Mahindra
École Centrale
COLLEGE OF ENGINEERING

MA 203: Problem Sheet 7: Statistics
Assignment submission deadline 15/11/2018

∗ Problems to be submitted as Assignment

1. Let X be a Gaussian random variable with mean 10 and variance 4. A sample of size 9 is obtained and the sample mean, minimum, and maximum of the sample are calculated.

   (a) Find the probability that the sample mean is less than 9.

   (b) Find the probability that the minimum is greater than 8.

   (c) Find the probability that the maximum is less than 12.

   (d) Find n so that the sample mean is within 1 of the true mean with probability 0.95.

*2. The lifetime of a device is an exponential random variable with mean 50 months. A sample of size 25 is obtained and the sample mean, maximum, and minimum of the sample are calculated.

   (a) Estimate the probability that the sample mean differs from the true mean by more than 1 month.

   (b) Find the probability that the longest-lived sample is greater than 100 months.

   (c) Find the probability that the shortest-lived sample is less than 25 months.

   (d) Find n so that the sample mean is within 5 months of the true mean with probability 0.9.

3. Let the signal X be a uniform random variable in the interval $[-3, 3]$ and suppose that a sample of size 50 is obtained.

   (a) Estimate the probability that the sample mean is outside the interval $[-0.5, 0.5]$.

   (b) Estimate the probability that the maximum of the sample is less than 2.5.

   (c) Estimate the probability that the sample mean of the squares of the samples is greater than 3.

4. Let the sample $X_1, X_2, \cdots, X_n$ consist of iid versions of the random variable X. The method of moments involves estimating the moments of X as follows:

$$\hat{m}_k = \frac{1}{n}\sum_{j=1}^{n} X_j^k.$$

(a) Suppose that X is a uniform random variable in the interval $[0, \theta]$. Use $\hat{m}_1$ to find an estimator for θ.

(b) Find the mean and variance of the estimator in part (a).

*5. Let $\mathbf{X} = (X, Y)$ be a pair of random variables with known means, m_1 and m_2. Consider the following estimator for the covariance of X and Y:

$$\hat{C}_{X,Y} = \frac{1}{n}\sum_{j=1}^{n}(X_j - m_1)(Y_j - m_2).$$

(a) Find the expected value and variance of this estimator.

(b) Explain the behavior of the estimator as n becomes large.

6. Let $\mathbf{X} = (X, Y)$ be a pair of random variables with unknown means and covariances. Consider the following estimator for the covariance of X and Y:

$$\hat{K}_{X,Y} = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X}_n)(Y_j - \bar{Y}_n),$$

where $\bar{X}_n$ and $\bar{Y}_n$ sample means of X and Y respectively.

(a) Find the expected value of this estimator.

(b) Explain why the estimator approaches the estimator in Problem n large.

7. Let the sample $X_1, X_2, \cdots, X_n$ consist of iid versions of the random variable X. Consider the maximum and minimum statistics for the sample:

$$W = min(X_1, X_2, \cdots, X_n), \quad Z = max(X_1, X_2, \cdots, X_n).$$

(a) Show that the pdf of Z is $f_Z(x) = n[F_X(x)]^{n-1} f_X(x)$.

(b) Show that the pdf of W is $f_W(x) = n[1 - F_X(x)]^{n-1} f_X(x)$.

8. Let the sample X_1, X_2, X_3, X_4 consist of iid versions of a Poisson random variable X with mean $\alpha = 4$. Find the mean and variance of the following estimators for α and determine whether they are biased or unbiased.

*(a) $\hat{\alpha}_1 = (X_1 + X_2)/2$.

(b) $\hat{\alpha}_2 = (X_3 + X_4)/2$.

*(c) $\hat{\alpha}_3 = (X_1 + 2X_3)/3$.

(d) $\hat{\alpha}_4 = (X_1 + X_2 + X_3 + X_4)/4$.

9. Let $\hat{\Theta}_1$ and $\hat{\Theta}_2$ be unbiased estimators for the parameter θ.

(a) Show that the estimator $\hat{\Theta} = p\hat{\Theta}_1 + (1-p)\hat{\Theta}_2$ is also an unbiased estimator for θ, where $0 \leq p \leq 1$.

(b) Find the value of p in part (a) that minimizes the mean square error.

(c) Let $\hat{\Theta}_1$ and $\hat{\Theta}_2$ be unbiased estimators for the first and second moments of X. Find an estimator for the variance of X. Is it biased?

*10. The output of a communication system is $Y = \theta + N$, where θ is an input signal and N is a noise signal that is uniformly distributed in the interval $[0, 2]$. Suppose the signal is transmitted n times and that the noise terms are iid random variables.

(a) Show that the sample mean of the outputs is a biased estimator for θ.

(b) Find the mean square error of the estimator.

11. Let X be an exponential random variable with mean $1/\lambda$.

(a) Find the maximum likelihood estimator $\hat{\Theta}_{ML}$ for $\theta = 1/\lambda$.

(b) Find the maximum likelihood estimator $\hat{\Theta}_{ML}$ for $\theta = \lambda$.

(c) Find the pdfs of the estimators in part (a).

(d) Is the estimator in part (a) unbiased and consistent?

12. Let $X = \theta + N$ be the output of a noisy channel where the input is the parameter θ and N is a zero-mean, unit-variance Gaussian random variable. Suppose that the output is measured n times to obtain the random sample $X_i = \theta + N_i$ for $i = 1, 2, \cdots, n$.

(a) Find the maximum likelihood estimator $\hat{\Theta}_{ML}$ for θ.

(b) Find the pdf of $\hat{\Theta}_{ML}$.

(c) Determine whether $\hat{\Theta}_{ML}$ is unbiased and consistent.

13. Let $\hat{\Theta}_{ML}$ be the maximum likelihood estimator for the mean of an exponential random variable. Suppose we estimate the variance of this exponential random variable using the estimator $\hat{\Theta}_{ML}^2$. What is the probability that $\hat{\Theta}_{ML}^2$ is within 5% of the true value of the variance? Assume that the number of samples is large.