Probability and Statistics – Problem Set 10

April 28, 2022
May 5, 2022 in class

Problem 1

Two children had a lemonade stand on a hot day last August, and recorded the following dollar amounts they received for each one-hour slot they worked.

$$3, 7, 8, 5, 12, 14, 21, 15, 35, 18, 14$$

Draw the box-and-whisker plot for this data set.

Problem 2

1. Let T_1 be an unbiased estimator for a parameter of interest θ, and W be a random variable with zero mean. Prove that $T_2 = T_1 + W$ is also an unbiased estimator for θ.

2. Let T_1 be an estimator for a parameter of interest θ such that $E[T_1] = a\theta + b$, where a and b are real numbers with $a \neq 0$. Is
$$T_2 = \frac{T_1 - b}{a}$$
an unbiased estimator for θ? Please explain.

Problem 3

Let X_1, X_2, X_3, …, X_n be a random sample from a Uniform$(0, \theta)$ distribution, where θ is the unknown parameter of interest. Define the estimator

$$T = \max(X_1, X_2, \dots, X_n)$$

Compute the bias of $B(T)$ of T.

Problem 4

When pollsters poll regarding sensitive or embarassing questions, they run into the difficulty that people polled may lie with their answer, which can pollute the quality of the poll. To address this difficulty, a pollster may use the following trick.

She asks every person being polled to roll a dice as he/she is being asked the question. If the result of the dice is 6, the person must respond without lying. If not, the person must say the opposite of what he/she believes. The key is that the pollster does not see the result of the dice, so the person may accept to participate in the poll and not be tempted to lie.

The pollster assumes that for the question being asked, the opinion of each person (before being polled) follows a Bernoulli distribution with parameter μ, μ being the probability of "yes", and $1 - \mu$ being the probability of "no". The pollster would like to estimate μ as accurately as possible.

1. Let X_i be the random variable corresponding to the answer a person polled gives to the pollster, after the dice is rolled. What is the probability mass function of X_i?

2. Consider the following estimator for μ:

$$T_n = \frac{\overline{X}_n + \frac{1}{6} - 1}{2 \cdot \frac{1}{6} - 1} = \frac{5}{4} - \frac{3}{2}\overline{X}_n$$

   where $\overline{X}_n$ is the sample mean. Is T_n a biased estimator? If yes, is it positively or negatively biased?

3. What is the variance of the estimator T_n?

Problem 5

Consider a continuous random variable whose probability density function is

$$f_\theta(x) = \frac{2}{\sqrt{\pi}\theta^{3/2}} x^2 e^{-x^2/\theta} \qquad , \qquad \theta > 0$$

for all $x \in \mathbb{R}$, where θ is the unknown parameter of interest.

1. n independent observations $(a_1, \dots, a_n)$ are taken from this distribution. What is the maximum likelihood estimate of θ?

2. Consider the maximum likelihood estimator $\hat{\theta}$ corresponding to the formula you derived to estimate θ. What is the bias of $\hat{\theta}$?

3. What is the variance of $\hat{\theta}$? What happens to the variance of $\hat{\theta}$ as one increases the number n of independent observations?

Remember to justify your answers!