SOR 201 Examples 3

1. A box contains N cards numbered 1 to N; n cards, $n \leq N$, are randomly selected without replacement. Let the random variable X be the largest number selected.

   (a) Define an outcome of this experiment. What values can X take? If $X = x$, what can be said about the other numbers selected? Explain why

   $$\mathrm{P}(X = x) = \binom{x-1}{n-1} \Big/ \binom{N}{n}$$

   for relevant values of x.

   (b) <u>Alternative method</u> Using the fact that the events 'largest number in the sample $\leq y$' and 'all numbers in the sample $\leq y$' are equivalent, find the cumulative distribution function of X i.e. $\mathrm{P}(X \leq y)$, and hence $\mathrm{P}(X = x)$.

2. (i) Suppose the random variable X has probability function $p(x)$, $\quad x = 0, \pm 1, \pm 2, \ldots$. Find the probability functions of the following transformations of X :

   (a) $Y = X^2$ (b) $W = |X|$

   (c) $Z = \mathrm{sgn}(X) = \begin{cases} X/|X|, & X \neq 0 \\ 0, & X = 0. \end{cases}$

   (ii) Prove that $\mathrm{E}[(X - a)^2]$ is minimized w.r.t. a when $a = \mathrm{E}(X)$.

3. (i) Let the random variable X have the geometric distribution with parameter p,

   $$\text{i.e.} \quad \mathrm{P}(X = x) = pq^{x-1}, \qquad x = 1, 2, \ldots.$$

   (a) Show that $\mathrm{P}(X > m) = q^m$, where m is a positive integer.

   (b) Show that

   $$\mathrm{P}(X > m + n | X > m) = \mathrm{P}(X > n),$$

   where m and n are positive integers i.e. the geometric distribution has the 'no-memory' property.

   (ii) Suppose the random variable X takes non-negative integer values, i.e. X is a count random variable. Prove that

   $$\mathrm{E}(X) = \sum_{x=0}^{\infty} \mathrm{P}(X > x).$$

   [<u>Hint</u>: Since all terms in the summation are non-negative, we may rearrange the order of the terms.]

   Hence show that the geometric distribution defined in (i) has mean $1/p$.

/continued overleaf

4. Let X and Y be independent random variables which have the same geometric distribution:

$$P(X = k) = P(Y = k) = pq^k, \qquad k = 0, 1, 2, \ldots,; \qquad p + q = 1.$$

(a) Calculate $P(X = Y)$.

(b) Show that the conditional distribution of X given $X+Y$ is the discrete uniform distribution, i.e.

$$P(X = x|X + Y = n) = \frac{1}{n+1}, \qquad x = 0, 1, \ldots, n.$$

(c) If $U = \min(X, Y)$, find the probability distribution of U.

5. In 5 tosses of a fair coin let X,Y and Z be the number of heads, the number of head runs and the length of the longest head run respectively.

(a) Tabulate the 32 outcomes in the sample space

$$\mathcal{S} = \{HHHHH, HHHHT, \ldots, TTTTT\},$$

together with the corresponding values of X,Y and Z.

(b) Tabulate the probability functions of the distributions of
(X, Y, Z), (X, Y), (Y, Z), X, Y, Z.

(c) Calculate the correlation coefficient $\rho(Y, Z)$.

(d) Tabulate the probability functions

$$\{P(X = x|Y = 1)\} \text{ and } \{P(Y = y, Z = z|X = 3)\}.$$

(e) Tabulate the probability function of the random variable $E(Y|X)$. Hence, or otherwise, calculate $E[E(Y|X)]$ and verify that it is equal to $E(Y)$.

6. A prisoner is trapped in a dark cell containing 3 doors. The first door leads to a tunnel that returns him to his cell after 2 days of travel, the second leads to a tunnel that returns him to his cell after 4 days, while the third door leads to freedom after 1 day. If it is assumed that the prisoner will always select doors 1,2 and 3 with probabilities 0.5, 0.3 and 0.2 respectively, what is the expected time for the prisoner to reach freedom?

Additional questions (NOT for handing in)

7. (a) Suppose the discrete random variable X takes the values 1,2, ..., $2N$ and its probability function is symmetrical about the value $(N + \frac{1}{2})$. Give a rough sketch of an example of such a probability function and hence write down a probability relationship for $\mathrm{P}(X = x), \quad x = 1, \ldots, N$.

   Write down a summation expression for $\mathrm{E}(X-N-\frac{1}{2})$. By looking at individual terms, calculate $\mathrm{E}(X-N-\frac{1}{2})$ and hence find $\mathrm{E}(X)$. Also calculate the median of X.

   (b) Suppose the discrete random variable Y takes the values 0,1, ..., $2M$ and its probability function is symmetrical about the value M. Repeat (a) for this distribution.

8. Two people toss a fair coin n times each. Show that, if X_i is the number of heads thrown by person i, then

$$\mathrm{P}(X_1 = X_2) = \mathrm{P}(X_1 + X_2 = n)$$

   [Hint: use the fact that

$$\mathrm{P}(\text{person throws } k \text{ heads}) = \mathrm{P}(\text{person throws } k \text{ tails}), \quad k = 0, ..., n.]$$

   Hence show that the probability of the two people throwing equal numbers of heads is

$$\binom{2n}{n}/(2)^{2n}.$$

9. For the random variables X, Y, Z in Question 5, determine also

   (a) the probability function of the distribution of (X,Z);

   (b) $\rho(X, Y)$ and $\rho(X, Z)$;

   (c) the probability function $\{\mathrm{P}(X = x|Y = 2, Z = 2)\}$;

   (d) the probability function of the random variable $\mathrm{E}(Z|X)$, and hence $\mathrm{E}[\mathrm{E}(Z|X)]$ (verifying that it is equal to $\mathrm{E}(Z)$).

10. In a certain dice game, a pair of fair dice are rolled. If the sum of the scores is 7, the game ends and the player wins 0. Otherwise, the player has the option of either stopping the game and receiving an amount equal to that sum, or starting over again. Suppose a player adopts a strategy of stopping when a sum greater or equal to k is obtained. What value of k maximises the expected return?

    [Hint: Let X_k be the return when the critical value k is used. Compute $\mathrm{E}(X_k)$ by conditioning on the value of the sum obtained in the first roll.]