HW#8 Solution (week10 HW)

Due 11/7/2019

HW 22.7 Productivity improvement - CH22PR07.txt

```r
HW7<- read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-10/CH22PR07.txt"))
names(HW7) = c("Response",  "Treatment", "units", "X")
HW7$Treatment = as.factor(HW7$Treatment)
str(HW7)
```

```
## 'data.frame':    27 obs. of  4 variables:
##  $ Response : num  7.6 8.2 6.8 5.8 6.9 6.6 6.3 7.7 6 6.7 ...
##  $ Treatment: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 2 ...
##  $ units    : int  1 2 3 4 5 6 7 8 9 1 ...
##  $ X        : num  8.2 7.9 7 5.7 7.2 7 6.5 7.9 6.3 8.8 ...
```

(a) residuals

```r
Indicator1 = (HW7$Treatment=="1")*1 + (HW7$Treatment=="3")*(-1)
Indicator2 = (HW7$Treatment=="2")*1 + (HW7$Treatment=="3")*(-1)

# center the observation
(meanX= mean( HW7$X))
```

```
## [1] 9.4
```

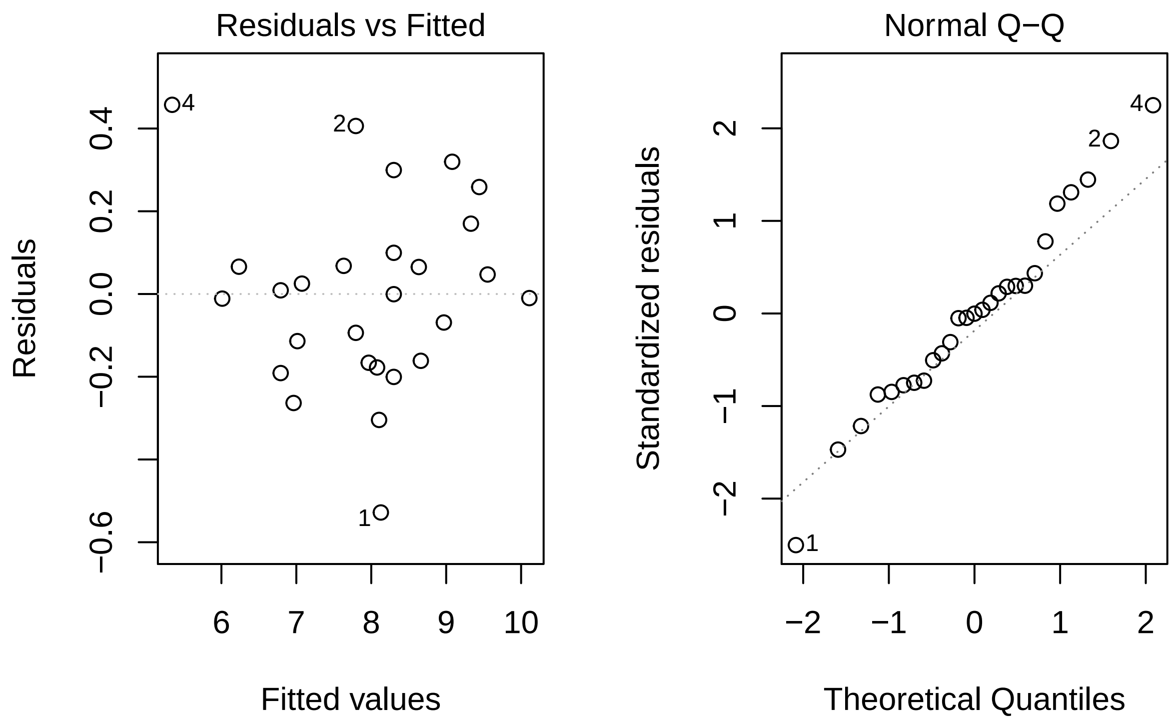
```r
X.centered = HW7$X - meanX

LM.full   =  lm( Response~ Indicator1 + Indicator2 + X.centered, data=HW7 )
LM.full$residuals
```

```
##             1             2             3             4             5
## -0.5281201408  0.4061297979  0.0088796137  0.4572960144 -0.1139536787
##             6             7             8             9            10
## -0.1911203863  0.0659628447 -0.0938702021 -0.0112038628 -0.2634585482
##            11            12            13            14            15
## -0.2004583026  0.3196251739  0.2995416974 -0.1662083640  0.0680415746
##            16            17            18            19            20
## -0.0689581799 -0.1776250102 -0.0004583026  0.0652917587  0.0251248056
##            21            22            23            24            25
##  0.0995416974 -0.1614862100  0.2585972666 -0.0099026107 -0.3044029790
##            26            27
##  0.0471806203  0.1700139128
```

(b) plots

```r
par(mfrow=c(1,2))
plot(LM.full, 1,  add.smooth = F)
plot(LM.full, 2)
```



Results: The constant variance and normal errors assumptions seem OK.

