Andrés Felipe Paredes

John Sebastián Marín

# Taller en clase: Variables Dummy

El modelo inicial del problema que proponemos es:

Modelo . Salario con base en Enseñanza, Experiencia y Años en la compañía

Donde ‘S’ es salario, ‘E’ es Educación, ‘X’ es Experiencia y ‘A’ es Años en la compañía. Si evaluamos este modelo podemos observar que ninguna de las variables explicativas es significativa.

El siguiente paso es tener en cuenta el género. Para revisar si esta variable cualitativa tiene algún efecto en el salario, proponemos el siguiente modelo:

Modelo . Salario con base en Enseñanza, Experiencia, Años en la compañía y Género

Donde la variable dummy D es 1 si el empleado es mujer y 0 si es hombre.

Evaluando el modelo 2 podemos ver que:

1. **La educación y los años en la compañía cobran significancia.**
2. Al comparar el modelo 1 con el modelo 2, mediante ANOVA podemos rechazar con 90% de confiabilidad que el modelo 1 sea mejor que el modelo 2. Por lo tanto, el modelo 2 es mejor que el modelo 1 y el género sí influye de alguna manera.

El resultado respectivo del modelo 2 sería:

Con estos datos podemos concluir que:

1. La parte fija del salario es mayor para las mujeres.
2. **La educación y los años en la compañía aumentan más el salario de los hombres que el de las mujeres.**
3. La experiencia aumenta más el salario de las mujeres que el de los hombres.

Teniendo en cuenta los puntos anteriores. Si la enseñanza y los años en la compañía son variables significativas y los hombres ganan más con base en estas 2 variables, podemos concluir que en la empresa sí hay discriminación de género.

A continuación se presentan los datos técnicos:

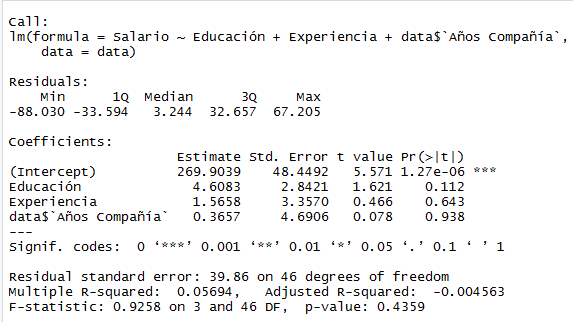


Tabla . Datos técnicos del Modelo 1

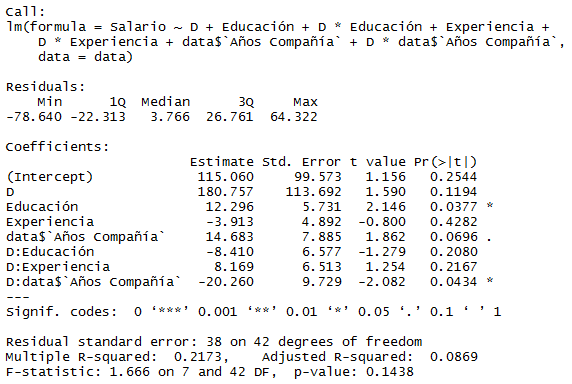


Tabla . Datos técnicos del Modelo 2

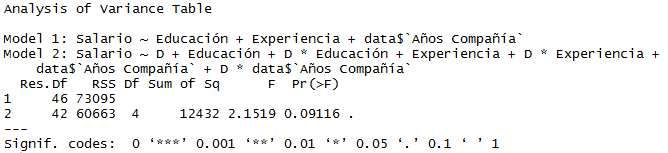


Tabla . Análisis ANOVA entre el Modelo 1 y el Modelo 2

## Código R

library(readxl)

data <- read\_excel(file.choose())

head(data)

str(data)

attach(data)

D <- Sexo=='F'

head(D)

View(data)

data$D <- as.numeric(D)

head(data)

plot(Salario,D)

res1 <- lm(Salario ~ Educación+Experiencia+data$`Años Compañía`,data=data)

summary(res1)

res2 <- lm(Salario ~ D+Educación+D\*Educación+Experiencia+D\*Experiencia+data$`Años Compañía`+

D\*data$`Años Compañía`,data=data)

summary(res2)

anova(res1,res2)

##Podemos rechazar con 90% de confianza que el 1 sea mejor que el 2. Es decir, el género

##de alguna manera sí afecta al Salario.