

SMM637 Exercises - Regression with Dummy Variables

1. The data for this exercise are contained in `wage1` from the R package `wooldridge`. The full description of the variables can be found here:

<https://cran.r-project.org/web/packages/wooldridge/wooldridge.pdf>.

- (a) Using R install the package and load the data.
- (b) Fit the following linear model:

$$\text{wage} = \alpha + \beta_1 \text{female} + \beta_2 \text{educ} + \beta_3 \text{exper} + \beta_4 \text{tenure} + \epsilon$$

- (c) Interpret the estimated coefficient $\hat{\alpha}$.
- (d) Comment of significance and meaning of the coefficient associated with `female`.
- (e) Now, fit the following linear model:

$$\text{wage} = \alpha + \beta_1 \text{female} + \epsilon$$

- (f) Comment on the estimated coefficients of α and β_1 .
- (g) Compare the estimate for β_1 obtained when fitting model in (b) and model in (e).
- (h) Carry out a comparison of means test between the two groups, which in this case are men and women. Comment on the results.
- (i) Modify the model fitted in (b) such that the new model allows the effect of `educ` to vary by gender. Is this model supported by the data?