     0      2    4     6    20        2      0
## 5     0      2    4    20    20        2      0
## 6     0      2    4    27    20        2      0
## 7     0      3    6     9    30        3      0
## 8     0      3    7    16    30        3      0
## 9     0      3    6    22    30        3      0
## 10    0      4    9    29    40        4      0
```

(a) Check the assumptions on the two-way main-effects model.

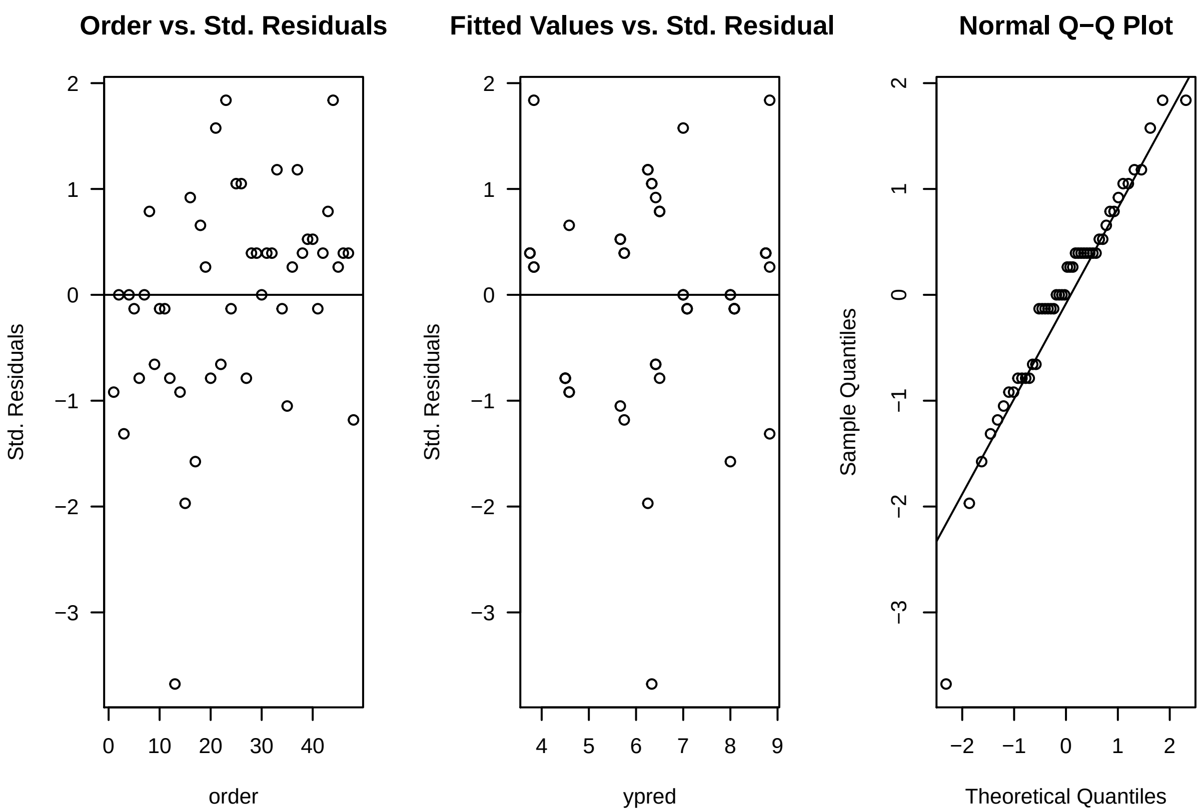
```
two.way_model <- aov(time ~ f_salt+f_burner, data = boiling.data)

boiling.data = within(boiling.data, {
  # Compute predicted, residual, and standardized residual values
  ypred = fitted(two.way_model)
  e = resid(two.way_model)
  z = rstandard(two.way_model)})
# Display first 10 lines of boiling.data, 4 digits per variable
print(head(boiling.data, 5), digits = 4)
```

```
##   salt burner time order trtmt f_burner f_salt          z          e ypred
## 1    0      1    7     7    10        1      0  2.502e-14  1.588e-14   7.0
## 2    0      1    8    21    10        1      0  1.576e+00  1.000e+00   7.0
## 3    0      1    7    30    10        1      0  0.000e+00 -3.965e-15   7.0
## 4    0      2    4     6    20        2      0 -7.878e-01 -5.000e-01   4.5
## 5    0      2    4    20    20        2      0 -7.878e-01 -5.000e-01   4.5
```

