Today

- Recap & questions from homework
- Results for bootstrapped confidence intervals
  - sharing, questions
- Coding a bootstrapped p-value algorithm (inference)

Homework check in

- Quantity
  - 33% too much
  - 66% just right
- Difficulty
  - 25% too hard
  - 75% just right or too easy
- Bottleneck: learning to program

Git skills

- Branching & merging
- Amending
- .gitignore
- Using a GUI

Coding style

- Working together
- Agreeing on a style

Sampling distribution algorithm

repeat very many times
    sample data from the population
    fit the model
    estimate the parameters
plot sampling distribution (histogram) of the parameter estimates

Bootstrap algorithm

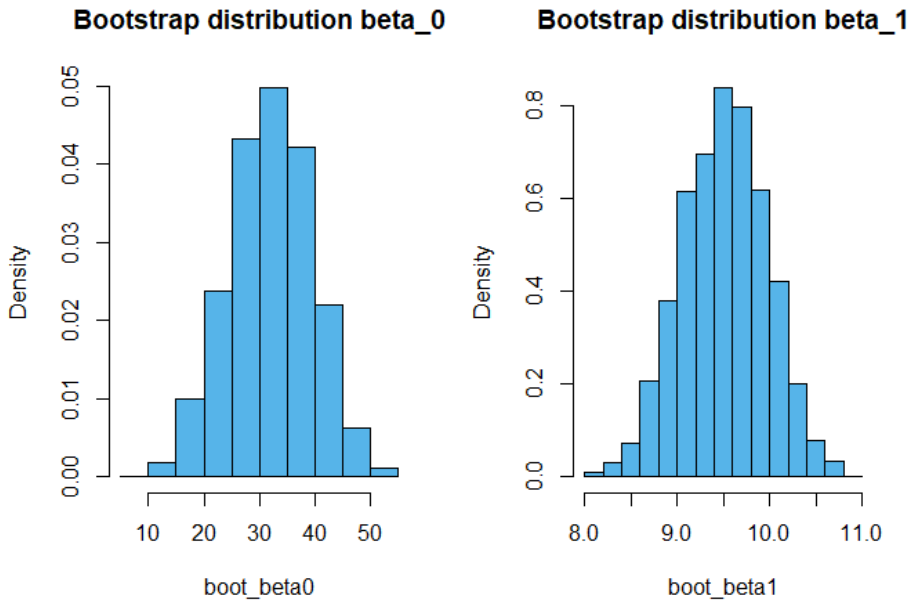
repeat very many times
    generate data based on the sample
    fit the model
    estimate the parameters
plot sampling distribution (histogram) of the parameter estimates

Bootstrap algorithms

- Non-parametric bootstrap
  - resample the data
- Empirical bootstrap
  - resample the residuals
- Parametric bootstrap
  - generate data from a distribution
  - use estimated parameters of the distribution

Code (e.g. empirical bootstrap)

```
for ( i in 1:10000 ) {
    e_boot <- sample(e_fit,replace=TRUE)
    df_boot$y <- coef(fit)[1] + coef(fit)[2]*df_boot$x + e_boot
    fit_boot <- lm(y ~ x, data=df_boot)
    boot_beta0[i] <- coef(fit_boot)[1]
    boot_beta1[i] <- coef(fit_boot)[2]
}
```



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