Math 423 Assignment 3

Jonathan Pearce 260672004

11/18/2017

```
file1<-"http://www.math.mcgill.ca/yyang/regression/data/cigs.csv"
data<-read.csv(file1,header=TRUE)
x1<-data$TAR
x2<-data$NICOTINE
x3<-data$WEIGHT
y<-data$CO
```

Question a.

```
Anova0123 <- anova(lm(y~x1+x2+x3))
Anova0123
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## x1         1 494.28  494.28 236.4843 6.651e-13 ***
## x2         1   0.97    0.97   0.4661    0.5023
## x3         1   0.00    0.00   0.0011    0.9735
## Residuals 21  43.89    2.09
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
SSRes0123 <- Anova0123[4,2]
#SSRes0123
```

Therefore we have,

$$SS_{Res}(\beta_0, \beta_1, \beta_2, \beta_3) = 43.8925856$$

Question b.

```
Anova012 <- anova(lm(y~x1+x2))
Anova012
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## x1         1 494.28  494.28 247.7322 1.858e-13 ***
## x2         1   0.97    0.97   0.4882     0.492
## Residuals 22  43.89    2.00
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
SSRes012 <- Anova012[3,2]
#SSRes012
```

Therefore we have,

$$SS_{Res}(\beta_0, \beta_1, \beta_2) = 43.8949429$$

Question c.

```
n <- 25
p <- 4
r <- 1
MSRes0123<-SSRes0123/(n-p)
F <- ((SSRes012-SSRes0123)/r)/MSRes0123
#F
```

Therefore the F statistic is,

$$F = 0.0011278$$

Question d.

```
Anova321 = anova(lm(y~x3+x2+x1))
Anova321
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## x3         1 116.06  116.06  55.526 2.522e-07 ***
## x2         1 346.20  346.20 165.636 1.982e-11 ***
## x1         1  33.00   33.00  15.789 0.0006921 ***
## Residuals 21  43.89    2.09
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
SSR321 = c(Anova321[1,2],Anova321[2,2],Anova321[3,2])
SSR321 = c(sum(SSR321),SSR321)
#SSR321
```

Therefore we have,

$$\overline{SS_R}(\beta_1, \beta_2, \beta_3 \mid \beta_0) = 495.2578144$$
$$\overline{SS_R}(\beta_3 \mid \beta_0) = 116.0565104$$
$$\overline{SS_R}(\beta_2 \mid \beta_0, \beta_3) = 346.1998825$$
$$\overline{SS_R}(\beta_1 \mid \beta_0, \beta_3, \beta_2) = 33.0014214$$

It follows,

$$\overline{SS_R}(\beta_3 \mid \beta_0) + \overline{SS_R}(\beta_2 \mid \beta_0, \beta_3) + \overline{SS_R}(\beta_1 \mid \beta_0, \beta_3, \beta_2)$$
$$= 116.0565104 + 346.1998825 + 33.0014214$$
$$= 495.2578144$$
$$= \overline{SS_R}(\beta_1, \beta_2, \beta_3 \mid \beta_0)$$

Question e.

```r
Anova12 = anova(lm(y~x1+x2))
Anova12
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## x1         1 494.28  494.28 247.7322 1.858e-13 ***
## x2         1   0.97    0.97   0.4882     0.492
## Residuals 22  43.89    2.00
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
SSR12 = c(Anova12[1,2],Anova12[2,2])
SSR12 = c(sum(SSR12),SSR12)
#SSR12
```

Therefore we have,

$$\overline{SS_R}(\beta_1, \beta_2 \mid \beta_0) = 495.2554571$$
$$\overline{SS_R}(\beta_1 \mid \beta_0) = 494.2813099$$
$$\overline{SS_R}(\beta_2 \mid \beta_0, \beta_1) = 0.9741472$$

It follows,

$$\begin{aligned}
&\overline{SS_R}(\beta_1 \mid \beta_0) + \overline{SS_R}(\beta_2 \mid \beta_0, \beta_1) \\
&= 494.2813099 + 0.9741472 \\
&= 495.2554571 \\
&= \overline{SS_R}(\beta_1, \beta_2 \mid \beta_0)
\end{aligned}$$

Question f.

```r
Anova01 <- anova(lm(y~x1))
Anova01
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## x1         1 494.28  494.28  253.37 6.552e-14 ***
## Residuals 23  44.87    1.95
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
SSRes01 <- Anova01[2,2]
#SSRes01

n <- 25
p <- 3
r <- 1
MSRes012<-SSRes012/(n-p)
F <- ((SSRes01-SSRes012)/r)/MSRes012
#F
```

Therefore the F statistic is,

$$F = 0.4882394$$

Question g.

```
Anova0 <- anova(lm(y~1))
Anova0
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals 24 539.15  22.465
```

```
SSRes0 <- Anova0[1,2]
#SSRes0

n <- 25
p <- 3
r <- 2
MSRes012<-SSRes012/(n-p)
F <- ((SSRes0-SSRes012)/r)/MSRes012
#F
```

Therefore the F statistic is,

$$F = 124.1101974$$