Biostatistics: Sheet 2 –
Discrete Random Variables
and their Probability
Distributions

1. Suppose that a random variable X has a discrete distribution with the following probability function:

$$f(x) = \begin{cases} cx & x = 1, 2, 3, 4, 5, \\ 0 & otherwise \end{cases}$$

   Determine the value of the constant c.

2. A discrete random variable has a pdf of the form f(x) = c(8-x) for x=0,1,2,3,4,5 and zero otherwise.

   (a) Find the constant c
   (b) Find the CDF, F(x)
   (c) Find P(X>2).

3. Is the following a valid CDF? If not, why not, and how could you modify it to be a valid one?

$$F_X(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{2} & 2 \leq x < 3 \\ \frac{3}{4} & 3 \leq x \leq 4 \\ 1 & x \geq 4 \end{cases}$$

4. Derive the expected value (the mean) for the following distributions:
   a) Bernoulli distribution
   b) Binomial Distribution
   c) Poisson Distribution
   d) Geometric distribution

5. If a discrete random variable X has the PMF $p_X[k]=1/5$ for k=0,1,2,3,4. If Y=sin((π/2)X), find E[Y].

6. Given a discrete random variable X with the following PMF:

$$p_X[k] = \frac{4/\pi^2}{k^2} \qquad k = 1, 2, \ldots \ldots$$

   This can be a PMF since it can be shown that it sums up to one. Try to find the mean of this PMF. [Hint: $\sum_{k=1}^{\infty}(1/k) \rightarrow \infty$ since it is a harmonic series]

Biostatistics: Sheet 2 –
Discrete Random Variables
and their Probability
Distributions

7. The expected value of a random variable is the best predictor of the outcome of the experiment, where best is to be interpreted as the value that minimized the mean square error (MSE)
   Example: The amount of rainfall in a particular area in winter is a random variable. A farmer in this area would like to be able to predict the amount of rainfall ahead in order to decide which crops to plant. We want to choose a predicted value b that is on the average is close to the true outcome of the random value (here the amount of rainfall).
   Prove that choosing b to be equal to the Expected value of the rainfall is the best estimate in the sense that it minimizes the mean square error.

8. A die is tossed. The probability of obtaining a 1, 2, or 3 is the same. Also, the probability of obtaining a 4, 5, or 6 is the same. However, a 5 is twice as likely to be observed as a 1. For a large number of tosses, what is the average value observed?

9. A fax machine dials a phone number which is typically busy 80% of the time. The machine dials it every 5 minutes until the line is clear and the fax is able to be transmitted. What is the probability that the fax machine will have to dial the number 9 times?

10. The arrival rate of calls at a mobile switching station is 1 per second. The probability of *k* calls in a *T* second interval is given by a passion PMF with λ= arrival rate x *T*. what is the probability that there will be more than 100 calls in a 1 minute interval?

11. Let the random variable X have a discrete uniform distribution on the integers 1<=x<=3. Determine the mean and variance of X.

12. The probability of a successful optical alignment in the assembly of an optical data storage product is 0.8. Assume the trials are independent.

(a) What is the probability that the first successful alignment requires exactly four trials?
(b) What is the probability that the first successful alignment requires at most four trials?
(c) What is the probability that the first successful alignment requires at least four trials?

13. A fault-tolerant system that processes transactions for a financial services firm uses three separate computers. If the operating computer fails, one of the two spares can be immediately switched online. After the second computer fails, the last computer can be immediately switched online. Assume that the probability of a failure during any transaction is and that the transactions can be considered to be independent events.
    (a) What is the mean number of transactions before all computers have failed?
    (b) What is the variance of the number of transactions before all computers have failed?