Module 4 Poll Q and A

Stat 406, W1 2020

Last modified - 2020-11-14

1. Which of the following are TRUE about bootstrap distributions and sampling distributions? (Multiple Choice)

   a. Bootstrap distributions are centered at the sample statistic while sampling distributions are centered at the population parameter. **TRUE**
   b. Both distributions are created by sampling with replacement from the population. **FALSE**, the bootstrap distribution is created by sampling with replacement from the sample.
   c. Both distributions get narrower as the standard deviation decreases. **TRUE**
   d. Both distributions depend on sample statistics. **TRUE**

2. Non-parametric bootstrap intervals are likely _____ because they _____ about the population (Single Choice)

   a. narrower; ignore information
   b. narrower; use more information
   c. wider; use more information
   d. wider; ignore information **CORRECT**

3. Before we bag trees, each tree, since it is unpruned, will have _____ variance and _____ bias. Therefore averaging many trees results in an estimator that has _____ variance.

   a. low; high; lower
   b. high; low; higher
   c. high; low; lower **CORRECT**
   d. low; high; higher

4. **TRUE**/FALSE. The difference between Random Forests and Bagged Trees is that Random Forests considers only a subset of predictors at each possible split.

5. TRUE/**FALSE**. Neural networks are called "universal approximators" because they can potentially estimate any continuous or discontinuous regression function arbitrarily well. **False**, Nnets can approximate piecewise continuous functions up to a certain precision as long as they are wide enough. They cannot approximate all functions arbitrarily well.

6. TRUE/**FALSE**. When training a neural network, it is recommended to input your data "as is", without making any transformations. As long as you have enough layers, the network can figure out any useful transformations. **False**. You should apply transformations for at least 2 reasons: (1) if you suspect that they will be beneficial (or know they will be) and (2) to ensure that the data is of commensurate size (centered and scaled to have unit norm for example). The second case is important because of the scale of the activation functions as well as helping with random starting values and learning rate choices.