HW#9 Solution (week11 HW)

Due 11/14/2019

HW 23.6 cash offers - CH19PR10.txt

```
HW6 = read.table(url(
  "https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-11/CH19PR10.txt"))
names(HW6) = c("response", "age", "gender", "units")
# to make six treatments
HW6$Treatment = HW6$gender*(HW6$age+2)
HW6$age = as.factor(HW6$age); HW6$gender = as.factor(HW6$gender)
HW6$Treatment = as.factor( HW6$Treatment)
HW6B = HW6[-c(16, 33),] #dim(HW6B) #34 ,5
```

a. $Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$

$$Y_{ijk} = \mu_{..} + \alpha_1 X_{ijk1} + \alpha_2 X_{ijk2} + \beta_1 X_{ijk3} + (\alpha\beta)_{11} X_{ijk1} X_{ijk3}$$
$$+ (\alpha\beta)_{21} X_{ijk2} X_{ijk3} + \epsilon_{ijk}$$

$$X_{ijk1} = \begin{array}{l} 1 \text{ if case from level 1 for factor } A \\ -1 \text{ if case from level 3 for factor } A \\ 0 \text{ otherwise} \end{array}$$

$$X_{ijk2} = \begin{array}{l} 1 \text{ if case from level 2 for factor } A \\ -1 \text{ if case from level 3 for factor } A \\ 0 \text{ otherwise} \end{array}$$

$$X_{ijk3} = \begin{array}{l} 1 \text{ if case from level 1 for factor } B \\ -1 \text{ if case from level 2 for factor } B \end{array}$$

b. $\boldsymbol{\beta}$ entries: $\mu_{..}$, α_1, α_2, β_1, $(\alpha\beta)_{11}$, $(\alpha\beta)_{21}$

X entries:

A	B	Freq.		X_1	X_2	X_3	X_1X_3	X_2X_3
1	1	6	1	1	0	1	1	0
1	2	6	1	1	0	−1	−1	0
2	1	5	1	0	1	1	0	1
2	2	6	1	0	1	−1	0	−1
3	1	6	1	−1	−1	1	−1	−1
3	2	5	1	−1	−1	−1	1	1

c. $\mathbf{X}\boldsymbol{\beta}$ entries:

A	B	
1	1	$\mu_{..} + \alpha_1 + \beta_1 + (\alpha\beta)_{11}$
1	2	$\mu_{..} + \alpha_1 - \beta_1 - (\alpha\beta)_{11} = \mu_{..} + \alpha_1 + \beta_2 + (\alpha\beta)_{12}$
2	1	$\mu_{..} + \alpha_2 + \beta_1 + (\alpha\beta)_{21}$
2	2	$\mu_{..} + \alpha_2 - \beta_1 - (\alpha\beta)_{21} = \mu_{..} + \alpha_2 + \beta_2 + (\alpha\beta)_{22}$
3	1	$\mu_{..} - \alpha_1 - \alpha_2 + \beta_1 - (\alpha\beta)_{11} - (\alpha\beta)_{21} = \mu_{..} + \alpha_3 + \beta_1 + (\alpha\beta)_{31}$
3	2	$\mu_{..} - \alpha_1 - \alpha_2 - \beta_1 + (\alpha\beta)_{11} + (\alpha\beta)_{21} = \mu_{..} + \alpha_3 + \beta_2 + (\alpha\beta)_{32}$

d. $Y_{ijk} = \mu_{..} + \alpha_1 X_{ijk1} + \alpha_2 X_{ijk2} + \beta_1 X_{ijk3} + \epsilon_{ijk}$

e. Full model:

$\hat{Y} = 23.56667 - 2.06667X_1 + 4.16667X_2 + .36667X_3 - .20000X_1X_3 - .30000X_2X_3,$

$SSE(F) = 71.3333$

Reduced model:

$\hat{Y} = 23.59091 - 2.09091X_1 + 4.16911X_2 + .36022X_3,$

$SSE(R) = 75.5210$

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

$F^* = (4.1877/2) \div (71.3333/28) = .82$, $F(.95; 2, 28) = 3.34$.

If $F^* \leq 3.34$ conclude H_0, otherwise H_a. Conclude H_0. P-value $= .45$

f. A effects:

$\hat{Y} = 23.50000 + .17677X_3 - .01010X_1X_3 - .49495X_2X_3,$

$SSE(R) = 359.9394$

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

$F^* = (288.6061/2) \div (71.3333/28) = 56.64$, $F(.95; 2, 28) = 3.34$.