```
# Generate residual plots
par(mfrow = c(1,3))
```

```r
plot(z ~ order, data=boiling.data,
     main= "Order vs. Std. Residuals",
     ylab="Std. Residuals", las=1)
abline(h=0)
plot(z ~ ypred, data=boiling.data,
      main= "Fitted Values vs. Std. Residuals",
     ylab="Std. Residuals", las=1)
abline(h=0)
qqnorm(boiling.data$z)
# Line through 1st and 3rd quantile points
qqline(boiling.data$z)
```



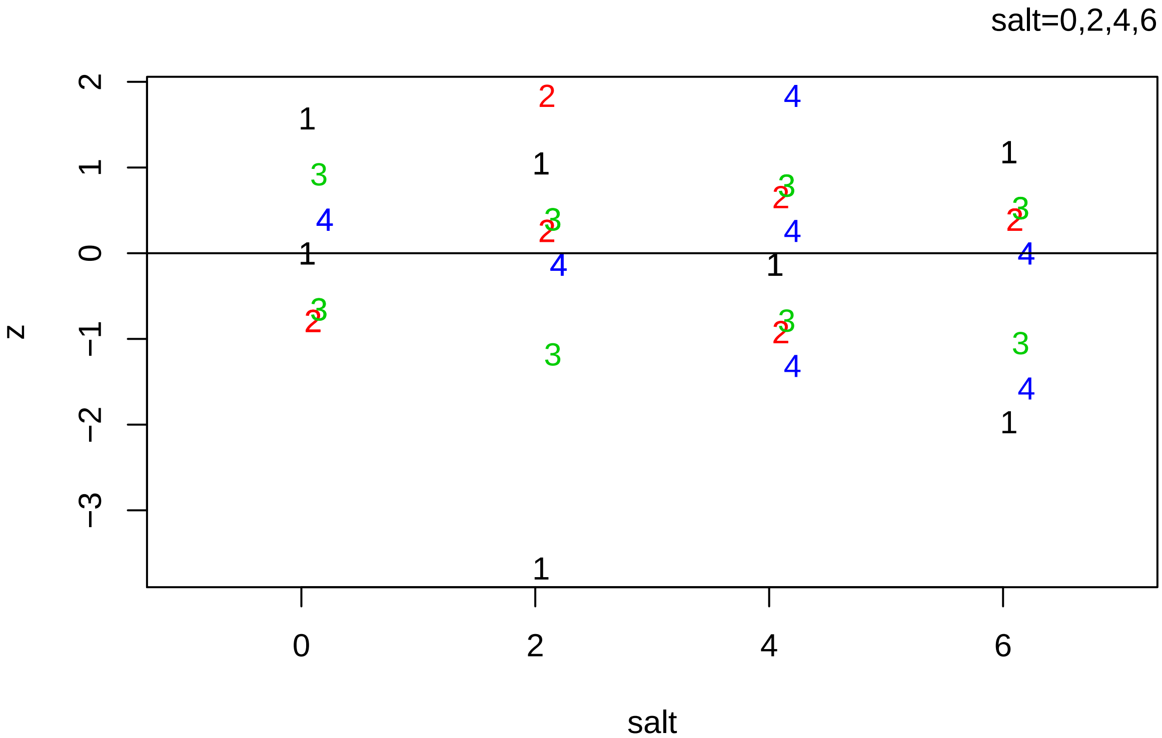
Assumption (a): The error have mean 0:

```r
attach(boiling.data)
```

```
## The following objects are masked from weld.data:
##
##     time, trtmt
```

```r
plot(z ~ salt, xaxt = "n", type = "n", xlim = c(-1, 7)) # Suppress x-axis, pts
axis(1, at = c(0, 2, 4,6))
points(x = as.numeric(salt) + as.numeric(burner) * 0.05, y = z,
pch = as.character(as.numeric(burner)), col = as.numeric(burner))
```

```r
mtext("salt=0,2,4,6", side = 3, adj = 1, line = 1)
abline(h = 0) # Horizontal line at zero
```




