

Indian Statistical Institute, Bangalore
B. Math (II), First semester 2016-2017
Mid-Semester Examination : Statistics (I)

Date: 13-09-2016

Maximum Score 50

Duration: 3 Hours

1. The probability mass function (*pmf*) of $NegBin(r, \theta)$ is given by

$$f(x|r, \theta) = \binom{r+x-1}{x} \theta^r (1-\theta)^x ; x = 0, 1, 2, \dots$$

Obtain mean and variance of $NegBin(r, \theta)$. Consider a sequence of *negative binomial distribution* $NegBin(r_n, \theta_n)$, $n \geq 1$. Let $r_n \rightarrow \infty$, and $\theta_n \rightarrow 1$ as $n \rightarrow \infty$ such that $\frac{r_n(1-\theta_n)}{\theta_n} \rightarrow \lambda > 0$ as $n \rightarrow \infty$. Let the mean and variance of $NegBin(r_n, \theta_n)$ be denoted respectively by μ_n and σ_n^2 . Find $\lim_{n \rightarrow \infty} \mu_n$ and $\lim_{n \rightarrow \infty} \sigma_n^2$. Also obtain limit of *pmf* of *negative binomial distribution*, $NegBin(r_n, \theta_n)$, as $n \rightarrow \infty$ and identify the limiting distribution. [5 + 2 + 7 + 1 = 15]

2. Let X_1, X_2, \dots, X_n be a random sample from the distribution whose probability density function (*pdf*) is given by

$$f(x|\theta) = 2(m+1)(x-\theta)^{2m+1} ; \theta < x < \theta+1, -\infty < \theta < \infty.$$

Obtain $E[(X_1)^r]$. Find *method of moments (MOM) estimator* for θ . Find *maximum likelihood estimator (MLE)* for θ . [3 + 4 + 6 = 13]

3. Let $X_1 \sim \chi_m^2$ and $X_2 \sim \chi_n^2$ be independent. Define $Y = X_1$ and $W = \frac{n}{m} \frac{X_1}{X_2}$. Obtain the joint density function $f_{YW}(y, w)$ of Y and W . Hence obtain the marginal density function $f_W(w)$ of W and identify it. [12]

4. Following is the data set of daily minimum temperature at a hill station recorded in $^{\circ}F$ during the month of April.

77, 80, 82, 68, 65, 59, 61, 57, 50, 62, 61, 70, 69, 64, 67,

70, 62, 65, 65, 73, 76, 87, 80, 82, 83, 79, 79, 71, 80, 77.

- (a) Make a stem and leaf plot of these data.
- (b) Find the sample mean \bar{X} .
- (c) Find 100 p -th percentile for $p = 0.25, 0.50$ and 0.75 .
- (d) Find the first quartile Q_1 , median M and the third quartile Q_3 .
- (e) Draw the box plot and identify the outliers.
- (f) Explain how to obtain the trimmed mean \bar{X}_T . Decide on trimming fraction just enough to eliminate the outliers and obtain the trimmed mean \bar{X}_T .
- (g) Explain how to obtain the trimmed standard deviation S_T .
- (h) Between the box plot and the stem and leaf plot what do they tell us about the above data set? In very general terms what can you say about the population from which the data arrived? [3 + 1 + 3 + 1 + 4 + 2 + 2 + 4 = 20]