

Ejemplo de variables ficticias, Wooldridge

Tenemos un fichero con las siguientes variables: *salario* (la variable explicada); *educacion*: años de educación; *experiencia*: años de experiencia; *tenure*: años en este mismo trabajo; *nonwhite*: =1 si no es de raza blanca; *mujer*: =1 si es mujer; y *casado*: =1 si está casado.

Tomamos el modelo simple:

$$\text{salario} = \beta_0 + \beta_1 \text{mujer} + \beta_2 \text{educacion} + \epsilon$$

Queremos ver si hay una diferencia en el salario entre mujeres y hombre de la misma educación. Por lo tanto si $\beta_1 < 0$ esto implicaría que las mujeres están en desventaja porque la variable *mujer*=1 para mujer y 0 para hombre. Por lo tanto si aceptamos que $E(\epsilon | \text{mujer}, \text{educacion}) = 0$ y además que el promedio de educación es el mismo para los dos sexos, entonces por definición:

Tomemos un modelo más realista:

$$\text{salario} = \beta_0 + \beta_1 \text{mujer} + \beta_2 \text{educacion} + \beta_3 \text{experiencia} + \beta_4 \text{tenure} + \epsilon$$

```
> data<-read.table("binario.txt", header=T)
> attach(data)
> binario.lm<-lm(salario~1+ mujer+educacion+experiencia+tenure)
> summary(binario.lm)
```

Call:

```
lm(formula = salario ~ 1 + mujer + educacion + experiencia +
    tenure)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.7675	-1.8080	-0.4229	1.0467	14.0075

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.56794	0.72455	-2.164	0.0309	*
mujer	-1.81085	0.26483	-6.838	2.26e-11	***
educacion	0.57150	0.04934	11.584	< 2e-16	***
experiencia	0.02540	0.01157	2.195	0.0286	*
tenure	0.14101	0.02116	6.663	6.83e-11	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.958 on 521 degrees of freedom

Multiple R-Squared: 0.3635, Adjusted R-squared: 0.3587

F-statistic: 74.4 on 4 and 521 DF, p-value: < 2.2e-16

Si quisieramos ver cual es la diferencia de salarios en promedio entre los dos sexos, nos desprenderíamos de las otras variables:

$$salario = \beta_0 + \beta_1mujer + \epsilon$$

de donde obtenemos

$$\widehat{salario} = 7.10 - 2.51mujer \quad R^2 = 0.116$$

(0.21) (0.30)