Bootstrap: further reading

Brief exposition:

James G, Witten D, Hastie T, Tibshirani R (2021). An Introduction to Statistical Learning: With Applications in R, Second edition. Springer, New York. Chapter 5.2.

Definitive references:

Davison AC, Hinkley DV (1997). Bootstrap Methods and Their Application. Cambridge University Press, Cambridge ; New York, NY, USA.

Efron B, Tibshirani R (1993). An Introduction to the Bootstrap. Chapman & Hall, New York.

Bootstrapped p-value

- Learning goals
  - Understand p-values by understanding their underlying sampling algorithm
  - Further understand how the sampling distribution is the basis for frequentist inference
  - Understand how bootstrap algorithms mimic the sampling algorithm
  - Formulate a bootstrap algorithm and translate it to R

Bootstrapped p-value

- Parametric bootstrap for H_0: $\beta_1 = 0$
  - what is the definition of p-value?
  - what is the algorithm for parametric bootstrap?
  - combine these concepts
- Pseudocode first!
- R code from pseudocode

Definition of a p-value

The probability of a sample statistic as large or larger than the one observed given that some hypothesis is true

Basic parametric bootstrap algorithm

repeat very many times
    sample from the error distribution
    create new y-values from the estimated parameters and errors
    fit the linear model
    estimate the parameters
plot sampling distribution (histogram) of the parameter estimates

plug in: create simulated data from model