If $F^* \leq 3.34$ conclude H_0, otherwise H_a. Conclude H_a. P-value $= 0+$

B effects:

$\hat{Y} = 23.56667 - 2.06667X_1 + 4.13229X_2 - .17708X_1X_3 - .31146X_2X_3,$

$SSE(R) = 75.8708$

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

$F^* = (4.5375/1) \div (71.3333/28) = 1.78$, $F(.95; 1, 28) = 4.20$.

If $F^* \leq 4.20$ conclude H_0, otherwise H_a. Conclude H_0. P-value $= .19$

e) Full model

```r
# set indicator
IndicatorB1 = (HW6B$gender   =="1")*1 + (HW6B$gender  =="2")*(-1)

IndicatorA1 =  (HW6B$age=="1")*1  + (HW6B$age=="3")*(-1)
IndicatorA2 =  (HW6B$age=="2")*1 + (HW6B$age=="3")*(-1)


LM.full  =  lm( response~ IndicatorA1 + IndicatorA2 + IndicatorB1+
                  IndicatorA1:IndicatorB1 + IndicatorA2:IndicatorB1, data=HW6B )
round(coef(LM.full),3)
```

```
##             (Intercept)             IndicatorA1             IndicatorA2
##                  23.567                  -2.067                   4.167
##             IndicatorB1 IndicatorA1:IndicatorB1 IndicatorA2:IndicatorB1
##                   0.367                  -0.200                  -0.300
```

Testing interaction effects

```r
Reduced.NoAB  =  lm( response~ IndicatorA1 + IndicatorA2+ IndicatorB1 , data=HW6B )
coef(Reduced.NoAB)
```

```
## (Intercept) IndicatorA1 IndicatorA2 IndicatorB1
##  23.5909091  -2.0909091   4.1691105   0.3602151
```

```r
# anova to compare the two models
anova(Reduced.NoAB, LM.full)
```

```
## Analysis of Variance Table
##
## Model 1: response ~ IndicatorA1 + IndicatorA2 + IndicatorB1
## Model 2: response ~ IndicatorA1 + IndicatorA2 + IndicatorB1 + IndicatorA1:IndicatorB1 +
##     IndicatorA2:IndicatorB1
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     30 75.521
## 2     28 71.333  2    4.1877 0.8219 0.4499
```

f) Testing effect A and B

```r
Reduced.NoA   =  lm( response~  IndicatorB1  +
                    IndicatorA1:IndicatorB1 + IndicatorA2:IndicatorB1, data=HW6B )
coef(Reduced.NoA)
```

```
##             (Intercept)             IndicatorB1 IndicatorB1:IndicatorA1
##             23.50000000              0.17676768             -0.01010101
## IndicatorB1:IndicatorA2
##             -0.49494949
```

```r
# anova to compare the two models
anova(Reduced.NoA, LM.full)
```

```
## Analysis of Variance Table
##
## Model 1: response ~ IndicatorB1 + IndicatorA1:IndicatorB1 + IndicatorA2:IndicatorB1
## Model 2: response ~ IndicatorA1 + IndicatorA2 + IndicatorB1 + IndicatorA1:IndicatorB1 +
##     IndicatorA2:IndicatorB1
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     30 359.94
## 2     28  71.33  2    288.61 56.642 1.442e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
Reduced.NoB   =  lm( response~  IndicatorA1 + IndicatorA2   +
                    IndicatorA1:IndicatorB1 + IndicatorA2:IndicatorB1, data=HW6B )

coef(Reduced.NoB)
```

```
##             (Intercept)             IndicatorA1             IndicatorA2
##              23.5666667              -2.0666667               4.1322917
## IndicatorA1:IndicatorB1 IndicatorA2:IndicatorB1
##              -0.1770833              -0.3114583
```

```
# anova to compare the two models
anova(Reduced.NoB, LM.full)
```

```
## Analysis of Variance Table
##
## Model 1: response ~ IndicatorA1 + IndicatorA2 + IndicatorA1:IndicatorB1 +
##     IndicatorA2:IndicatorB1
## Model 2: response ~ IndicatorA1 + IndicatorA2 + IndicatorB1 + IndicatorA1:IndicatorB1 +
##     IndicatorA2:IndicatorB1
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     29 75.871
## 2     28 71.333  1    4.5375 1.7811 0.1928
```

