E1 222 Stochastic Models and Applications

Problem Sheet 2.3

(You need not submit the solutions)

1. Consider the random experiment of five independent tosses of a fair coin. In any outcome (of this random experiment) we say a change-over has occurred at the i^{th} toss if the result of the i^{th} toss differs from that of the $(i-1)^{th}$ toss. Let X be a random variable whose value is the number of change-overs. For example, if the outcome of the random experiment is $HTTHH$ then the value of X would be 2. Note that the minimum value of X is 0 (e.g., when the outcome is $HHHHH$) and the maximum value of X is 4 (e.g., when the outcome is $HTHTH$). Find the probability mass function of X. Generalize this to the case of n tosses.

2. Let p be a number such that $0 < p < 1$ and let U be a random variable distributed uniformly over $(0,\ 1)$. Let X be a random variable defined by
$$X = \mathrm{Int}\left(\frac{\log(1-U)}{\log(1-p)}\right) + 1$$
where $\mathrm{Int}(x)$ is the largest integer smaller than or equal to x. Find the distribution of X.

3. Let F be the distribution function of a random variable that takes values in $\{0, 1, 2, \cdots\}$. Consider the following procedure (written as a pseudo-code) for determining the value of a random variable X

   1 Generate Z uniform over $[0,\ 1]$.

   2 Set k=0.

   3 *while* $Z > F(k)$ set $k = k + 1$

   4 Set $X = k$ and exit

   What would be the distribution function of X? What is the expected number of steps spent in the while loop?

4. The price of some commodity is Rs. 2 per gram this week. Next week the price would be either Rs.1 per gm or Rs. 4 per gram, each with

probability 0.5. You have a capital of Rs.1000. What would be your strategy if (i) you want to maximize expected amount of money with you (next week), (ii) you want to maximize the expected quantity of the commodity with you.

5. Children from a school went to a picnic in four buses. Different buses carried different number of students. Define two random variables, X, Y, as follows. We select one of the four drivers at random and X is the number of students in the bus driven by that driver. We select a student at random and Y is the number of students in the bus in which the selected student travelled. Can you say whether $EX > EY$ or $EY > EX$ (or the information given is not sufficient to decide which of EX, EY is greater)?

6. For a continuous random variable, X, the real number a that satisfies $\int_{-\infty}^{a} f_X(x)\, dx = 0.5$ is called the median of X. Show that for a continuous random variable, X, the number x_0 that minimizes $E[\, |X - x_0|\,]$ is the median of X.

7. Suppose X is a discrete random variable taking positive integer values. Assume $E|X| < \infty$. Show that

$$E[X] = \sum_{k=0}^{\infty} P[X > k].$$

8. Let X be a continuous random variable with $E|X|^k < \infty$ for some $k > 0$. Then show that $n^k P[|X| > n] \to 0$ as $n \to \infty$. (Hint: Write the expectation integral of $|X|^k$ as two parts one for $|x| \leq n$ and the other for $|x| > n$. Since the integral is finite, argue that the second part goes to zero. Then try and bound the second integral in terms of $P[|X| > n]$).

9. Let X be a nonnegative continuous random variable and suppose EX exists. Show that

$$EX = \int_0^\infty (1 - F(x))\, dx$$

(Hint: Integrate by parts and use the previous problem).

10. Consider the following density function (called Beta density)

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \ 0 \le x \le 1.$$

where $\Gamma(\cdot)$ is the gamma function and $a, b \ge 1$ are parameters. Show that this is a density as follows. By definition of gamma function, we have

$$\Gamma(a)\Gamma(b) = \int_0^\infty x^{a-1} e^{-x} \, dx \int_0^\infty y^{b-1} e^{-y} \, dy$$

First bring the integral over y inside the integral over x. Now in the inner integral change the variable from y to t using $t = y + x$. Now change the order of the x and t integrals so that the x integral becomes the inner integral. Now, in the inner integral change the variable from x to s using $x = ts$. The final expression you get can then be used to show that the above $f(x)$ is a density.

11. Suppose an experiment can result in one of r possible outcomes and the i^{th} outcome has probability p_i, $i = 1, 2, \cdots, r$. (Note that $\sum_{i=1}^{r} p_i = 1$). Suppose we have n independent repetitions of this experiment. Argue that the probability that the first outcome occurs x_1 times, the second x_2 times and so on, is

$$\frac{n!}{x_1! x_2! \cdots x_r!} \, p_1^{x_1} p_2^{x_2} \cdots p_r^{x_r}$$

where $x_1 + x_2 + \cdots + x_r = n$. This is known as the mutinomial distribution. What would this be if $n = 2$?

12. A coin having probability p of coming up heads is successively tossed till the r^{th} head appears. (p and r are parameters). Let X denote the number of tosses needed. Find the mass function of X. (Hint: To calculate $P[X = n]$, think of how many heads are allowed in the first $n - 1$ tosses).

13. Consider a random variable X with the mass function

$$f(x) = {}^{(\alpha+x-1)}C_x \, p^\alpha (1-p)^x, \quad x = 0, 1, \cdots$$

where $\alpha > 0$. Is this realted to the X in the previous problem? This is known as the negative binomial distribution. The motivation for the name can be seen as follows. For any positive real number α and a nonnegative integer x we have

$$\begin{aligned} {}^{-\alpha}C_x &= \frac{-\alpha(-\alpha-1)(-\alpha-x+1)}{x!} \\ &= \frac{(-1)^x(\alpha)(\alpha+1)(\alpha+x-1)}{x!} \\ &= {}^{(\alpha+x-1)}C_x\ (-1)^x \end{aligned}$$

Thus ${}^{(\alpha+x-1)}C_x\ p^\alpha(1-p)^x = {}^{-\alpha}C_x p^\alpha(-1)^x(1-p)^x$. Thus our distribution can be seen to involve binomial coefficients for negative index and hence the name. What will this distribution be for $\alpha = 1$?

14. The binomial distribution can be approximated by the Poisson distribution for large n. Consider a binomial distribution with parameters n and p. Since, the expectation is np, if we want an approximation as n tends to infinity we need to ensure that the expectation is finite. So, let us write p_n as the probability of success when we consider n trials and let us assume that as $n \to \infty$, $np_n \to \lambda$. Noting that, as $n \to \infty$, we have (i). $(1-\frac{\lambda}{n})^n \to e^{-\lambda}$, (ii). $(1-\frac{\lambda}{n})^{-k} \to 1$, (iii). $(n(n-1)\cdots(n-k+1))/(n^k) \to 1$, show that

$$\lim_{n\to\infty} {}^nC_k(p_n)^k(1-p_n)^{n-k} = \frac{\lambda^k}{k!}e^{-\lambda}$$

15. This problem is about calculating the mode of binomial distribution. That is, if we consider n tosses of a coin with probability of heads as p, we want to know what is the most probable number of heads. Let X be binomial with parameters, n, p. Then, as k goes from 0 to n, $P[X=k]$ first increases monotonically and then decreases monotonically. You can show this as follows. Derive a condition on k to satisfy $P[X=k] < P[X=k+1]$ and similarly for $P[X=k] > P[X=k+1]$. Using these, show the following.

    a. If $(n+1)p$ is an integer then $P[X=k]$ attains its maximum value at $(n+1)p-1$ or at $(n+1)p$

    b. If $(n+1)p$ is not an integer then $P[X=k]$ attains its maximum value when k satisfies $(n+1)p-1 < k < (n+1)p$.