MA 4710 Homework 9

Benjamin Hendrick

April 6, 2016

Problem 7.3

Load the data and rename the columns.

```
brand <- read.table("~/GitHub/MA-4710/Homework 8/brand.txt", quote="\"", comment.char="")
names(brand) <- c("Yi","Xi1", "Xi2")
```

Part A

Find the linear models for $Y + i = \beta_1 X_1$, $Y_i = \beta_2 X_2$, and $Y_i = \beta_1 X_1 + \beta_2 X_2$

```
brand.fit1 <- lm(Yi ~ Xi1, data = brand)
brand.fit2 <- lm(Yi ~ Xi2, data = brand)
brand.fit3 <- lm(Yi ~ Xi1 + Xi2, data = brand)
```

Find the ANOVA table for the third model ($Y_i = \beta_1 X_1 + \beta_2 X_2$) using the `anova` function.

```
anova(brand.fit3)
```

```
## Analysis of Variance Table
##
## Response: Yi
##           Df  Sum Sq Mean Sq F value    Pr(>F)
## Xi1        1 1566.45 1566.45 215.947 1.778e-09 ***
## Xi2        1  306.25  306.25  42.219 2.011e-05 ***
## Residuals 13   94.30    7.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the above outpout, the following ANOVA table is produced:

Name	*df*	SS	MS	F^*
Responce	2	1872.8	936.4	215.947
X_1	1	1566.45	1566.45	216.06
X_2	1	306.25	306.25	42.219
Error	13	94.30	7.25	
Total	17	1967.1		

Part B

$$H_0 : \beta_2 = 0$$

$$H_a : \beta_2 \neq 0$$

```
f <- (anova(brand.fit3)[2,2]/anova(brand.fit3)[2,1])/
                ((anova(brand.fit3)[2,2]+anova(brand.fit3)[3,2])/
                   (anova(brand.fit3)[2,1]+anova(brand.fit3)[3,1]))
```

The F test statistic to see if X_2 can be dropped is 10.704032. The critical value is 3.8979442

Becasue the test statistic is greater than the critical value, we reject the null hypothesis H_0. Therefore we should not drop X_2.

Problem 7.12

```
R2_Y1 <- anova(brand.fit1)[1,2]/(anova(brand.fit1)[1,2]+anova(brand.fit1)[2,2])
R2_Y2 <- anova(brand.fit2)[1,2]/(anova(brand.fit2)[1,2]+anova(brand.fit2)[2,2])
R2_12 <- (anova(brand.fit3)[1,2]+anova(brand.fit3)[2,2])/
          (anova(brand.fit3)[1,2]+anova(brand.fit3)[2,2]+anova(brand.fit3)[3,2])
R2_Y12 <- anova(brand.fit3)[1,2]/(anova(brand.fit3)[1,2]+anova(brand.fit3)[3,2])
R2_Y21 <- anova(brand.fit3)[2,2]/(anova(brand.fit3)[2,2]+anova(brand.fit3)[3,3])
R2 <- (anova(brand.fit3)[1,2]+anova(brand.fit3)[2,2])/
  (anova(brand.fit3)[1,2]+anova(brand.fit3)[2,2]+anova(brand.fit3)[3,2])
```

$R^2_{Y1} = 0.796365$ percent of the error in the total model can be explained by X_1.

$R^2_{Y2} = 0.155694$ percent of the error in the total model can be explained by X_2.

$R^2_{12} = 0.952059$ percent of the error in the total model can be explained by X_1 and X_2.

$R^2_{Y1|2} = 0.9432184$ percent of the error in X_2 can be explained by X_1.

$R^2_{Y2|1} = 0.976862$ percent of the error in X_1 can be explained by X_2.

$R^2 = 0.952059$ percent of the error in the total model can be explained by X_1 and X_2.

Problem 7.16

Part A

The following code transforms the data and standardizes the regression model.

```
n <- dim(brand)[1]
brand.trans <- data.frame(1/sqrt(n-1)*scale(brand))
names(brand.trans) <- c("tYi", "tXi1", "tXi2")
brand.trans.fit <- lm(tYi ~ -1 + tXi1 + tXi2, data = brand.trans)
```

Part B

```
brand.trans.fit$coefficients
```

```
##      tXi1      tXi2
## 0.8923929 0.3945807
```

The transformed coefficient $^*b_1$ means that the transformed X_1 increased 0.8924 per standard deviation.

Part C

Problem 7.24

Part A

```r
brand.simple <- lm(Yi ~ Xi1, data = brand)
brand.simple$coefficients
```

```
## (Intercept)         Xi1
##      50.775       4.425
```

The simple regression model is $Y_i = 50.775 + 4.425X_1$.

Part B

```r
brand.lm.6.5 <- lm(Yi ~ Xi1 + Xi2, data=brand)
brand.lm.6.5$coefficients
```

```
## (Intercept)         Xi1         Xi2
##      37.650       4.425       4.375
```

The coefficients for moisture content are the same between Problem 7.24a and Problem 6.5b.

Part C

```r
anova(brand.simple)[1,2]
```

```
## [1] 1566.45
```

```r
anova(brand.lm.6.5)[1,2]
```

```
## [1] 1566.45
```

$SSR(X_1) = 1566.45$ and $SSR(X_1|X_2) = 1566.45$ are equal.

Part D

Recall the correlation matrix from Problem 6.5a.

```r
cor(brand)
```

```
##            Yi       Xi1       Xi2
## Yi  1.0000000 0.8923929 0.3945807
## Xi1 0.8923929 1.0000000 0.0000000
## Xi2 0.3945807 0.0000000 1.0000000
```

According to the matrix, there is no correlation between X_1 and X_2. Therefore, Part B is justified because X_2 has no influence on X_1. Likewise, Part C is justified because the error of X_2 has no influence on X_1.