

# rm\_examen2\_solucion.R

Ruben.MontesdeOca

Mon Dec 17 12:49:54 2018

```
#####  
#programador:  
#examen 2, diciembre del 2018  
#####  
  
library(car)  
  
## Loading required package: carData  
library(ggplot2)  
library(reshape2)  
  
#####  
  
#1. Lee las dos bases de datos siguientes:  
#1a. datos hombres:  
  
Input = ("  
Distancia Tiempo  
100      9.9  
200      19.8  
400      43.8  
800      103.7  
1000     136.0  
1500     213.1  
2000     296.2  
3000     457.6  
5000     793.0  
10000    1650.8  
20000    3464.4  
25000    4495.6  
30000    5490.4  
      ")  
hombres <- read.table(textConnection(Input), header=TRUE)  
  
hombres$DistanciaLog <- log(hombres$Distancia) #agregando log(Distancia)  
hombres$TiempoLog <- log(hombres$Tiempo) #agregando log(Tiempo)  
  
#1b. datos mujeres:  
  
Input = ("  
Distancia Tiempo  
60      7.2  
100     10.8  
200     22.1  
400     51.0  
800     117.0  
1500    241.4
```

```

    ")
mujeres <- read.table(textConnection(Input), header=TRUE)

##### pon abajo tu codigo de los datos mujeres con log(Distancia) y log(Tiempo) agregados
##### y ejecuta esta parte

    mujeres$DistanciaLog <- log(mujeres$Distancia) #agregando log(Distancia)
    mujeres$TiempoLog <- log(mujeres$Tiempo) #agregando log(Tiempo)

#2a. Para la base de datos hombres, estima la ordenada al origen y la beta
#del modelo en el pdf (adjunto):

#####Escribe y ejecuta abajo el codigo del modelo
#TiempoLog~ DistanciaLog
#de los datos hombres, no olvides ejecutar summary y anova

    hombres.lm <- lm(TiempoLog~ DistanciaLog, data=hombres)
    summary(hombres.lm)

##
## Call:
## lm(formula = TiempoLog ~ DistanciaLog, data = hombres)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.083987 -0.031813 -0.006206  0.044394  0.060398
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.823196   0.061725  -45.74 6.67e-14 ***
## DistanciaLog   1.112214   0.007785  142.87 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04965 on 11 degrees of freedom
## Multiple R-squared:  0.9995, Adjusted R-squared:  0.9994
## F-statistic: 2.041e+04 on 1 and 11 DF,  p-value: < 2.2e-16

    anova(hombres.lm)

## Analysis of Variance Table
##
## Response: TiempoLog
##              Df Sum Sq Mean Sq F value    Pr(>F)
## DistanciaLog  1  50.304   50.304   20410 < 2.2e-16 ***
## Residuals    11   0.027    0.002
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#2b. Para la base de datos mujeres, estima la ordenada al origen y la beta
#del modelo en el pdf (adjunto):

#####Escribe y ejecuta abajo el codigo del modelo

