Engineering Maths EMAT 30007, Bootstrapping

Ksenia Shalonova

1 Bootstrap confidence intervals:

MATLAB function ci = bootci(nboot,bootfun,...) works out bootstrap confidence interval. The default value of alpha is 0.05 (95% confidence interval). The first argument `nboot` is the number of bootstrap replications used in the computation. The second argument `bootfun` is the parameter you estimate. The third argument is input data. The last argument defines the type of the bootstrap confidence interval. There are three major ones described in the lecture:

(1) basic percentile bootstrap CI - `'percentile'`
(2) bootstrap-t or studentized CI - `'student'`
(3) CI with bootstrapped standard error - `'normal'`

Notice that the default bootstrap method is the `'accelerated percentile method'`.

(a) The random sample of 10 elements in the code below represents the auditory nerve fiber responses rates in cats. We want to estimate the population mean response rate.

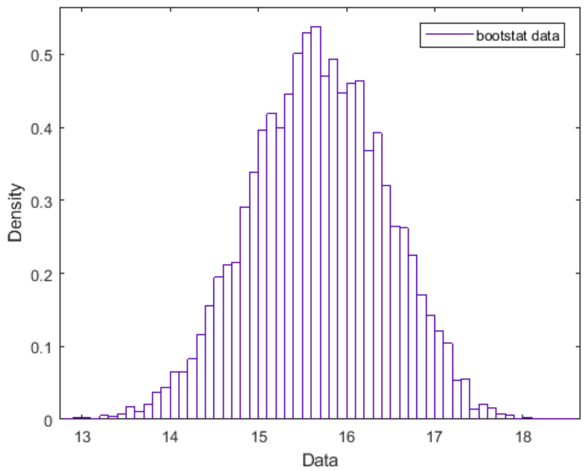
```
>> x=[15.1 14.6 12.0 19.2 16.1 15.5 11.3 18.7 17.1 17.2] % input vector
>> parameter=@(y)mean(y) % we estimate mean for each bootstrap sample
>> ci=bootci(3000,{parameter, x}, 'alpha', 0.05,'type', 'percentile')

OR

>> [ci,bootstat] =bootci(3000,{parameter, x}, 'alpha', 0.05,'type', 'percentile')
% returns the bootstrapped statistic computed for each of the bootstrap sample
```

Work out:

- Run the function `bootci` several times. Do you get a different confidence interval each time? Explain.
- Plot the bootstrap distribution on the graph. Does the distribution look normal? What is the bootstrap standard error?



- Using `bootstat` output (i.e. bootstrap distribution), work out bootstrap percentile CI 'by hand'. You can use MATLAB function `prctile`.
- Apply bootstrap-t CI method by using `bootci` function. Does it give similar result to the percentile method?
- Applying CI with bootstrapped standard error method. Compare the result with two previous CI bootstrap methods.

(b) **Compare the performance of three bootstrap methods for any skewed distribution.** You can perform the following steps:

- Apply Monte Carlo simulations for generating random samples in order to get a sample distribution. You can work either with lognormal or exponential distribution using Matlab `***rnd` function.
- Apply three CI bootstrap methods by taking one random sample from the distribution as an input.
- Compare the results of three bootstrap methods with the original sample distribution.
- Does the number of replications affect the result?

(c) Estimate the age of glacial drifts:
Download file 'TILLRATIO.xls' from Blackboard that provides data on the ratios of aluminum (Al) to beryllium (Be) chemical makeup of buried glacial drifts in Wisconsin. Use the bootstrap procedure to estimate the true proportion of Al to Be using a 95% confidence interval.

2 Bootstrap hypothesis testing: *OPTIONAL, but important for understanding the concept of a p-value*

Use the method described in the recommended book for the unit (written by W.Mendenhall et al.), section 'Alternative Testing Procedures: Bootstrapping'.

Download file 'ROUGHPIPE.xls' from Blackboard. The file presents data for 20 sampled pipe sections (in micrometers). Use the bootstrap procedure to test whether mean surface roughness differs from 2 micrometers. Make significance level $\alpha = 0.05$.