

tarea

Diego Dell'Era

La **tarea** era: ver qué onda la variable *sexo2*.

Cargamos el dataset:

```
library(foreign)
dataset <- read.spss('GSS_EEA2010_minimod.sav', to.data.frame=TRUE)
```

Antes de mandarnos a hacer algo, comprobamos que las variables de sexo sean categóricas:

```
is.factor(dataset$sexo1) && is.factor(dataset$sexo2)
```

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

```
levels(dataset$sexo1)
```

```
## [1] "Mujer" "Hombre"
```

```
levels(dataset$sexo2)
```

```
## [1] "Hombre" "Mujer"
```

Miramos cómo aparecen estos factores en el dataset:

```
str(dataset$sexo1)
```

```
## Factor w/ 2 levels "Mujer","Hombre": 2 1 1 1 1 1 1 1 1 1 ...
```

```
str(dataset$sexo2)
```

```
## Factor w/ 2 levels "Hombre","Mujer": 1 2 2 2 2 2 2 2 2 2 ...
```

¿Son lo mismo? Comparamos todos los registros:

```
all(dataset$sexo1 == dataset$sexo2)
```

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

Si no entendí mal, son lo mismo, pero con etiquetas diferentes nomás.

Hacemos el modelo de regresión con la variable *sexo1* como una de las dependientes:

```
fit_sexo1 <- lm(horastv ~ educ + sexo1, data=dataset)
coefficients(fit_sexo1)
```

```
## (Intercept)          educ sexo1Hombre
## 12.9225562   -0.6838044   -0.4047501
```

Ahora sacamos *sexo1* y ponemos *sexo2*:

```
fit_sexo2 <- lm(horastv ~ educ + sexo2, data=dataset)
coefficients(fit_sexo2)
```

```
## (Intercept)          educ  sexo2Mujer
## 12.5178061   -0.6838044    0.4047501
```

El coeficiente sobre la variable regresora en cuestión cambió de signo:

```
coefficients(fit_sexo1)[3]
```

```
## sexo1Hombre
## -0.4047501
```

```
coefficients(fit_sexo2)[3]
```

```
## sexo2Mujer
## 0.4047501
```

Recodificamos *sexo2* con 1 y 2:

```
library(plyr)
dataset$sexo2_recodificado <- revalue(dataset$sexo2, c("Hombre"="1", "Mujer"="2"))
fit_sexo2_recodificado <- lm(horastv ~ educ + sexo2_recodificado, data=dataset)
coefficients(fit_sexo2_recodificado)
```

```
##          (Intercept)          educ sexo2_recodificado2
##      12.5178061      -0.6838044          0.4047501
```

Recodificamos al revés:

```
dataset$sexo2_recodificado <- revalue(dataset$sexo2, c("Hombre"="2", "Mujer"="1"))
fit_sexo2_recodificado <- lm(horastv ~ educ + sexo2_recodificado, data=dataset)
coefficients(fit_sexo2_recodificado)
```

```
##          (Intercept)          educ sexo2_recodificado1
##      12.5178061      -0.6838044          0.4047501
```

Los coeficientes se mantienen iguales a los que se obtenían para *sexo2* originalmente.