Lista5

Claudia Meneses

1/1/2020

```
library(MASS)
dataset = read.csv("/Users/claudiameneses/desktop/EA2/Lista2/cableTV.csv", header=T,
                   colClasses = c("numeric", "numeric", "numeric","numeric", "numeric", "numeric","numeric", "num
eric", "numeric"))
```

Resumen de los datos

```
summary(dataset)
```

```
##       obs.          colonia.        manzana.       adultos.    
##  Min.   : 1.00   Min.   :1.000   Min.   : 2.0   Min.   :1.000  
##  1st Qu.:10.75   1st Qu.:1.000   1st Qu.: 7.0   1st Qu.:2.000  
##  Median :20.50   Median :2.000   Median :11.5   Median :2.000  
##  Mean   :20.50   Mean   :1.625   Mean   :13.0   Mean   :2.375  
##  3rd Qu.:30.25   3rd Qu.:2.000   3rd Qu.:20.5   3rd Qu.:3.000  
##  Max.   :40.00   Max.   :2.000   Max.   :25.0   Max.   :4.000  
##      ninos.          teles.         renta.          tvtot.     
##  Min.   :0.000   Min.   :0.00   Min.   : 0.00   Min.   : 0.00  
##  1st Qu.:1.000   1st Qu.:1.75   1st Qu.:50.00   1st Qu.:27.75  
##  Median :1.500   Median :3.00   Median :62.50   Median :39.00  
##  Mean   :1.475   Mean   :2.45   Mean   :58.25   Mean   :43.98  
##  3rd Qu.:2.000   3rd Qu.:3.00   3rd Qu.:70.00   3rd Qu.:63.50  
##  Max.   :3.000   Max.   :5.00   Max.   :85.00   Max.   :86.00  
##     X.valor      
##  Min.   : 79928  
##  1st Qu.:162187  
##  Median :216393  
##  Mean   :227966  
##  3rd Qu.:284262  
##  Max.   :370325  
```

Con base en los datos de la encuesta sobre servivio de televisión por cable se considera el siguiente modelo de regresión lineal múltiple:

$$Renta = \beta_0 + \beta_1 * niños + \beta_2 * adultos + \beta_3 * tvtotal + \beta_4 * valor$$

Regresión y resumen del modelo

```
modAdju <- lm( renta. ~ ninos. + adultos. + tvtot. + X.valor , data = dataset)
summary(modAdju)
```

```
## 
## Call:
## lm(formula = renta. ~ ninos. + adultos. + tvtot. + X.valor, data = dataset)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -22.9975  -7.1922   0.8115   8.5131  24.3167 
## 
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)  9.806e+00  1.033e+01   0.950 0.348838    
## ninos.      -4.914e+00  2.735e+00  -1.797 0.080973 .  
## adultos.     2.640e+00  2.442e+00   1.081 0.287065    
## tvtot.       4.505e-01  1.144e-01   3.936 0.000375 ***
## X.valor      1.299e-04  3.141e-05   4.135 0.000211 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 11.99 on 35 degrees of freedom
## Multiple R-squared:  0.5916, Adjusted R-squared:  0.545 
## F-statistic: 12.68 on 4 and 35 DF,  p-value: 1.772e-06
```

1. Conrucción de la matriz X, calculo de la matriz X’X sus valores propios.

```
X <- model.matrix(modAdju)
XX <- t(X)%*%X
eig <- eigen(XX)
eigen_values <- eig$values
eigen_values
```

```
## [1] 2.311723e+12 3.968089e+04 2.909169e+01 2.531849e+01 1.314027e+00
```

2. Calculo de $\hat{\beta}$ por medio de las ecuaciones normales.

```
Y <- dataset$renta.
beta <- solve(crossprod(X), crossprod(X,Y))
beta
```

```
##                      [,1]
## (Intercept)  9.8056592389
## ninos.      -4.9143208260
## adultos.     2.6400569545
## tvtot.       0.4505260385
## X.valor      0.0001298918
```

