Statistics-III

Trishan Mondal

Assignment-6

> *For the sake of calculation we will use R. In the following codes, we will upload the data and we will consider the linear model of SF and SJ.

```
 ozone = read.table("ozone.txt")
 ozone.lm_sf = lm(ozone[,3]~ozone[,1]+ozone[,2])
ozone.lm_sj = lm(ozone[,4]~ozone[,1]+ozone[,2])
```

Calculating SF

```
anova(ozone.lm_sf)

## Analysis of Variance Table
##
## Response: ozone[, 3]
##             Df Sum Sq Mean Sq F value    Pr(>F)
## ozone[, 1]   1 9.0569  9.0569  87.119 2.979e-06 ***
## ozone[, 2]   1 1.3112  1.3112  12.612  0.005256 **
## Residuals   10 1.0396  0.1040
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

9.0569+1.3112

## [1] 10.3681

(9.0569+1.3112)/2

## [1] 5.18405

9.0569+1.3112 +1.0396

## [1] 11.4077

5.18405/0.1040

## [1] 49.84663
```

(a) We will compute the **ANoVA for regression**.

Source of variation	d.f	sum of squares	mean squares	F-ratio
regression	$r-1=2$	$SS_{reg}=10.3681$	$MS_{reg}=SS_{reg}/(r-1)=5.18405$	$F_{reg}=49.85$
residual square	$n-r=10$	$SSE=RSS=1.0396$	$MSE=SSE/(n-r)=0.1040$	
Total	12	$SST=11.4077$		

(b) The confidence interval for β_y is $(-0.25165688, -0.13974968)$ and for β_r confidence interval is $(0.01277538, 0.05580002)$.

```
confint(ozone.lm_sf,level = 0.95)

##                    2.5 %       97.5 %
## (Intercept) 277.95501784 498.86914764
## ozone[, 1]   -0.25165688  -0.13974968
## ozone[, 2]    0.01277538   0.05580002
```

It's not hard to see that confidence interval for $\beta_r - \beta_y$ is $(0.17656196, 0.28342)$.

Calculating SJ

```
anova(ozone.lm_sj)

## Analysis of Variance Table
##
## Response: ozone[, 4]
##             Df Sum Sq Mean Sq F value   Pr(>F)
## ozone[, 1]   1 8.9679  8.9679  19.285 0.001353 **
## ozone[, 2]   1 4.7188  4.7188  10.148 0.009727 **
## Residuals   10 4.6502  0.4650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

8.9679+4.7188

## [1] 13.6867

(8.9679+4.7188)/2

## [1] 6.84335

4.6502/10

## [1] 0.46502

 6.84335+13.6867

## [1] 20.53005

  6.84335/ 0.46502

## [1] 14.71625
```

(a) ANOVA table for this model is given by,

Source of variation	d.f	sum of squares	mean squares	F-ratio
regression	$r-1=2$	$SS_{reg}=13.6867$	$MS_{reg}=SS_{reg}/(r-1)=6.84335$	$F_{reg}=14.71$
residual square	$n-r=10$	$SSE=RSS=6.84335$	$MSE=SSE/(n-r)=0.46502$	
Total	12	$SST=20.53005$		

(c) We know, $R^2 = 1 - \frac{RSS}{SST}$. In this case $R^2_{SF} = 0.9088$ and $R^2SJ = 0.74640$. It means that the proportion of variation explained by SF's regressors is greater than SJ's regressors.

(d) We can see that $\hat{\sigma}_{SF} = MSE_{SF} = 0.1040$ and $\hat{\sigma}_{SJ} = MSE_{SJ} = 0.465$. So estimated random error for SF model is less than SJ model.