(C) test slope

```r
# Adding interaction term
LM.Interaction  =  lm( Response~ Indicator1 + Indicator2 + X.centered+
                         Indicator1:X.centered + Indicator2:X.centered,
                         data=HW7 )

# compare two models
anova(LM.full, LM.Interaction)
```

```
## Analysis of Variance Table
##
## Model 1: Response ~ Indicator1 + Indicator2 + X.centered
## Model 2: Response ~ Indicator1 + Indicator2 + X.centered + Indicator1:X.centered +
##     Indicator2:X.centered
##   Res.Df     RSS Df Sum of Sq     F  Pr(>F)
## 1     23 1.31753
## 2     21 0.95718  2   0.36035 3.953 0.03491 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: P=0.035 , not significant at level of 0.01.

(d) test if the linear effect of X is significant.

```
summary(LM.full)
```

```
##
## Call:
## lm(formula = Response ~ Indicator1 + Indicator2 + X.centered,
##     data = HW7)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.52812 -0.16385 -0.00046  0.08379  0.45730
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.80627    0.05082 153.599  < 2e-16 ***
## Indicator1   1.65885    0.19386   8.557 1.33e-08 ***
## Indicator2  -0.17431    0.06418  -2.716   0.0123 *
## X.centered   1.11417    0.07116  15.658 9.27e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2393 on 23 degrees of freedom
## Multiple R-squared:  0.9629, Adjusted R-squared:  0.958
## F-statistic: 198.8 on 3 and 23 DF,  p-value: < 2.2e-16
```

Results: The linear coefficient estimate= 1.114, with a p-value <0.001, suggesting a significant effect. The MSE for the model has df=23.

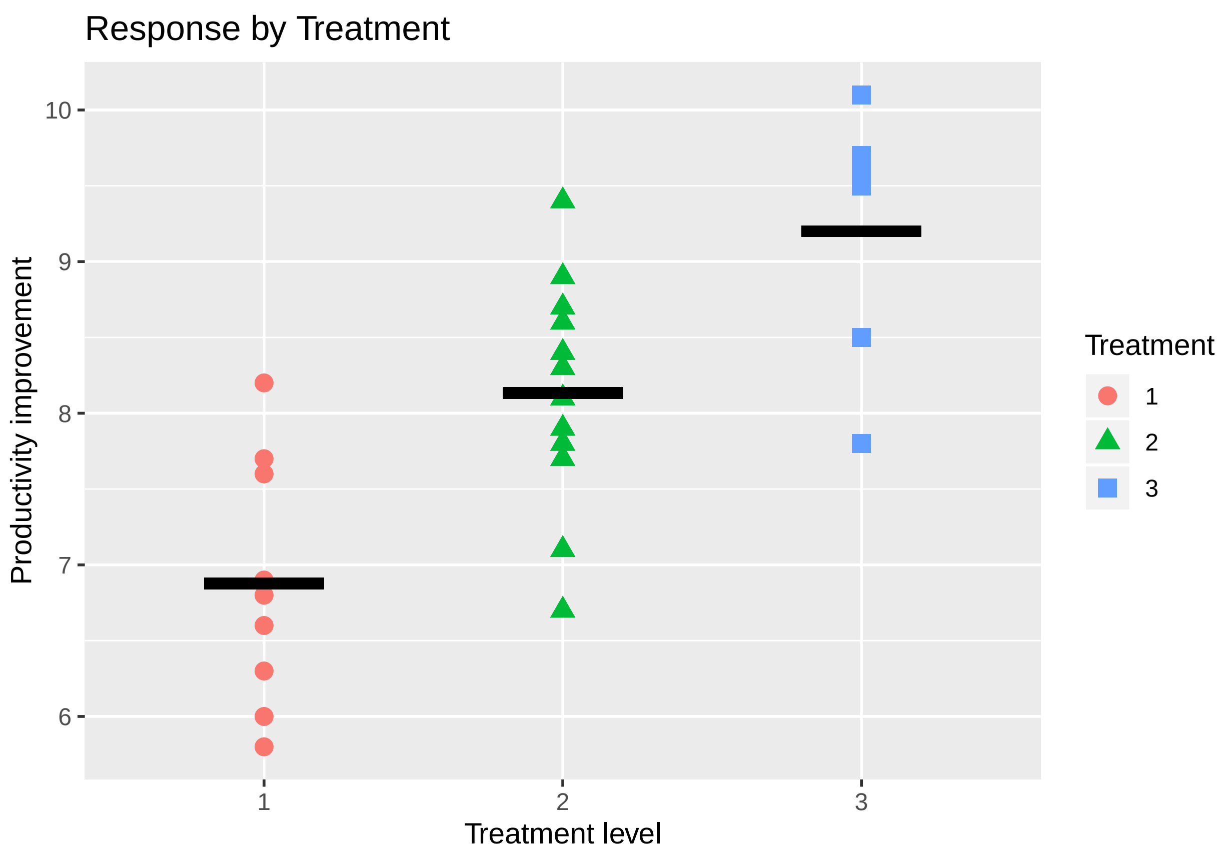
HW 22.8 Productivity improvement

(a) plot treatment means

