MA2040: Probability, Statistics and Stochastic Processes
Problem Set-VI

Sivaram Ambikasaran

April 3, 2019

1. Let X be the height of a randomly chosen individual from a population. In order to estimate the mean and variance of X, we observe a random sample $X_1, X_2, \ldots, X_7$. Thus, X_i's are i.i.d. and have the same distribution as X. We obtain the following values (in centimeters): $166.8, 171.4, 169.1, 178.5, 168.0, 157.9, 170.1$ Find the values of the sample mean, the sample variance, and the sample standard deviation for the observed sample.

2. Prove the following:
   - If $\hat{\Theta}$ is an unbiased estimator for Θ, then so is $\hat{\Theta} + W$, where W is a zero mean random variable.
   - If $\hat{\Theta}_1$ is an estimator for Θ such that $\mathbb{E}\left(\hat{\Theta}_1\right) = a\Theta + b$, where $a \neq 0$, show that
   $$\hat{\Theta}_2 = \frac{\Theta_1 - b}{a}$$
   is an unbiased estimator for Θ.

3. Let $\{X_i\}_{i=1}^n$ be a random sample from a uniform distribution $(0, \Theta)$, where Θ is unknown. Define the estimator
$$\hat{\Theta} = \max\{X_1, X_2, \ldots, X_n\}$$
   - Find the bias of $\hat{\Theta}$
   - Find the $\mathbb{E}\left(\left(\hat{\Theta} - \Theta\right)^2\right)$
   - Is $\hat{\Theta}$ a consistent estimator of Θ?

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a geometric distribution with parameter p, where p is unknown. Find the maximum likelihood estimator of p based on the random sample $\{X_i\}_{i=1}^n$.

5. Let $X_1, X_2, \ldots, X_n$ be a random sample from a uniform distribution $(0, \Theta)$, where Θ is unknown. Find the maximum likelihood estimator of Θ based on the random sample $\{X_i\}_{i=1}^n$.

6. Let X_1, X_2, X_3, X_4 be a random sample from a normal distribution with mean μ and $\sigma = 3$. If s^2 denotes the sample variance, and $\mathbb{P}\left(s^2 \leq k\right) = 0.05$, find k.

7. If $X \sim \mathcal{N}\left(\mu = 10, \sigma^2 = 25\right)$ and $X_1, X_2, \ldots, X_{501}$ is a random sample from X, find $\mathbb{E}\left(s^2\right)$.

8. If $T \sim t_{19}$, find C so that $\mathbb{P}\left(|T| \leq X\right) = 0.95$.

9. Let $X_1, X_2, \ldots, X_{11}$ be a random sample of size 11 from a normal distribution with unknown mean μ and variance $\sigma^2 = 9.9$. If $\displaystyle\sum_{i=1}^{11} X_i = 132$, find k so that $[12 - k\sqrt{0.9}, 12 + k\sqrt{0.9}]$ is a 90% confidence interval for μ.

10. If $X_1, X_2, \ldots, X_n$ is a random sample from a normal population with variance $\sigma^2 = 25$, how large must the sample size be so that the length of a 95% confidence interval for μ is 1.96?

11. A building contractor usually gets his light bulbs from shop A, where he knows that the bulbs last for an expected time of 100 months with a standard deviation of 3 months. A new shop B contacts the building contractor and say that he can supply bulbs which will last more than those from shop A maintaining the same standard deviation. The building contractor has to now decide whether he needs to make an order of 1 lakh light bulbs from shop A or shop B. To do this, he asks 16 customers who purchased from shop B and finds out that that these light bulbs lasted for an average of 103 months. The contractor doesn't like to take risk and decides to use hypothesis testing with a very low significance value of 1% to decide. Will he purchase from shop A or shop B? You may assume that the times are normally distributed.

12. A dealer purchases an idli grinder from manufacturer A, where he knows that the expected time to grind 2 kgs of flour to a desired consistency is 15 minutes with a standard deviation of 3 minutes. A new manufacturer B contacts the dealer and proposes that his grinders can grind faster than those from manufacturer A with the same standard deviation of 3 minutes. The dealer samples 10 grinders from B and finds that they take 14 minutes to do the same process. The dealer is willing to take risk and decides to use hypothesis testing with a reasonably high significance value of 10% to decide. Will he purchase from shop A or shop B? You may assume that the times are normally distributed.

13. A premium gold ball production line must produce all of its balls of 45.75 grams in order to get the top rating. Samples are drawn hourly and checked. If the production line gets out os sync with a statistical significance of more than 1%, it must be shut down and repaired. A sample of 18 balls has a mean of 45 grams and a sample standard deviation of 2 grams. Should the production line be repaired?

14. In the first problem, the second shop owner (B) says that his bulbs last for an average of 105 months with a standard deviation of 3 months. If the contractor wants his type I error to be 1% and type II error to be 5%, how many samples should he check and what is his critical value for accepting and rejecting the null hypothesis?