g. $\hat{D}_1 = \hat{\alpha}_1 - \hat{\alpha}_2 = -6.23334$, $\hat{D}_2 = \hat{\alpha}_1 - \hat{\alpha}_3 = 2\hat{\alpha}_1 + \hat{\alpha}_2 = .03333$, $\hat{D}_3 = \hat{\alpha}_2 - \hat{\alpha}_3 = 2\hat{\alpha}_2 + \hat{\alpha}_1 = 6.26667$, $s^2\{\hat{\alpha}_1\} = .14625$, $s^2\{\hat{\alpha}_2\} = .15333$, $s\{\hat{\alpha}_1, \hat{\alpha}_2\} = -.07313$, $s\{\hat{D}_1\} = .6677$, $s\{\hat{D}_2\} = .6677$, $s\{\hat{D}_3\} = .6834$, $q(.90; 3, 28) = 3.026$, $T = 2.140$

$$\begin{array}{rr} -6.23334 \pm 2.140(.6677) & -7.662 \leq D_1 \leq -4.804 \\ .03333 \pm 2.140(.6677) & -1.396 \leq D_2 \leq 1.462 \\ 6.26667 \pm 2.140(.6834) & 4.804 \leq D_3 \leq 7.729 \end{array}$$

h. $\hat{L} = .3\bar{Y}_{12.} + .6\bar{Y}_{22.} + .1\bar{Y}_{32.} = .3(21.33333) + .6(27.66667) + .1(20.60000) = 25.06000$, $s\{\hat{L}\} = .4429$, $t(.975; 28) = 2.048$, $25.06000 \pm 2.048(.4429)$, $24.153 \leq L \leq 25.967$

R analysis (g)

```
library(emmeans)

LM.full2  =  lm( response~  gender*age, data=HW6B )

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

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

```
# CI with adjustment for MCP
confint(pairs(fit.emm), adjust = "tukey", level=0.9)
```

```
##  contrast estimate    SE df lower.CL upper.CL
##  1 - 2     -6.2333 0.668 28    -7.66    -4.80
##  1 - 3      0.0333 0.668 28    -1.40     1.46
##  2 - 3      6.2667 0.683 28     4.80     7.73
##
## Results are averaged over the levels of: gender
## Confidence level used: 0.9
## Conf-level adjustment: tukey method for comparing a family of 3 estimates
```

R analysis (f)

```
# get the treatment level
LM.mod2  =  lm( response~ Treatment, data=HW6B )
fit.emm2 <- emmeans( LM.mod2, ~ Treatment)

#Set the c1-c6 corresponding to treatment levels: male=0, female only
L = list( L1= c(0,0, 0, .30, .60,  0.1))

confint(contrast(fit.emm2, L, adjust='none'))
```

```
##  contrast estimate    SE df lower.CL upper.CL
##  L1           25.1 0.443 28     24.2       26
##
## Confidence level used: 0.95
```

HW 23.18: Auditor training data (CH21PR05.txt)

```
HW18 = read.table(
  url("https://raw.githubusercontent.com/npmldabook/Stat3119/master/Week-11/CH21PR05.txt"))
HW18 = HW18[-6,]
names(HW18) = c("Response", "block", "Training")
dim(HW18)
```

```
## [1] 29  3
```

```
table(HW18$block,HW18$Training)
```

```
##
##      1 2 3
##   1  1 1 1
##   2  1 1 0
##   3  1 1 1
##   4  1 1 1
##   5  1 1 1
##   6  1 1 1
##   7  1 1 1
##   8  1 1 1
##   9  1 1 1
##   10 1 1 1
```

. a. $Y_{ij} = \mu_{..} + \rho_i + \tau_j + \epsilon_{ij}$

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

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

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

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

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

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

c. <u>Full model</u>: $\hat{Y} = 77.15556 + 4.84444X_1 + 4.40000X_2 + 2.17778X_3$
$+3.17778X_4 + 1.17778X_5 + .84444X_6 - 1.15556X_7$
$-3.82222X_8 - 4.15556X_9 - 6.55556X_{10} - 2.55556X_{11}$

$SSE(F) = 110.6667$

<u>Reduced model</u>: $\hat{Y} = 76.70000 + 5.30000X_1 + .30000X_2 + 2.63333X_3$
$+3.63333X_4 + 1.63333X_5 + 1.30000X_6 - .70000X_7$
$-3.36667X_8 - 3.70000X_9$

$SSE(R) = 1,311.3333$

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

$F^* = (1,200.6666/2) \div (110.6667/17) = 92.22$, $F(.95; 2, 17) = 3.59$.

If $F^* \leq 3.59$ conclude H_0, otherwise H_a. Conclude H_a.

d. $\hat{L} = \hat{\tau}_2 - \hat{\tau}_3 = 2\hat{\tau}_2 + \hat{\tau}_1 = -11.66667$, $s^2\{\hat{\tau}_i\} = .44604$ $(i = 1, 2)$, $s\{\hat{\tau}_1, \hat{\tau}_2\} = -.20494$, $s\{\hat{L}\} = 1.1876$, $t(.975; 17) = 2.11$,

$-11.66667 \pm 2.11(1.1876)$, $-14.17 \leq L \leq -9.16$

(a)