```
library(ggplot2)

Trt.means= as.numeric(with(HW7, by(Response , Treatment, mean) ))

ggplot(data=HW7, aes(x= Treatment, y= Response,color= Treatment, shape=  Treatment))+
  geom_point(size=3) +
  geom_segment(aes(x = 0.8, y = Trt.means[1], xend = 1.2, yend = Trt.means[1]),size=2, col=1)+
  geom_segment(aes(x = 1.8, y = Trt.means[2], xend = 2.2, yend = Trt.means[2]),size=2, col=1)+
  geom_segment(aes(x = 2.8, y = Trt.means[3], xend = 3.2, yend = Trt.means[3]),size=2, col=1)+
  labs(title="Response by Treatment",
       x="Treatment level", y="Productivity improvement")
```



(b)-(d)

b. Full model: $Y_{ij} = \mu_{.} + \tau_1 I_{ij1} + \tau_2 I_{ij2} + \gamma x_{ij} + \varepsilon_{ij}$, $(\bar{X}_{..} = 9.4)$.

Reduced model: $Y_{ij} = \mu_{.} + \gamma x_{ij} + \varepsilon_{ij}$.

c. Full model: $\hat{Y} = 7.80627 + 1.65885 I_1 - .17431 I_2 + 1.11417x$, $SSE(F) = 1.3175$

Reduced model: $\hat{Y} = 7.95185 + .54124x$, $SSE(R) = 5.5134$

H_0: $\tau_1 = \tau_2 = 0$, H_a: not both τ_1 and τ_2 equal zero.

$F^* = (4.1959/2) \div (1.3175/23) = 36.625$, $F(.95; 2, 23) = 3.42$.

If $F^* \leq 3.42$ conclude H_0, otherwise H_a. Conclude H_a. P-value $= 0+$

```
options(contrasts = c("contr.sum", "contr.poly"))
library(car)
```

```
## Loading required package: carData
```

```
# Each factor is a categorical variable in the model
LM.full2  =  lm( Response ~ Treatment+X, data=HW7 )

# use Anova function in car package to get SS3
Anova(LM.full2, type="III")
```

```
## Anova Table (Type III tests)
##
## Response: Response
##              Sum Sq Df F value    Pr(>F)
## (Intercept)  0.8622  1  15.052 0.0007582 ***
## Treatment    4.1958  2  36.623 7.095e-08 ***
## X           14.0447  1 245.176 9.274e-14 ***
## Residuals    1.3175 23
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results (d): MSE = 1.3175/23= .0573

(e) Estimate treatment =2 and X=9

```
# center x
9-9.4
```

```
## [1] -0.4
```

```
# get coefficent and variance-covariance
coef(LM.full)
```

```
## (Intercept)  Indicator1  Indicator2  X.centered
##   7.8062717   1.6588482  -0.1743132   1.1141665
```

```
vcov(LM.full)
```

```
##              (Intercept)   Indicator1    Indicator2   X.centered
## (Intercept)  0.0025829193 -0.003248320 -0.0004516222 -0.001200159
## Indicator1  -0.0032483197  0.037582855 -0.0041739889  0.012957962
## Indicator2  -0.0004516222 -0.004173989  0.0041192942 -0.001078267
## X.centered  -0.0012001585  0.012957962 -0.0010782674  0.005063169
```