```r
# Margin text, top-rt, line 1 abline(h = 0)
```

From the plot above we can see that the burner factor with level 2 is consistently above or below the 0 line, except for when the salt level is 4, which is an indication that the the errors might not have mean zero and therefore the two-way main effects model is probably not the best model to represent the data.

Assumption (b): The error have constant variance:

By plotting the standardized residuals against the fitted values we can see that the spread of the standardized residuals are roughly centered around zero. However there seems to be only one apparent outlier. The outlier may be of interest to inspect further but might not be too concerning. Hence, we can say that the equal variance assumption is roughly satisfied.

Assumption (c): The error are normally distributed:

From the qq-plot above we see that the data is fairly straight. However, there is just one apparent outliers. Therefore, the normality assumption is fairly reasonable.

Assumption (d): The errors are independent:

From the plot of Order vs. Std. Residuals we can see that there isn't an apparent pattern as time increases. Therefore we feel comfortable to say that the assumption that errors are independent is approximately satisfies.

(b) Calculate a 99% set of Tukey confidence intervals for pairwise differences between the levels of salt, and calculate separately a 99% set of intervals for pairwise differences between the levels of burner.

```
em.salt <- emmeans(two.way_model, specs = ~f_salt)
em.burner <- emmeans(two.way_model, specs = ~f_burner)

summary(contrast(em.salt, method = "tukey"),infer = c(TRUE,TRUE), level=.99)
```

```
##  contrast    estimate        SE df    lower.CL  upper.CL t.ratio p.value
##  0 - 2     0.66666667 0.2803405 41 -0.26256979 1.5959031   2.378  0.0974
##  0 - 4    -0.08333333 0.2803405 41 -1.01256979 0.8459031  -0.297  0.9907
##  0 - 6     0.75000000 0.2803405 41 -0.17923646 1.6792365   2.675  0.0503
##  2 - 4    -0.75000000 0.2803405 41 -1.67923646 0.1792365  -2.675  0.0503
##  2 - 6     0.08333333 0.2803405 41 -0.84590313 1.0125698   0.297  0.9907
##  4 - 6     0.83333333 0.2803405 41 -0.09590313 1.7625698   2.973  0.0244
##
## Results are averaged over the levels of: f_burner
## Confidence level used: 0.99
## Conf-level adjustment: tukey method for comparing a family of 4 estimates
## P value adjustment: tukey method for comparing a family of 4 estimates
```

From the table above we see that none of the p-values are below $\alpha = .01$, so we cannot detect any pairwise differences between the salt levels at the $\alpha = .01$ α-level. Equivalently, all the confidence intervals contain zero. If we let factor A represent the salt, then the we can see that the levels of A can be expressed as

$$\alpha_6 = \alpha_2 = \alpha_0 = \alpha_4$$

```
summary(contrast(em.burner, method ="tukey"), infer = c(TRUE,TRUE), level=.99)
```

```
##  contrast   estimate        SE df   lower.CL   upper.CL t.ratio p.value
##  1 - 2     2.5000000 0.2803405 41  1.5707635  3.4292365   8.918  <.0001
##  1 - 3     0.5833333 0.2803405 41 -0.3459031  1.5125698   2.081  0.1764
##  1 - 4    -1.7500000 0.2803405 41 -2.6792365 -0.8207635  -6.242  <.0001
##  2 - 3    -1.9166667 0.2803405 41 -2.8459031 -0.9874302  -6.837  <.0001
##  2 - 4    -4.2500000 0.2803405 41 -5.1792365 -3.3207635 -15.160  <.0001
##  3 - 4    -2.3333333 0.2803405 41 -3.2625698 -1.4040969  -8.323  <.0001
##
## Results are averaged over the levels of: f_salt
## Confidence level used: 0.99
## Conf-level adjustment: tukey method for comparing a family of 4 estimates
## P value adjustment: tukey method for comparing a family of 4 estimates
```

From the table above we see that only once of the p-values is not below $\alpha = .01$, so we cannot detect any differences between the burner level 1 and burner level 3 at the $\alpha = .01$ α-level. If we let factor B represent the burner, then equivalently the confidence interval for $\beta_1 - \beta_3$ is the only interval that contains zero. So we can see that the levels of B can be expressed as

$$\beta_2 < \beta_3 = \beta_1 < \beta_4$$