```r
attach(HW18)
I.B1= (block=='1')*1+ (block=='10')*(-1)
I.B2= (block=='2')*1+ (block=='10')*(-1)
I.B3= (block=='3')*1+ (block=='10')*(-1)
I.B4= (block=='4')*1+ (block=='10')*(-1)
I.B5= (block=='5')*1+ (block=='10')*(-1)
I.B6= (block=='6')*1+ (block=='10')*(-1)
I.B7= (block=='7')*1+ (block=='10')*(-1)
I.B8= (block=='8')*1+ (block=='10')*(-1)
I.B9= (block=='9')*1+ (block=='10')*(-1)
I.T1= (Training=='1')*1+ (Training=='3')*(-1)
I.T2= (Training=='2')*1+ (Training=='3')*(-1)

LM2.full <- lm(Response~ I.B1+I.B2+I.B3+I.B4+I.B5+I.B6+I.B7+I.B8+I.B9+
              I.T1+I.T2   )
coef(LM2.full)
```

```
## (Intercept)        I.B1        I.B2        I.B3        I.B4        I.B5
##  77.1555556   4.8444444   4.4000000   2.1777778   3.1777778   1.1777778
##        I.B6        I.B7        I.B8        I.B9        I.T1        I.T2
##   0.8444444  -1.1555556  -3.8222222  -4.1555556  -6.5555556  -2.5555556
```

(b-c)

```r
LM2.reduced <- lm(Response~ I.B1+I.B2+I.B3+I.B4+I.B5+I.B6+I.B7+I.B8+I.B9  )
anova(LM2.full,LM2.reduced )
```

```
## Analysis of Variance Table
##
## Model 1: Response ~ I.B1 + I.B2 + I.B3 + I.B4 + I.B5 + I.B6 + I.B7 + I.B8 +
##     I.B9 + I.T1 + I.T2
## Model 2: Response ~ I.B1 + I.B2 + I.B3 + I.B4 + I.B5 + I.B6 + I.B7 + I.B8 +
##     I.B9
##   Res.Df     RSS Df Sum of Sq     F    Pr(>F)
## 1     17  110.67
## 2     19 1311.33 -2   -1200.7 92.22 7.474e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(d) get the variance=0.446 and covariance =-0.205 from variance and covariance matrix.

```r
round(vcov(LM2.full)[11:12, 11:12],3)
```

```
##        I.T1   I.T2
## I.T1  0.446 -0.205
## I.T2 -0.205  0.446
```

HW 23.20: cash offer , same data as problem 23.6

23.20. See Problem 19.10a. $L_1 = .3\mu_{11} + .6\mu_{21} + .1\mu_{31}$, $L_2 = .3\mu_{12} + .6\mu_{22} + .1\mu_{32}$.

H_0: $L_1 = L_2$, H_a: $L_1 \neq L_2$.

$\hat{L}_1 - \hat{L}_2 = 25.43332 - 25.05001 = .38331$, $MSE = 2.3889$, $s\{\hat{L}_1 - \hat{L}_2\} = .6052$,

$t^* = .38331/.6052 = .63$, $t(.975; 30) = 2.042$.

If $|t^*| \leq 2.042$ conclude H_0, otherwise H_a. Conclude H_0. P-value $= .53$

```
# this is based on the data without setting the missing
LM.new1  =  lm( response~  Treatment, data=HW6 )

fit.new <- emmeans( LM.new1, ~ Treatment)

# CI with adjustment for MCP
L = list( L1= c(.30, .60,  0.1 ,0, 0, 0),
          L1= c(0, 0, 0, .30, .60,  0.1),
          L=  c(.30, .60,  0.1, -.30, -.60,  -0.1))

contrast(fit.new, L, adjust='none')
```

```
##  contrast estimate    SE df t.ratio p.value
##  L1         25.433 0.428 30 59.429  <.0001
##  L1         25.050 0.428 30 58.534  <.0001
##  L           0.383 0.605 30  0.633  0.5313
```