e. $\hat{Y} = \hat{\mu}_{.} + \hat{\tau}_2 - .4\hat{\gamma} = 7.18629$, $s^2\{\hat{\mu}_{.}\} = .00258$, $s^2\{\hat{\tau}_2\} = .00412$, $s^2\{\hat{\gamma}\} = .00506$, $s\{\hat{\mu}_{.}, \hat{\tau}_2\} = -.00045$, $s\{\hat{\tau}_2, \hat{\gamma}\} = -.00108$, $s\{\hat{\mu}_{.}, \hat{\gamma}\} = -.00120$, $s\{\hat{Y}\} = .09183$, $t(.975; 23) = 2.069$, $7.18629 \pm 2.069(.09183)$, $6.996 \leq \mu_{.} + \tau_2 - .4\gamma \leq 7.376$

(f) Pairwise difference with Bonferroni

```
library(emmeans)
fit.emm <- emmeans( LM.full2, ~ Treatment)

# CI with adjustment for MCP
confint(pairs(fit.emm), adjust = "Bonferroni", level=.9 )
```

```
##  contrast estimate    SE df lower.CL upper.CL
##  1 - 2        1.83 0.224 23    1.327     2.34
```

```
##  1 - 3        3.14 0.371 23    2.303     3.98
##  2 - 3        1.31 0.193 23    0.873     1.75
##
## Confidence level used: 0.9
## Conf-level adjustment: bonferroni method for 3 estimates
```

HW 22.15 Cash offers - CH22PR15.txt

```r
HW15 <- read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-10/CH22PR15.txt"))
names(HW15) = c("Response",  "factorA","factorB", "units", "X")

HW15$factorA = as.factor(HW15$factorA)
HW15$factorB = as.factor(HW15$factorB)

str(HW15)
```

```
## 'data.frame':    36 obs. of  5 variables:
##  $ Response: num  21 23 19 22 22 23 21 22 20 21 ...
##  $ factorA : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
##  $ factorB : Factor w/ 2 levels "1","2": 1 1 1 1 1 1 2 2 2 2 ...
##  $ units   : int  1 2 3 4 5 6 1 2 3 4 ...
##  $ X       : num  3 5.1 1 4.4 2.7 4.9 3.5 4.2 2.2 3.1 ...
```

(a) residuals

```r
(meanX= mean( HW15$X))
```

```
## [1] 3.408333
```

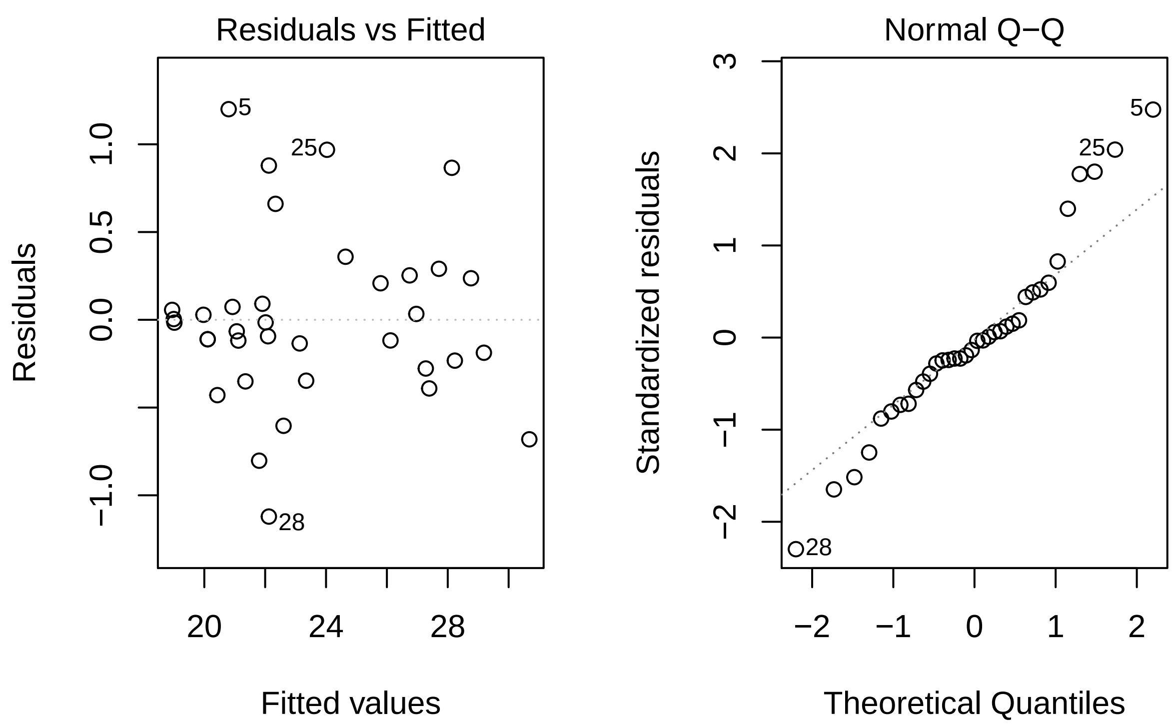
```r
X.centered = HW15$X - meanX

LM15.full  =  lm( Response~ factorA*factorB +
                         X.centered, data=HW15 )
LM15.full$residual
```

```
##            1            2            3            4            5
## -0.118369221 -0.346933033  0.004072505 -0.604078429  1.199997038
##            6            7            8            9           10
## -0.134688861 -0.351020348 -0.093874952  0.028566774  0.073467998
##           11           12           13           14           15
## -0.016334449  0.359194977 -0.680942649  0.865987422 -0.117692938
##           16           17           18           19           20
##  0.290475767 -0.391157974  0.033330372  0.208156858 -0.187749904
##           21           22           23           24           25
##  0.253058081 -0.232651127 -0.277552350  0.236738441  0.968713286
##           26           27           28           29           30
## -0.014967074  0.878910839 -1.121089161  0.091155012 -0.802722902
##           31           32           33           34           35
##  0.660550504  0.056457266 -0.110885684 -0.065984460 -0.429251942
##           36
## -0.110885684
```