3. Estimación de la varianza mediante el cuadrado medio de los residuales. Varias formas.

```
s2 <- deviance(modAdju)
s2
```

```
## [1] 5034.184
```

```
sum(resid(modAdju)^2)
```

```
## [1] 5034.184
```

```
anova(modAdju)
```

```
## Analysis of Variance Table
##
## Response: renta.
##           Df Sum Sq Mean Sq F value    Pr(>F)
## ninos.     1  793.1  793.08  5.5138 0.0246456 *
## adultos.   1 2198.8 2198.82 15.2873 0.0004048 ***
## tvtot.     1 1841.7 1841.66 12.8041 0.0010366 **
## X.valor    1 2459.8 2459.76 17.1014 0.0002106 ***
## Residuals 35 5034.2  143.83
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(modAdju)["Residuals", "Sum Sq"]
```

```
## [1] 5034.184
```

```
with(summary(modAdju), df[2] * sigma^2)
```

```
## [1] 5034.184
```

4. Calculo de la matriz de covarianzas de los coeficientes estimados $cov(\hat{\beta})$.

```
vcov(modAdju)
```

```
##               (Intercept)         ninos.       adultos.         tvtot.
## (Intercept)  1.066331e+02 -1.517955e+01 -7.446102e+00 -8.515453e-02
## ninos.      -1.517955e+01  7.478973e+00  1.469803e+00 -1.781787e-01
## adultos.    -7.446102e+00  1.469803e+00  5.963906e+00 -1.271224e-01
## tvtot.      -8.515453e-02 -1.781787e-01 -1.271224e-01  1.309878e-02
## X.valor     -2.597677e-04  3.725414e-05 -1.445797e-05  3.240173e-07
##                   X.valor
## (Intercept) -2.597677e-04
## ninos.       3.725414e-05
## adultos.    -1.445797e-05
## tvtot.       3.240173e-07
## X.valor      9.865796e-10
```

5. Determine los errores estándar de los coeficientes $s_j = ee(\hat{\beta}_j)$

```
se <- sqrt(diag(vcov(modAdju)))
se
```

```
##  (Intercept)        ninos.      adultos.        tvtot.       X.valor
## 1.032633e+01 2.734771e+00 2.442111e+00 1.144499e-01 3.140986e-05
```

6. Verificar*

```
a <- se^(2)
a <- sum(a)
a
```

```
## [1] 120.0891
```

```
b <- eigen_values^(-1)
b <- sum(b)
b <- b*(sum(resid(modAdju)^2))
b
```

```
## [1] 4203.117
```

8. Intervalo de 90% de confianza para β_3.

```
confint(modAdju, 'ninos.', level=0.90)
```

```
##              5 %       95 %
## ninos. -9.534915 -0.2937269
```

```
confint(modAdju, 'adultos.', level=0.90)
```

```
##                5 %    95 %
## adultos. -1.486067 6.76618
```

```
confint(modAdju, 'tvtot.', level=0.90)
```

```
##              5 %      95 %
## tvtot. 0.2571547 0.6438974
```

```
confint(modAdju, 'X.valor', level=0.90)
```

```
##                  5 %        95 %
## X.valor 7.682254e-05 0.000182961
```

9. Verificar que la renta que esperaría pagar una casa habitación a un nivel promedio de los regresores ($\bar{x} = (1, \bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4)'$) es $\bar{y}$, la renta promedio.

```
y_mean <- mean(Y)
x_prom <- c(1,mean(dataset$ninos.), mean(dataset$adultos.), mean(dataset$tvtot.), mean(dataset$X.valor))
y_est <- sum(beta*x_prom)
print(y_mean == y_est)
```

```
## [1] TRUE
```

10. Respuesta media

```
x_0 <- c(1,1,2,60,300000)
y_0 <- sum(beta*x_0)
c <- ginv(XX)
aux <- sqrt(s2*t(x_0)%*%c%*%x_0)
int_inf <- y_0 - qt(0.05,35)*aux
int_sup <- y_0 + qt(0.05,35)*aux
int_inf
```

```
##          [,1]
## [1,] 100.8056
```

```
int_sup
```

```
##          [,1]
## [1,] 51.53551
```

11. Respuesta nueva estimación

```
x_0 <- c(1,2,3,70,350000)
y_0 <- sum(beta*x_0)
c <- ginv(XX)
aux <- sqrt(s2*(1 + t(x_0)%*%c%*%x_0))
int_inf <- y_0 - qt(0.025,35)*aux
int_sup <- y_0 + qt(0.025,35)*aux
int_inf
```

```
##          [,1]
## [1,] 233.0202
```

```
int_sup
```

```
##           [,1]
## [1,] -63.22791
```