Homework 5

PSTAT 5A: Spring 2023, with Ethan P. Marzban

Instructions

- Please submit your work to Gradescope by no later than **11:59pm on Wednesay, May 10**. As a reminder, late homework will not be accepted.
- Recall that you will be asked to upload a **single** PDF containing your work for *both* the programming and non-programming questions to Gradescope.
  - You can merge PDF files using either Adobe Acrobat, or using adobe's online PDF merger at this link.

Problem 1: Browsing Habits

Suppose that the time a randomly-selected person spends on their phone in a day is found to be a random variable that follows the normal distribution with mean 195 minutes and standard deviation 50 minutes.

a. Define the random variable of interest, and call it X.

b. What is the probability that a randomly-selected person spends less than one hour per day on their phone?

c. What proportion of the population spends greater than 173 minutes per day on their phone?

d. What proportion of the population spends between 170 minutes and 200 minutes on their phone in a given day?

Problem 2: Trees!

The heights of trees in a particular forest are found to vary uniformly between 10ft and 100ft. A park ranger is interested in the heights of randomly-selected trees.

a. Define the random variable of interest, and call it X.

b. What proportion of trees have heights between 25 and 87 feet?

c. What proportion of trees have heights between 25 and 150 feet?

d. What proportion of trees have heights larger than 63 feet?

e. Suppose the ranger collects a sample of 50 different tree heights (assume the sample was taken *with* replacement). What is the probability that exactly 22 of these trees had heights larger than 63 feet? **Hint:** You may need to define another random variable, and, consequently, use *another* distribution!

Problem 3: Parameter or Statistic?

In each of the parts below, determine whether the provided quantity is a population parameter or a sample statistic. Use this to further determine whether the quantity is a deterministic (i.e. non-random) constant, or a random variable.

a. The median score of 80 students, sampled from a class of 100.

b. The maximum amount of time (in minutes) any human can hold their breath under water.

c. The true IQR of incomes in Brazil.

d. The standard deviation of the times it took 40 randomly-selected runners to complete a marathon.

Problem 4: Reducing Blood Pressure

A new drug is advertised to significantly reduce systolic blood pressure. To test these claims, a clinician takes a representative sample of 120 volunteers to whom she administers the drug. She records the difference in (systolic) blood pressure pre- and post- administration of the drug for each of the 120 volunteers, and finds that the volunteers had an average difference of -8 mm Hg (millimeters of mercury).

a. Identify the population of interest.

b. Identify the sample.

c. Is the mean difference of -8 mm HG a population parameter or an observed instance of a sample statistic?

Problem 5: College Degrees

Suppose that 33% of a particular country's population has a college degree. A representative sample of 243 people is taken, and the proportion of these people who have a college degree is recorded.

a. Define the parameter, and use the notation discussed in Lecture 10.

b. Define the random variable of interest, and use the notation discussed in Lecture 10.

c. Check whether the success-failure conditions are satisfied.

d. What is the probability that over 30% of the sample have college degrees?

e. What is the probability that the proportion of people in the sample who have college degrees lies within 5% of the true proportion of 33%.

There is no programming part on this homework.