(b) plot

```
par(mfrow=c(1,2))
plot(LM15.full, 1,  add.smooth = F)
plot(LM15.full, 2)
```



Results: The constant variance and normal errors assumptions seem OK.

(C) test constant slope

```
(meanX= mean( HW15$X))
```

```
## [1] 3.408333
```

```
X.centered = HW15$X - meanX

LM15.Int  =  lm( Response~ factorA*factorB* X.centered, data=HW15 )

anova( LM15.full, LM15.Int)
```

```
## Analysis of Variance Table
##
## Model 1: Response ~ factorA * factorB + X.centered
## Model 2: Response ~ factorA * factorB * X.centered
```

```
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     29 8.2941
## 2     24 6.1765  5    2.1176 1.6456 0.1863
```

$$\begin{aligned}Y_{ijk} &= \mu_{..} + \alpha_1 I_{ijk1} + \alpha_2 I_{ijk2} + \beta_1 I_{ijk3} + (\alpha\beta)_{11} I_{ijk1} I_{ijk3} \\ &+ (\alpha\beta)_{21} I_{ijk2} I_{ijk3} + \gamma x_{ijk} + \delta_1 I_{ijk1} x_{ijk} + \delta_2 I_{ijk2} x_{ijk} \\ &+ \delta_3 I_{ijk3} x_{ijk} + \delta_4 I_{ijk1} I_{ijk3} x_{ijk} + \delta_5 I_{ijk2} I_{ijk3} x_{ijk} + \epsilon_{ijk}\end{aligned}$$

H_0: all δ_i equal zero ($i = 1, ..., 5$), H_a: not all δ_i equal zero.

$SSE(R) = 8.2941$, $SSE(F) = 6.1765$,

$F^* = (2.1176/5) \div (6.1765/24) = 1.646$, $F(.99; 5, 24) = 3.90$.

If $F^* \leq 3.90$ conclude H_0, otherwise H_a. Conclude H_0. P-value $= .19$

HW 22.16 Cash offers

(a-e)

a. $$\begin{aligned}Y_{ijk} &= \mu_{..} + \alpha_1 I_{ijk1} + \alpha_2 I_{ijk2} + \beta_1 I_{ijk3} + (\alpha\beta)_{11} I_{ijk1} I_{ijk3} \\ &+ (\alpha\beta)_{21} I_{ijk2} I_{ijk3} + \gamma x_{ijk} + \epsilon_{ijk}\end{aligned}$$

$$I_{ijk1} = \begin{cases} 1 & \text{if case from level 1 for factor } A \\ -1 & \text{if case from level 3 for factor } A \\ 0 & \text{otherwise} \end{cases}$$

$$I_{ijk2} = \begin{cases} 1 & \text{if case from level 2 for factor } A \\ -1 & \text{if case from level 3 for factor } A \\ 0 & \text{otherwise} \end{cases}$$

$$I_{ijk3} = \begin{cases} 1 & \text{if case from level 1 for factor } B \\ -1 & \text{if case from level 2 for factor } B \end{cases}$$

