

Resampling Methods

Exercise 1 (Conceptional)

Exercise 2(p. 197): Derive the probability that a given sample is part of a bootstrap sample.

Exercise 2 (Applied: CV)

Exercise 8 (p. 200): Perform cross-validation on a simulated data set.

Exercise 3 (Applied: Bootstrap)

Exercise 9 (p. 201): Perform bootstrap on `Boston` housing data.

Exercise 4 (To be graded in detail)

Let be X a random variable with unknown distribution function F . We are able to sample seven observations:

$$\mathbf{x} = (4.63, 2.61, 4.95, 1.95, 3.31, 4.85, 4.57, 4.37, 1.98, 2.97).$$

We are interested in the mean θ of the distribution F and the corresponding standard error. Since the sample size is very small and the distribution F is unknown, we generate the bootstrap samples $\mathbf{x}^{*1}, \dots, \mathbf{x}^{*10}$:

b	bootstrap sample \mathbf{x}^{*b}							mean $\hat{\theta}^*(b)$	$(\hat{\theta}^*(b))^2$
1	4.85	2.61	1.95	4.95	2.61	1.95	4.63	3.36	11.32
2	3.31	4.95	2.61	3.31	4.63	4.95	4.63	4.06	16.45
3	1.95	2.61	3.31	1.95	2.61	4.85	1.95	2.75	7.55
4	1.95	3.31	2.61	4.85	1.95	1.95	4.95	3.08	9.50
5	1.95	4.95	4.57	4.95	4.57	4.63	3.31	4.13	17.08
6	4.63	4.85	4.85	3.31	4.95	3.31	1.95	3.98	15.83
7	4.95	4.95	4.57	1.95	4.57	4.85	1.95	3.97	15.76
8	4.95	1.95	1.95	2.61	4.95	2.61	1.95	3.00	8.97
9	4.57	4.85	1.95	3.31	4.57	4.85	2.61	3.82	14.56
10	1.95	2.61	1.95	4.85	4.95	4.63	4.57	3.64	13.28
sum								35.79	130.29

- (a) Calculate the bootstrap mean estimate $\hat{\theta}^*(\cdot)$ and compare it to the sample mean.
- (b) Calculate the bootstrap estimate for the standard error $\hat{se}(\hat{\theta}(\mathbf{x}))$ of the mean of X

This homework is due at the beginning of the discussion section on Feb 9, 2016 at 3.30pm.

All references refer to the textbook: James, Witten, Hastie and Tibshirani (2013). *An Introduction to Statistical Learning with Applications in R*. Springer. Available online: <http://www-bcf.usc.edu/~gareth/ISL/>.
