HW5-ADA

Jing Qian (jq2282)

1.a. Give the design matrix corresponding the the model $y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$ where $\alpha_1+\alpha_2 = 0,\ \beta_1+\beta_2 = 0,\ (\alpha\beta)_{11}+(\alpha\beta)_{21} = 0,\ (\alpha\beta)_{12}+(\alpha\beta)_{22} = 0,\ (\alpha\beta)_{11}+(\alpha\beta)_{12} = 0,\ (\alpha\beta)_{21}+(\alpha\beta)_{22} = 0$.

The design matrix is as following if we define $A_i = 1$ when A is High and $A_i = -1$ when A is low; $B_j = 1$ when B is New and and $B_j = -1$ when B is Old:

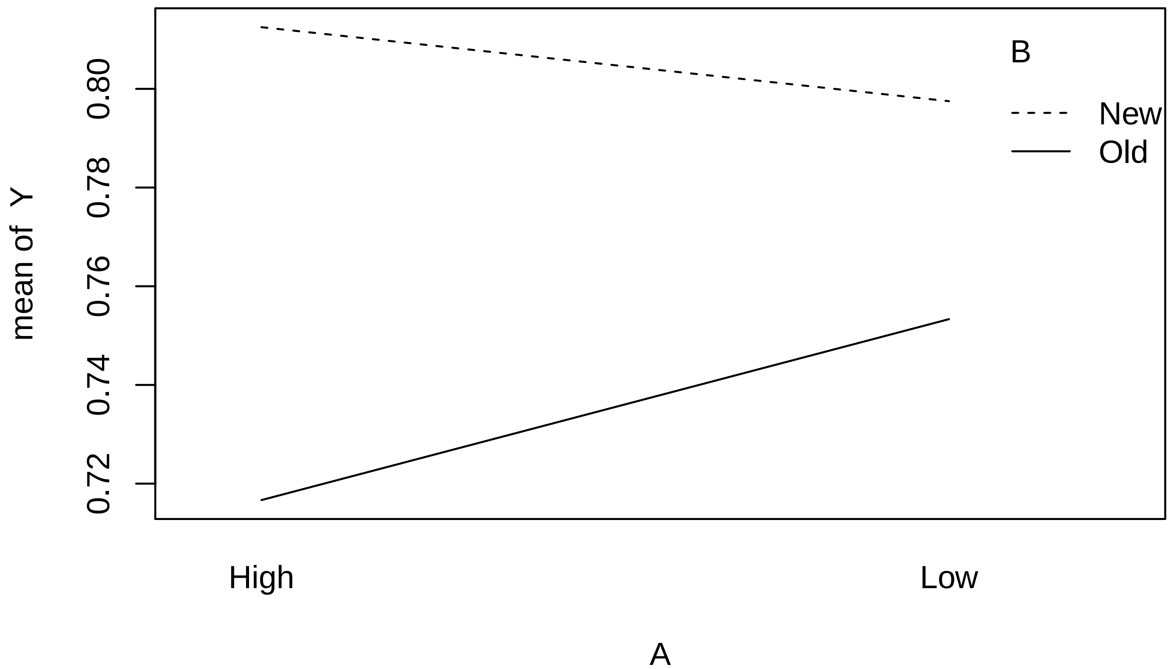
```
Y<-c(0.83, 0.78, 0.75, 0.83, 0.86, 0.67, 0.73, 0.72, 0.81, 0.85, 0.87, 0.68, 0.73, 0.74)
A<-c(rep('Low', 7), rep('High', 7))
B<-c(rep(c(rep('New', 4), rep('Old', 3)), 2))
data<-data.frame(Y, A, B)
#print(data)
```

```
model.matrix(~A*B, contrasts=list(A=contr.sum, B=contr.sum))
```

```
##    (Intercept) A1 B1 A1:B1
## 1            1 -1  1    -1
## 2            1 -1  1    -1
## 3            1 -1  1    -1
## 4            1 -1  1    -1
## 5            1 -1 -1     1
## 6            1 -1 -1     1
## 7            1 -1 -1     1
## 8            1  1  1     1
## 9            1  1  1     1
## 10           1  1  1     1
## 11           1  1  1     1
## 12           1  1 -1    -1
## 13           1  1 -1    -1
## 14           1  1 -1    -1
## attr(,"assign")
## [1] 0 1 2 3
## attr(,"contrasts")
## attr(,"contrasts")$A
##      [,1]
## High    1
## Low    -1
##
## attr(,"contrasts")$B
##     [,1]
## New    1
## Old   -1
```

1.b. Construct an interaction plot. Does it suggest that there is an interaction between fertilizer type and wheat type?

```
interaction.plot(A, B, Y)
```



The lines in the interaction plot are not parallel to each other, which suggests that there is interaction between fertilizer type and wheat type.