$x_{ijk} = X_{ijk} - \bar{X}_{...} \qquad (\bar{X}_{...} = 3.4083)$

$\hat{Y} = 23.55556 - 2.15283 I_1 + 3.68152 I_2 + .20907 I_3 - .06009 I_1 I_3 - .04615 I_2 I_3 + 1.06122 x$

$SSE(F) = 8.2941$

b. Interactions:

$Y_{ijk} = \mu_{..} + \alpha_1 I_{ijk1} + \alpha_2 I_{ijk2} + \beta_1 I_{ijk3} + \gamma x_{ijk} + \epsilon_{ijk}$

$\hat{Y} = 23.55556 - 2.15400 I_1 + 3.67538 I_2 + .20692 I_3 + 1.07393 x$

$SSE(R) = 8.4889$

Factor A:

$Y_{ijk} = \mu_{..} + \beta_1 I_{ijk3} + (\alpha\beta)_{11} I_{ijk1} I_{ijk3} + (\alpha\beta)_{21} I_{ijk2} I_{ijk3} + \gamma x_{ijk} + \epsilon_{ijk}$

$\hat{Y} = 23.55556 + .12982 I_3 + .01136 I_1 I_3 + .06818 I_2 I_3 + 1.52893x$

$SSE(R) = 240.7835$

Factor B:

$$Y_{ijk} = \mu_{..} + \alpha_1 I_{ijk1} + \alpha_2 I_{ijk2} + (\alpha\beta)_{11} I_{ijk1} I_{ijk3} + (\alpha\beta)_{21} I_{ijk2} I_{ijk3} + \gamma x_{ijk} + \epsilon_{ijk}$$

$\hat{Y} = 23.55556 - 2.15487 I_1 + 3.67076 I_2 - .05669 I_1 I_3 - .04071 I_2 I_3 + 1.08348x$

$SSE(R) = 9.8393$

c. H_0: $(\alpha\beta)_{11} = (\alpha\beta)_{21} = 0$, H_a: not both $(\alpha\beta)_{11}$ and $(\alpha\beta)_{21}$ equal zero.

$F^* = (.1948/2) \div (8.2941/29) = .341$, $F(.95; 2, 29) = 3.33$.

If $F^* \leq 3.33$ conclude H_0, otherwise H_a. Conclude H_0. P-value $= .714$

d. H_0: $\alpha_1 = \alpha_2 = 0$, H_a: not both α_1 and α_2 equal zero.

$F^* = (232.4894/2) \div (8.2941/29) = 406.445$, $F(.95; 2, 29) = 3.33$.

If $F^* \leq 3.33$ conclude H_0, otherwise H_a. Conclude H_a. P-value $= 0+$

e. H_0: $\beta_1 = 0$, H_a: $\beta_1 \neq 0$.

$F^* = (1.5452/1) \div (8.2941/29) = 5.403$, $F(.95; 1, 29) = 4.18$.

If $F^* \leq 4.18$ conclude H_0, otherwise H_a. Conclude H_a. P-value $= .027$

(a-e) using R

```
IndicatorA1 = (HW15$factorA =="1")*1 + (HW15$factorA =="3")*(-1)
IndicatorA2 = (HW15$factorA =="2")*1 + (HW15$factorA =="3")*(-1)

IndicatorB = (HW15$factorB =="1")*1 + (HW15$factorB =="2")*(-1)

# center the observation
LM15.fullB  =  lm( Response~ IndicatorA1 + IndicatorA2 + IndicatorB+
                     IndicatorA1*IndicatorB + IndicatorA1*IndicatorB+
                         X.centered, data=HW15 )
```

generate SS III for testing

```
options(contrasts = c("contr.sum", "contr.poly"))
library(car)

# use Anova function in car package to get SS3
Anova(LM15.full, type="III")
```

```
## Anova Table (Type III tests)
##
## Response: Response
##                   Sum Sq Df    F value    Pr(>F)
## (Intercept)      19975.1  1 69842.2532 < 2.2e-16 ***
## factorA            232.5  2   406.4455 < 2.2e-16 ***
## factorB              1.5  1     5.4027   0.02731 *
## X.centered          63.4  1   221.5799 4.092e-15 ***
## factorA:factorB      0.2  2     0.3405   0.71422
## Residuals            8.3 29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: 1) no interaction (P-value=0.71) 2) significant factor A ($P<0.001$) 3) significant factor A ($P=0.027$).

(f)

