STAT 102B: Homework 2

George Michailidis

Due electronically through **Gradescope**
on April 24 at 11:59 pm

Problem 1: Recall Problem 1 in Homework 1.

Consider the correlation coefficient r between two random variables X and Y. Recall that it is a measure of *linear association* between X and Y and takes values in the interval $[-1, 1]$; i.e., $r(X, Y) \in [-1, 1]$.

(a) Use the code from Homework 1 to simulate data from a *bivariate normal distribution* with mean vector $\mu = [0\ 0]$ and correlation matrix

$$R = \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$$

for the following cases:

1. Sample size $n \in \{50, 200\}$ and correlation coefficient $r = 0$
2. Sample size $n \in \{50, 200\}$ and correlation coefficient $r = 0.5$
3. Sample size $n \in \{50, 200\}$ and correlation coefficient $r = 0.85$

(b) Obtain the following Bootstrap Confidence Intervals for the correlation coefficient r for the three cases above.

1. Normal Bootstrap CI

2. Basic Bootstrap CI

3. Percentile Bootstrap CI

4. Bootstrap-t (Studentized) Bootstrap CI

Calculate the length and the shape of each type of Bootstrap CI and report them as well.

Discuss how you selected the number of bootstrap replicates B.

Hint: The results for the population median in Lecture2-2 and Lecture3-1 provide good guidance on what is an appropriate B.

Comment on the results; in particular how the various bootstrap CI behave as a function of the sample size n, and the value of the correlation coefficient r.

Problem 2:

The data set "cats" can be obtain through the following `R` code:

```
> library(MASS)
> data(cats)
> summary(cats)
 Sex         Bwt             Hwt
 F:47   Min.   :2.000   Min.   : 6.30
 M:97   1st Qu.:2.300   1st Qu.: 8.95
        Median :2.700   Median :10.10
        Mean   :2.724   Mean   :10.63
        3rd Qu.:3.025   3rd Qu.:12.12
        Max.   :3.900   Max.   :20.50
```

It contains the body weight (Bwt) in kilograms and the (Hwt) in grams of 47 female and 97 male cats.

Part (a):

Construct the following bootstrap CI for the difference of the body weight means between female and male cats.

1. Normal Bootstrap CI
2. Basic Bootstrap CI
3. Percentile Bootstrap CI
4. Bootstrap-t (Studentized) Bootstrap CI

Calculate the length and the shape of each type of Bootstrap CI and report them as well.

Discuss how you selected the number of bootstrap replicates B and comment on the results.

Part (b):

Using your code from Problem 1 above to:
construct the following bootstrap CI for the correlation coefficient between the body weight and the heart weight of female cats.

1. Normal Bootstrap CI
2. Basic Bootstrap CI
3. Percentile Bootstrap CI
4. Bootstrap-t (Studentized) Bootstrap CI

Calculate their length and their shape and report those as well for each type of bootstrap CI.

Discuss how you selected the number of bootstrap replicates B and comment on the results.

Part (c):

Using your code from Problem 1 above to:
construct the following bootstrap CI for the correlation coefficient between the body weight and the heart weight of male cats.

1. Normal Bootstrap CI
2. Basic Bootstrap CI
3. Percentile Bootstrap CI
4. Bootstrap-t (Studentized) Bootstrap CI

Calculate their length and their shape and report those as well for each type of bootstrap CI.

Discuss how you selected the number of bootstrap replicates B and comment on the results.