Indian Institute of Technology Guwahati
Probability Theory and Stochastic Processes (MA225)
Problem Set 02

1. Are the following functions cumulative distribution functions?

   (a) $F_1(x) = \begin{cases} 0 & \text{if } x < -5 \\ x & \text{if } -5 \le x \le 0.5 \\ 1 & \text{if } x > 0.5. \end{cases}$

   (b) $F_2(x) = \dfrac{1}{2} + \dfrac{1}{\pi} \tan^{-1}(x),\ -\infty < x < +\infty.$

   (c) $F_3(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-x} & \text{if } x \ge 0. \end{cases}$

   (d) $F_4(x) = \begin{cases} 0 & \text{if } x < 0 \\ \dfrac{1}{\Gamma(\alpha)} \displaystyle\int_0^x y^{\alpha-1} e^{-y} dy & \text{if } x \ge 0. \end{cases}$

2. Let X be a random variable having the distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{2}{3} & \text{if } 0 \le x < 1 \\ \frac{7-6c}{6} & \text{if } 1 \le x < 2 \\ \frac{4c^2-9c+6}{4} & \text{if } 2 \le x \le 3 \\ 1 & \text{if } x > 3, \end{cases}$$

   where c is a real constant.

   (a) Find the value of constant c. (Ans: 1/4.)

   (b) Using the distribution function, find $P(\{1 < X < 2\})$, $P(\{2 \le X < 3\})$, $P(\{0 < X \le 1\})$, $P(\{1 \le X \le 2\})$, $P(\{X \ge 3\})$, and $P(\{X = 2.5\})$. (Ans: 0, 1/12, 1/4, 1/3, 0, 0.)

   (c) Find the conditional probabilities $P(\{X = 1\} | \{1 \le X \le 2\})$, $P(\{1 \le X < 2\} | \{X > 1\})$, and $P(\{1 \le X \le 2\} | \{X = 1\})$. (Ans: 3/4, 0, 1.)

   (d) Find the PMF of X. (Ans: $2\mathbb{I}_{\{0\}}(x)/3 + \mathbb{I}_{\{1\}}(x)/4 + \mathbb{I}_{\{2\}}(x)/12$.)

3. Let X be a random variable having CDF

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \left(1 - (1-p)^{[x]}\right) & \text{if } x \ge 0. \end{cases}$$

   Determine whether X is DRV or CRV. Find the PMF/PDF, whatever applicable, of X.

4. Twenty distinguishable balls are placed at random in six boxes that are labeled as $B_1, \ldots, B_6$. Find the probability that boxes with labels B_1, B_2, and B_3 all together contain six balls. (Ans: $\binom{20}{6}(1/2)^{20}$.)

5. Each child in a family is equally likely to be a boy or a girl. Find the minimum number of children the family should have so that the probability of it having at least a boy and at least a girl is at least 0.90. (Ans: 5.)

6. Let $X \sim \text{Bin}(n, p)$. Find the mode of X. [m is said to be mode of a DRV X if PMF of X becomes maximum at m.] (Hint: PMF cannot be maximized using differentiation (why?). Ans: For $(n+1)p$ integer, there are two modes, *viz.*, $(n+1)p$ and $(n+1)p+1$. For $(n+1)p$ not an integer, there is one mode, *viz.*, $[(n+1)p]+1$.)

7. Let $X \sim \text{Bin}(n, p)$ and let $k \in \{1, 2, \ldots, n\}$. Show that

$$P(X \ge k) = k\binom{n}{k} \int_0^p t^{k-1} (1-t)^{n-k}\, dt.$$

   Hence show that $P(X \ge k) \le \binom{n}{k} p^k$. (Hint: Use by-parts taking $(1-t)^{n-k}$ as first function and t^{k-1} as second function.)

8. Let $n\,(\geq 2)$ and $r \in \{1, 2, \ldots, n-1\}$ be fixed integers and let $p \in (0,1)$ be a fixed real number. Using the probabilistic arguments show that

$$\sum_{j=r}^{n} \binom{n}{j} p^j (1-p)^{n-j} - \sum_{j=r}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = \binom{n-1}{r-1} p^r (1-p)^{n-r}.$$

(Hint: Let a Bernoulli experiment be repeated independently. Let the probability of success in each trail is p. Change the side of the second term on the left hand side. Then try to interpret both the side in terms of probabilities of events of the above mentioned experiment.)

9. (Negative Binomial Distribution) Let r be a given positive integer. Suppose that we keep performing independent Bernoulli trials until the rth success is observed. Further suppose that the probability of success in each trial is $p \in (0,1)$. Let X be a RV that denotes the number of failures preceding the r-th success. Find the PMF of X. (The probability distribution having this PMF is called the negative binomial distribution and is denoted by $\mathrm{NB}(r,\, p)$. We will use $X \sim \mathrm{NB}(r,\, p)$ to denote that random variable X follows a negative binomial distribution.)

10. (Lack of memory property of Geometric distribution) Let $X \sim \mathrm{Geo}(p)$. Then prove that

$$P\left(\{X \geq m+n\} | \{X \geq m\}\right) = P\left(\{X \geq n\}\right).$$

11. A mathematician carries at all times two match boxes, one in his left pocket and one in his right pocket. To begin with each match box contains n matches. Each time the mathematician needs a match he is equally likely to take it from either pockets. Consider the moment when the mathematician for the first time discovers that one of the match boxes is empty. Find the probability that at that moment the other box contains exactly k matches, where $k \in \{0,\, 1,\, \ldots,\, n\}$. (Hint: You may try to use negative binomial distribution with proper choice of parameters. (Ans: $\binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}$.)

12. (Hypergeometric Distribution) Consider a population comprising of $N\,(\geq 2)$ units out of which $a \in \{1,\, 2,\, \ldots,\, N-1\}$ are labeled as S (success) and rest $N-a$ are labeled as F (failure). Suppose that we are interested in drawing a sample of size n from this population, drawing one unit at a time. Let X denote the number of S in the drawn sample.

    (a) Assuming that the draws are independent and sampling is with replacement, find the PMF of X.

    (b) Assuming that sampling is without replacement, find the PMF of X. Are the draws independent? (The probability distribution having this p.m.f. is called the Hypergeometric distribution and is denoted by $\mathrm{Hyp}(a,\, n,\, N)$, $a,\, n,\, N \in \mathbb{N}$, $N \geq 2$, $n \leq N-1$, $a \leq N-1$. We shall use the notation $X \sim \mathrm{Hyp}(a,\, n,\, N)$ to indicate that the random variable X follows $\mathrm{Hyp}(a,\, n,\, N)$ distribution.)

13. Suppose that a particular trait (such as eye color or left handedness) of a person is classified on the basis of one pair of genes. Suppose that d and r represent a dominant gene and a recessive gene, respectively. Thus a person with dd genes is pure dominant and one with rr is pure recessive. One with rd is hybrid. The pure dominance and the hybrid are alike in appearance. Children receive one gene from each parent. If, with respect to a particular trail, two hybrid parents have a total of four children, what is the probability that exactly three of the four children have the outward appearance of the dominant genes? (Ans: 27/64.)