1.c. Test H0 : No interaction against Ha : there is an interaction, use $\alpha = 0.05$.

```
summary(aov(Y~A*B))
```

```
##             Df  Sum Sq  Mean Sq F value Pr(>F)
## A            1 0.00018 0.000179   0.046 0.8346
## B            1 0.01680 0.016800   4.321 0.0643 .
## A:B          1 0.00229 0.002288   0.588 0.4607
## Residuals   10 0.03888 0.003888
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the ANOVA table above, we could see that p value of the test is 0.4607, which is larger than α. So we could not reject H_0 and there is no interaction between A and B.

1.d. It is decided to fit a model without an interaction. Test H0 : there is no A effect against Ha : there is an A effect. Use $\alpha = 0.05$.

```
library(car)
model<-lm(Y~A+B)
Anova(model, type=3)
```

```
## Anova Table (Type III tests)
##
## Response: Y
##             Sum Sq Df  F value    Pr(>F)
## (Intercept) 3.2698  1 873.6183 7.838e-12 ***
## A           0.0002  1   0.0477   0.83110
## B           0.0168  1   4.4885   0.05771 .
## Residuals   0.0412 11
```

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

From the above ANOVA table with type III SS, we could see that the p value of the test is 0.83110, which is larger than α. So we accept H0 and there is no A effect.

1.e. It is decided to fit a model without an interaction. Test H0 : there is no B effect against Ha : there is an B effect. Use $\alpha = 0.05$.

From the above ANOVA table with type III SS, we could see that the p value of the test is 0.05771, which is larger than α. So we accept H0 and there is no B effect.

2.a. Estimate the β_is and interpret your result (see note for how to fit this model)

```r
y<-c(28.0,28.6,27.4,33.3,34.5,33.0,32.0,35.6,34.4,35.0,34.0,33.3,34.7,33.5,32.3,33.4,33.0,32.0,29.6,30.
D1<-factor(c(rep(0,3),rep(1,2),rep(0,4),rep(1,3),rep(0,3),1,rep(0,2),rep(1,2),rep(0,2)))
D2<-factor(c(rep(0,12),1,rep(0,3),rep(1,2),rep(0,2),rep(1,2)))
x1<-c(rep('A',3),rep('B',2),rep('A',4),rep('B',3),'C',rep('A',2),'B',rep('C',2),rep('B',2),rep('C',2))
x2<-c(rep(0,5),rep(1,8),rep(2,5),rep(3,4))
#data2<-data.frame(y,D1,D2,x1,x2)
#print(data2)
```

```r
fit2<-lm(y~D1+D2+x2)
summary(fit2)
```

```
##
## Call:
## lm(formula = y ~ D1 + D2 + x2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6171 -1.6321  0.5508  1.3756  4.0021
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  32.0171     1.0005  32.002   <2e-16 ***
## D11           1.5218     1.2650   1.203    0.245
## D21           0.5252     1.6194   0.324    0.749
## x2           -0.4192     0.6042  -0.694    0.497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.532 on 18 degrees of freedom
## Multiple R-squared:  0.09453,    Adjusted R-squared:  -0.05638
## F-statistic: 0.6264 on 3 and 18 DF,  p-value: 0.6072
```

When type A, the response is modelled as $Y_i = \beta_0 + \beta_3 x_2 + \epsilon_i$; When type B, the response is modelled as $Y_i = \beta_0 + \beta_1 + \beta_3 x_2 + \epsilon_i$; When type C, the response is modelled as $Y_i = \beta_0 + \beta_2 + \beta_3 x_2 + \epsilon_i$.

From the script above, the estimated $\beta_0 = 32.017,\ \beta_1 = 1.522,\ \beta_2 = 0.525,\ \beta_3 = -0.419$. However, among all these β_i, only β_0 is significant while others are not. The fitted result suggests that neither premium gasoline unleaded type (A, B, C) or amount of gasoline additive VST (0, 1, 2, 3 units) has significant contribution to the gasoline mileage obtained by an automobile called Encore.

2.b. Construct a 95% confidence interval for β_3 and interpret your result

```r
confint(fit2,"x2",level = 0.95)
```

```
##        2.5 %    97.5 %
## x2 -1.688644 0.8502126
```

The 95% confidence interval for β_3 is [-1.688644, 0.8502126], which covers 0. From the confidence interval result, we could see that amount of gasoline additive VST is not a significant variable to the gasoline mileage y, which agrees with the conclusion in 2.a.

2.c. Test H0 :$\beta_1 = \beta_2 = 0$ against Ha :Not H0 using α=0.05.

```r
Anova(lm(y~x1+x2),type=3)
```

```
## Anova Table (Type III tests)
##
## Response: y
##             Sum Sq Df   F value Pr(>F)
## (Intercept) 6567.0  1 1024.1237 <2e-16 ***
## x1             9.7  2    0.7574 0.4832
## x2             3.1  1    0.4814 0.4967
## Residuals    115.4 18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since β_1 and β_2 are dummy variables for variable x_1, we could turn the $H_0 : \beta_1 = \beta_2 = 0$ to H_0 : there is no x_1 effect. From the above ANOVA table with type III SS, we could see that the p value for the test is 0.4832, which is larger than α. So H_0 could not be rejected and the premium gasoline unleaded type (A, B, C) is not a significant variable, which agrees with the concluion in 2.a.