```

```
#TiempoLog~ DistanciaLog
#de los datos mujeres, no olvides ejecutar summary y anova
```

```
mujeres.lm <- lm(TiempoLog~ DistanciaLog, data=mujeres)
summary(mujeres.lm)
```

```
##
## Call:
## lm(formula = TiempoLog ~ DistanciaLog, data = mujeres)
##
## Residuals:
##      1      2      3      4      5      6
## 0.11466 -0.04774 -0.10227 -0.03657  0.02322  0.04870
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.69216    0.18197  -14.79 0.000122 ***
## DistanciaLog   1.11167    0.03151   35.28 3.85e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08679 on 4 degrees of freedom
## Multiple R-squared:  0.9968, Adjusted R-squared:  0.996
## F-statistic: 1245 on 1 and 4 DF, p-value: 3.852e-06
```

```
anova(mujeres.lm)
```

```
## Analysis of Variance Table
##
## Response: TiempoLog
##              Df Sum Sq Mean Sq F value    Pr(>F)
## DistanciaLog  1  9.3768   9.3768  1244.7 3.852e-06 ***
## Residuals    4  0.0301   0.0075
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#3. Pon dummy variable para sexo, y concatena los datos
#Dummy variable para identificar hombres de mujeres de la siguiente manera:
hombres$Sexo <- 1
mujeres$Sexo <- 0
```

```
#ejecuta esta concatenacion:
data.hm <- as.data.frame(rbind(hombres, mujeres))
```

```
#4. Modelo de regresion
```

```
#####Escribe y ejecuta abajo el codigo del modelo
# TiempoLog~DistanciaLog+ Sexo+ Sexo:DistanciaLog
#de los datos data.hm no olvides ejecutar summary y anova
```

```
hombresmujeres.lm <- lm(TiempoLog~DistanciaLog+ Sexo+ Sexo:DistanciaLog, data=data.hm)
summary(hombresmujeres.lm)
```

```
##
```

```
## Call:
## lm(formula = TiempoLog ~ DistanciaLog + Sexo + Sexo:DistanciaLog,
##     data = data.hm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.102266 -0.038941 -0.006206  0.046545  0.114664
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -2.6921619   0.1295181  -20.786  1.8e-12 ***
## DistanciaLog    1.1116747   0.0224271   49.568 < 2e-16 ***
## Sexo          -0.1310339   0.1505799   -0.870   0.398
## DistanciaLog:Sexo  0.0005397   0.0244299    0.022   0.983
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06178 on 15 degrees of freedom
## Multiple R-squared:  0.9993, Adjusted R-squared:  0.9991
## F-statistic: 6898 on 3 and 15 DF,  p-value: < 2.2e-16
```

```
anova(hombresmujeres.lm)
```

```
## Analysis of Variance Table
##
## Response: TiempoLog
##              Df Sum Sq Mean Sq    F value    Pr(>F)
## DistanciaLog    1  78.927   78.927  20681.3827 < 2.2e-16 ***
## Sexo            1   0.049    0.049   12.8950  0.002675 **
## DistanciaLog:Sexo 1   0.000    0.000    0.0005  0.982666
## Residuals      15   0.057    0.004
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##Estimacion de tiempos si hombres corrieran 60, mujeres corrieran 200
##hay diferentes respuestas porque son diferentes modelos.
##Compara tus resultados de la opcion que elegiste.
##(estos calculos se pueden hacer a mano o siguiendo esta parte del programa):
```

```
##### 4 opcion I: tiempos estimados por los modelos separados:
```

```
#a. hombres corren la distancia 60 usando el modelo de hombres:
TiempoLog= -2.823196+ 1.112214* log(60)
exp(TiempoLog)
```

```
## [1] 5.64399
```

```
#b. mujeres corren la distancia 200 usando el modelo de mujeres:
TiempoLog= -2.69216+ 1.11167* log(200)
exp(TiempoLog)
```

```
## [1] 24.47912
```

```
##### 4 opcion II: tiempos estimados por los modelos juntos usando datos
#####concatenados:
```

```
#a. hombres corren la distancia 60 usando el modelo de hombres y mujeres
```

```

#(coeficientes significantes)
TiempoLog= -2.6921619+ 1.1116747* log(60)
exp(TiempoLog)

```

```
## [1] 6.419995
```

```

#b. mujeres corren la distancia 200 usando el modelo de hombres y mujeres
#(coeficientes significantes):
TiempoLog= -2.6921619+ 1.1116747* log(200)
exp(TiempoLog)

```

```
## [1] 24.47968
```

```

##### 4 opcion III: tiempos estimados por los modelos juntos usando
#####datos concatenados (hombres es Sexo=1):

#a. hombres corren la distancia 60 usando el modelo de hombres y mujeres
#(coeficientes significantes o no)
TiempoLog= -2.6921619+ 1.1116747* log(60)- 0.1310339*1+ 0.0005397*1
exp(TiempoLog)

```

```
## [1] 5.634583
```

```

#b. mujeres corren la distancia 200 usando el modelo de hombres y mujeres
#(coeficientes significantes o no):
TiempoLog= -2.6921619+ 1.1116747* log(200)- 0.1310339*0+ 0.0005397*0
exp(TiempoLog)

```

```
## [1] 24.47968
```

```

#####
#end of program
#####

```