```
library(emmeans)
fit.emmA <- emmeans( LM15.full, ~ factorA )
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
fit.emmB <- emmeans( LM15.full, ~ factorB  )
```

```
## NOTE: Results may be misleading due to involvement in interactions
```

```
# CI with adjustment for MCP
pairs(fit.emmA)
```

```
##  contrast estimate    SE df t.ratio p.value
##  1 - 2      -5.834 0.220 29 -26.507 <.0001
##  1 - 3      -0.624 0.223 29  -2.793 0.0241
##  2 - 3       5.210 0.231 29  22.555 <.0001
##
## Results are averaged over the levels of: factorB
## P value adjustment: tukey method for comparing a family of 3 estimates
```

```
pairs(fit.emmB)
```

```
##  contrast estimate   SE df t.ratio p.value
##  1 - 2       0.418 0.18 29 2.324   0.0273
##
## Results are averaged over the levels of: factorA
```

Results: We can get the paired difference estimates and SE from the R output. Then use Bonferroni methods as follows:

f. $\hat{D}_1 = \hat{\alpha}_1 - \hat{\alpha}_2 = -5.83435$, $\hat{D}_2 = \hat{\alpha}_1 - \hat{\alpha}_3 = 2\hat{\alpha}_1 + \hat{\alpha}_2 = -.62414$,

$\hat{D}_3 = \hat{\alpha}_2 - \hat{\alpha}_3 = 2\hat{\alpha}_2 + \hat{\alpha}_1 = 5.21021$, $\hat{D}_4 = \hat{\beta}_1 - \hat{\beta}_2 = 2\hat{\beta}_1 = .41814$,

$s^2\{\hat{\alpha}_1\} = .01593$, $s^2\{\hat{\alpha}_2\} = .01708$, $s\{\hat{\alpha}_1, \hat{\alpha}_2\} = -.00772$, $s^2\{\hat{\beta}_1\} = .00809$,

$s\{\hat{D}_1\} = .22011$, $s\{\hat{D}_2\} = .22343$, $s\{\hat{D}_3\} = .23102$, $s\{\hat{D}_4\} = .17989$,

$B = t(.9875; 29) = 2.364$

$$\begin{array}{rr}
-5.83435 \pm 2.364(.22011) & -6.355 \leq D_1 \leq -5.314 \\
-.62414 \pm 2.364(.22343) & -1.152 \leq D_2 \leq -.096 \\
5.21021 \pm 2.364(.23102) & 4.664 \leq D_3 \leq 5.756 \\
.41814 \pm 2.364(.17989) & -.007 \leq D_4 \leq .843
\end{array}$$

HW 22.19 Auditor training - CH22PR19.txt

b. $Y_{ij} = \mu_{..} + \rho_1 I_{ij1} + \rho_2 I_{ij2} + \rho_3 I_{ij3} + \rho_4 I_{ij4} + \rho_5 I_{ij5} + \rho_6 I_{ij6}$
$+\rho_7 I_{ij7} + \rho_8 I_{ij8} + \rho_9 I_{ij9} + \tau_1 I_{ij10} + \tau_2 I_{ij11} + \gamma x_{ij} + \epsilon_{ij}$

$$I_{ij1} = \begin{array}{l} 1 \text{ if experimental unit from block } 1 \\ -1 \text{ if experimental unit from block } 10 \\ 0 \text{ otherwise} \end{array}$$

$I_{ij2}, \ldots, I_{ij9}$ are defined similarly

$$I_{ij10} = \begin{array}{l} 1 \text{ if experimental unit received treatment } 1 \\ -1 \text{ if experimental unit received treatment } 3 \\ 0 \text{ otherwise} \end{array}$$

$$I_{ij11} = \begin{array}{l} 1 \text{ if experimental unit received treatment } 2 \\ -1 \text{ if experimental unit received treatment } 3 \\ 0 \text{ otherwise} \end{array}$$

$x_{ij} = X_{ij} - \bar{X}_{..} \qquad (\bar{X}_{..} = 80.033333)$

c. $\hat{Y} = 77.10000 + 4.87199 I_1 + 3.87266 I_2 + 2.21201 I_3 + 3.22003 I_4$
$+1.23474 I_5 + .90876 I_6 - 1.09124 I_7 - 3.74253 I_8 - 4.08322 I_9$
$-6.50033 I_{10} - 2.49993 I_{11} + .00201 x$

$SSE(F) = 112.3327$

d. $Y_{ij} = \mu_{..} + \rho_1 I_{ij1} + \rho_2 I_{ij2} + \rho_3 I_{ij3} + \rho_4 I_{ij4} + \rho_5 I_{ij5} + \rho_6 I_{ij6}$
$+\rho_7 I_{ij7} + \rho_8 I_{ij8} + \rho_9 I_{ij9} + \gamma x_{ij} + \epsilon_{ij}$

$\hat{Y} = 77.10000 + 6.71567 I_1 + 5.67233 I_2 + 3.61567 I_3 + 4.09567 I_4$
$+1.14233 I_5 + .33233 I_6 - 1.66767 I_7 - 5.33100 I_8 - 5.18767 I_9 - .13000 x$

$SSE(R) = 1,404.5167$

e. H_0: $\tau_1 = \tau_2 = 0$, H_a: not both τ_1 and τ_2 equal zero.

$F^* = (1,292.18/2) \div (112.3327/17) = 97.777$, $F(.95; 2, 17) = 3.59$.

If $F^* \leq 3.59$ conclude H_0, otherwise H_a. Conclude H_a. P-value $= 0+$

f. $\hat{\tau}_1 = -6.50033$, $\hat{\tau}_2 = -2.49993$, $\hat{L} = -4.0004$, $L^2\{\hat{\tau}_1\} = .44162$, $s^2\{\hat{\tau}_2\} = .44056$,
$s\{\hat{\tau}_1, \hat{\tau}_2\} = -.22048$, $s\{\hat{L}\} = 1.1503$, $t(.975; 17) = 2.11$,

$-4.0004 \pm 2.11(1.1503)$, $-6.43 < L < -1.57$

(g)

