Lab Assignment 2

May 7, 2024

The Boston Housing Data Set: We will work with `Boston`, the Boston Housing data set, which contains 506 observations on housing values in the suburbs of Boston. Boston data set comes with the package MASS. The packages MASS and AER are required for the interactive R exercises below.

Instructions:

1. Load both the package `MASS` and `AER` and the data set `Boston` from `MASS`.

2. Get yourself an overview of the data.

3. Estimate a simple linear regression model that explains the median house value of districts (`medv`) by the percent of households with low socioeconomic status, `lstat`, and a constant. Save the model to `bh-mod`.

4. Print a coefficient summary to the console that reports standard errors.

Now, let us expand the approach from the previous exercise by adding additional regressors to the model and estimating it again.

5. Regress the median housing value in a district, `medv`, on the average age of the buildings, age, the per-capita crime rate, `crim`, the percentage of individuals with low socioeconomic status, `lstat`, and a constant. Estimate the model.

6. Print a coefficient summary to the console that reports standard errors for the augmented model.

7. The R^2 of the simple regression model is stored in `R2-res`. Save the multiple regression model's R^2 to `R2-unres` and check whether the augmented model yields a higher R^2.

We can conduct hypothesis tests about the coefficients in multiple regression models. The most common hypothesis is $\mathrm{H}_0 : \beta_j = 0 \text{vs.} H_1 : \beta_j \neq 0$

8. Compute t-statistics for each coefficient. Assign them to `tstat`
9. Compute p-values for each coefficient and assign them to `pval`.
10. Check whether the hypotheses are rejected at the 5% significance level.