Homework 10 Solutions

1. The interaction term is significant according to the F-test

$$H_0 : \gamma_{ij} = 0 \quad \forall i,j$$
$$H_1 : \text{not } H_0.$$

See attached output:

$$F = 4.503$$
$$p = 0.01288 < .05$$

$\Rightarrow$ Reject H_0.

$\Rightarrow$ Model chosen is two-way anova model with interactions.

2. H_0 : linear regression model is "good"
   H_1 : one-way anova model is "good".

   See R output for the appropriate F-test.

$$F = 4.138$$
$$p = 0.058 > .05$$

$\Rightarrow$ Do not reject H_0.

3. H_0 : Linear regression
H_1 : quadratic regression.

$F = 1.249$
$p = 0.301$

$\Rightarrow$ Do not reject H_0.
$\Rightarrow$ Linear regression prefered.

4a. H_0 : Interactions $= 0$.
H_1 : Interactions $\neq 0$ for some.

$F = 0.0136$
$p = .9865$

$\Rightarrow$ Interactions are insignificant.

b. H_0 : drugs $= 0$
H_1 : not H_0.

$F = 32.8$
$p < .05$

$\Rightarrow$ Drugs have an effect.

H_0 : gender has no effect
H_1 : not H_0.

$F = 0.0016$
$p = .9688$

$\Rightarrow$ Gender has no effect.

c. We lose orthogonality & hence SS don't decompose nicely

* F-tests still work fine though.

9.5.10.
note these are cell means.

$$\hat{\mu} = \frac{6+7+7+12+\ldots+10}{12}$$

$$= 8$$

$$\hat{\alpha}_1 = \frac{6+7+7+12}{4} - 8.$$

$$= 0$$

$$\hat{\alpha}_2 = \frac{10+3+11+8}{4} - 8$$

$$= 0.$$

& $\hat{\alpha}_3 = 0$ as $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = 0$.

$$\hat{\beta}_1 = \frac{6+10+8}{3} - 8$$
$$= 0.$$

$$\hat{\beta}_2 = \frac{7+3+5}{3} - 8.$$
$$= -3.$$

$$\hat{\beta}_3 = \frac{7+11+9}{3} - 8 = 1$$

$$\hat{\beta}_4 = \frac{12+8+10}{3} - 8 = 2.$$

$$\hat{\gamma}_{11} = 6 - \cancel{8} - 8 + 8 = \cancel{4} - 2.$$
$$\hat{\gamma}_{12} = 7 - 8 - 5 + 8 = 2.$$

etc. (too lazy to do the rest ☺).

9.6.17.

$$\frac{L_0}{L_1} = \frac{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(\sum y_i^2)}}{(2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(y_i - \hat{\beta}x_i)^2}}$$

$$= C \cdot \exp\left[\left(\sum y_i^2 - \sum(y_i - \hat{\beta}x_i)^2\right) C_2\right].$$

$$= C \cdot \exp\left[C_2 \cdot \left(2\hat{\beta}\sum y_i x_i - \hat{\beta}^2 \sum x_i^2\right)\right]$$

What is $\hat{\beta}$?

$$\hat{\beta} = (A^TA)^{-1}A^Ty \quad , \quad A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \frac{\sum x_i y_i}{\sum x_i^2} \qquad \Rightarrow A^TA = \sum x_i^2$$

$$(A^TA)^{-1} = \frac{1}{\sum x_i^2}$$

$$\Rightarrow \frac{L_1}{L_0} = C \cdot \exp\left\{ C_2 \cdot \left(2 \cdot \frac{(\sum x_i y_i)^2}{\sum x_i^2} - \frac{(\sum x_i y_i)^2}{\sum x_i^2} \right)\right.$$

$$= C \cdot \exp\left\{ C_2 \cdot \left(\frac{(\sum x_i y_i)^2}{\sum x_i^2} \right)\right.$$

Depends on $\dfrac{(\sum x_i y_i)^2}{\sum x_i^2}$

So it's something to do with sums of normals.

~~$\Rightarrow \chi^2$ ish in dist'n~~.

$\sum x_i y_i$ – Normal distributed.

Question 1:

```
> anova(model.full, model.add)
Analysis of Variance Table

Model 1: data ~ A * B
Model 2: data ~ A + B
  Res.Df     RSS Df Sum of Sq     F  Pr(>F)
1     12 0.27500
2     18 0.89417 -6  -0.61917 4.503 0.01288 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> summary(model.full)

Call:
lm(formula = data ~ A * B)

Residuals:
   Min     1Q Median     3Q    Max 
  -0.2   -0.1    0.0    0.1    0.2 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  42.5500     0.1070 397.502  < 2e-16 ***
A2           -0.3500     0.1514  -2.312  0.03933 *  
A3            1.1500     0.1514   7.597 6.36e-06 ***
B2           -0.4500     0.1514  -2.973  0.01164 *  
B3            1.2000     0.1514   7.927 4.13e-06 ***
B4           -0.2000     0.1514  -1.321  0.21109    
A2:B2        -0.1500     0.2141  -0.701  0.49688    
A3:B2         0.1500     0.2141   0.701  0.49688    
A2:B3        -0.3500     0.2141  -1.635  0.12802    
A3:B3        -0.7500     0.2141  -3.503  0.00436 ** 
A2:B4        -0.4500     0.2141  -2.102  0.05735 .  
A3:B4        -0.5500     0.2141  -2.569  0.02459 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1514 on 12 degrees of freedom
Multiple R-squared:  0.9833,	Adjusted R-squared:  0.968 
F-statistic:  64.2 on 11 and 12 DF,  p-value: 6.436e-09
```

Question 2:

```
> model.linear <- lm(y~x)
> model.anova <- lm(y~factor(x))
> anova(model.linear, model.anova)
Analysis of Variance Table

Model 1: y ~ x
```

```
Model 2: y ~ factor(x)
  Res.Df    RSS Df Sum of Sq      F  Pr(>F)
1     10 36.621
2      8 18.000  2    18.621 4.1379 0.05837 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Question 3:

```
> anova(model.linear, model.quad)
Analysis of Variance Table

Model 1: y ~ x
Model 2: y ~ x + I(x^2)
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1      8 163019
2      7 138340  1     24679 1.2488 0.3007
```

Question 4:

```
> anova(model.int, model.add)
Analysis of Variance Table

Model 1: IQ ~ gender * drug
Model 2: IQ ~ gender + drug
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1      8 131.00
2     10 131.45 -2  -0.44554 0.0136 0.9865

> anova(model.add, model.drug)
Analysis of Variance Table

Model 1: IQ ~ gender + drug
Model 2: IQ ~ drug
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     10 131.45
2     11 131.47 -1 -0.021122 0.0016 0.9688

> anova(model.add, model.gender)
Analysis of Variance Table

Model 1: IQ ~ gender + drug
Model 2: IQ ~ gender
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1     10 131.45
2     12 993.88 -2   -862.43 32.806 4.046e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```