```
HW19 <- read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-10/CH22PR19.txt"))
names(HW19) = c("Response",  "block","Treatment", "X")

LM19.full =  lm( Response~ factor(block )+ factor(Treatment ) +X, data=HW19 )


LM19.reduce =  lm( Response~ factor(block )+ factor(Treatment ) , data=HW19 )
anova( LM19.reduce, LM19.full)
```

```
## Analysis of Variance Table
##
## Model 1: Response ~ factor(block) + factor(Treatment)
## Model 2: Response ~ factor(block) + factor(Treatment) + X
##   Res.Df    RSS Df  Sum of Sq     F Pr(>F)
## 1     18 112.33
## 2     17 112.33  1 0.00066854 1e-04 0.9921
```

```
Anova( LM19.full)
```

```
## Anova Table (Type II tests)
##
## Response: Response
##                    Sum Sq Df F value    Pr(>F)
## factor(block)       74.38  9  1.2507    0.3298
## factor(Treatment) 1292.18  2 97.7771 4.735e-10 ***
## X                    0.00  1  0.0001    0.9921
## Residuals          112.33 17
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova( LM19.reduce)
```

```
## Anova Table (Type II tests)
##
## Response: Response
##                    Sum Sq Df  F value    Pr(>F)
## factor(block)      433.37  9   7.7157 0.0001316 ***
## factor(Treatment) 1295.00  2 103.7537 1.315e-10 ***
## Residuals          112.33 18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: The error variance (RSS) did not reduce substantially by adding the concomitant variable. The model MSE after adding X is about the same as before.

HW 22.21 Productivity improvement

```r
HW7$Difference = HW7$Response - HW7$X
summary(aov(Difference~Treatment, data=HW7))
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## Treatment    2 25.582  12.791   209.5 6.38e-16 ***
## Residuals   24  1.465   0.061                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(LM.full)
```

```
## 
## Call:
## lm(formula = Response ~ Indicator1 + Indicator2 + X.centered, 
##     data = HW7)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -0.52812 -0.16385 -0.00046  0.08379  0.45730 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)  7.80627    0.05082 153.599  < 2e-16 ***
## Indicator1   1.65885    0.19386   8.557 1.33e-08 ***
## Indicator2  -0.17431    0.06418  -2.716   0.0123 *  
## X.centered   1.11417    0.07116  15.658 9.27e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.2393 on 23 degrees of freedom
## Multiple R-squared:  0.9629, Adjusted R-squared:  0.958 
## F-statistic: 198.8 on 3 and 23 DF,  p-value: < 2.2e-16
```

(b)

For ANOVA model, MSE= 0.061 For ANCOVA mode, MSE= 0.2393^2 =0.0573,

The relative efficience= 0.061/0.0573 =1.065.
Because $\hat{\gamma} = 1.11$ in ANCOVA. Use of differences with the regular ANOVA model would yield similar MSE compared to ANCOVA model.