CSSC 432 Probability and Statistics - Internal 2
Department of Computer Science, Pondicherry University

Marks: 25

I. <u>Answer any 5</u> **5 x 5 = 25**

1. (a). State properties of a Normal Distribution.

The normal distribution, also known as the Gaussian distribution or bell curve, is a common probability distribution that is widely used in statistical analysis. Some of the key properties of a normal distribution are:

1. Symmetry: The normal distribution is symmetric about its mean, which means that the left and right halves of the distribution are mirror images of each other.
2. Unimodality: The normal distribution has a single peak or mode.
3. Bell-shaped: The normal distribution has a bell-shaped curve, which means that the majority of the data falls near the mean, and the further away from the mean the data is, the less likely it is to occur.
4. Continuous: The normal distribution is a continuous probability distribution, which means that the values it can take are not restricted to discrete values.
5. Mean and standard deviation: The normal distribution is completely determined by its mean and standard deviation. The mean represents the center of the distribution, while the standard deviation represents the spread of the data.
6. Empirical rule: The normal distribution follows the empirical rule, which states that approximately 68% of the data falls within one standard deviation of the mean, 95% falls within two standard deviations, and 99.7% falls within three standard deviations.
7. Z-score: The z-score, which represents the number of standard deviations a data point is from the mean, can be used to standardize normal distributions so that they can be compared across different datasets.

(b). The weights in kilogram of parcels arriving at a package delivery company's warehouse can be modelled by an N(5, 16) normal random variable, X. Calculate the standard deviation of the distribution and the probability that a randomly selected parcel weighs more than 13 kilogram.

Variance $\sigma_x^2 = 16$. Hence, Standard Deviation $\sigma_x = 4$

The probability that a randomly selected parcel weighs more than 13 kilogram is

$P[X > 13] = 1 - P[X \le 13] = 1 - F_X(13) = 1 - \Phi(13 - 5)/ 4 = 1 - \Phi(2) = 1 - 0.9772 = 0.0228$

2. (a). State Markov's and Chebychev's Inequality

The Chebyshev inequality is a statement that places a bound on the probability that an experimental value of a random variable X with finite mean $E[X] = \mu_X$ and variance σ_X^2 will differ from the mean by more than a fixed number a. The statement says that the bound is directly proportional to the variance and inversely proportional to a^2. That is,

$$P[|X - E[X]| \geq a] \leq \frac{\sigma_X^2}{a^2} \quad a > 0$$

The Markov inequality applies to random variables that take only nonnegative values. If X is a random variable that takes only nonnegative values, then for any value $a > 0$,

$$P[X \geq a] \leq \frac{E[X]}{a}$$

(b). A random variable X has a mean of 2 and a variance of 1. Obtain an upper bound for $P[|X - 2| \geq 3]$ and $P[X \geq 3]$.

$$P[|X - 2| \geq 3] \leq \sigma_X^2 / 3^2$$
$$\leq 1/9$$

$$P[X \geq 3] \leq E[X] / 3$$
$$\leq 2/3$$

3. (a). Find the variance of a continuous random variable X whose PDF is given by $f_X(x) = 2e^{-2x}$; $x \geq 0$

<u>Note:</u> Substitute **λ** by 2 in the following solution.

Solution From Example 3.4, the expected value of X is given by $1/\lambda$. The second moment of X is given by

$$E[X^2] = \int_0^\infty x^2 \lambda e^{-\lambda x} dx$$

Let $dv = \lambda e^{-\lambda x} dx$ and $u = x^2$. This means that $v = -e^{-\lambda x}$ and $du = 2xdx$. Thus, integrating by parts we obtain

$$\begin{aligned} E[X^2] &= [uv]_0^\infty - \int_0^\infty v du \\ &= [-x^2 e^{-\lambda x}]_0^\infty + \int_0^\infty 2xe^{-\lambda x} dx = 0 + 2\left[-\frac{xe^{-\lambda x}}{\lambda}\right]_0^\infty + \int_0^\infty \frac{2e^{-\lambda x}}{\lambda} dx \\ &= \frac{2}{\lambda^2} \end{aligned}$$

Thus, the variance of X is given by

$$\sigma_X^2 = E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Thus, $\boldsymbol{\sigma^2_X = 1/4}$

(b). Find the expected value of a discrete random variable K whose PMF is given by

$$p_K(k) = (5^k e^{-5}) / k! \; ; k = 0, 1, 2, \ldots$$

Note: Substitute **λ** by 5 in the following solution.

Solution $E[K]$ is given by

$$\begin{aligned} E[K] &= \sum_{k=0}^{\infty} k p_K(k) = \sum_{k=0}^{\infty} k\left(\frac{\lambda^k}{k!}e^{-\lambda}\right) \\ &= \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda} \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

Since

$$\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$$

we obtain

$$E[K] = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Thus **E[K] = 5**

4. (a). The joint PDF of the random variables X and Y is defined as follows

$$f_{XY}(x, y) = 25e^{-5y} \; ; 0 < x < 0.2 \; ; \; y > 0 \text{ and } f_{XY}(x, y) = 0 \text{ otherwise}$$

Find the covariance of X and Y if the marginal distributions of X and Y are uniform and exponential with **λ** = 5, respectively.

Thus, X has a uniform distribution, and Y has an exponential distribution. The expected values of X and Y are given by

$$\begin{aligned} E[X] &= \mu_X = \frac{0+0.2}{2} = 0.1 \\ E[Y] &= \mu_Y = 1/5 = 0.2 \end{aligned}$$

Also,

$$\begin{aligned} E[XY] &= \int_{x=0}^{0.2}\int_{y=0}^{\infty} xyf_{XY}(x,y)dydx = \int_{x=0}^{0.2}\int_{y=0}^{\infty} 25xye^{-5y}dydx \\ &= \int_{x=0}^{0.2} x\left\{\int_{y=0}^{\infty} 25ye^{-5y}dy\right\}dx = \int_{x=0}^{0.2} xdx = \left[\frac{x^2}{2}\right]_0^{0.2} \\ &= 0.02 \end{aligned}$$

Thus, the covariance of X and Y is given by

$$\sigma_{XY} = E[XY] - \mu_X\mu_Y = 0.02 - (0.1)(0.2) = 0$$

(b). A student doing a summer internship in a company was asked to model the lifetime of certain equipment that the company makes. After a series of tests, the student proposed that the lifetime of the equipment can be modelled by a random variable X that has the PDF

$$f(x) = (xe^{-x/10})/10 \; ; \; x \geq 0 \text{ and } f(x) = 0 \text{ otherwise}$$

Check whether f(x) is a valid PDF.

Solution: The given f(x) is **NOT** a valid PDF as the integral over (-Inf, Inf) is not equal to 1 i.e., the total probability is not equal to 1

5. A test engineer discovered that the CDF of the lifetime of equipment in years is given by

$$F_X(x) = 0 \; ; \; x < 0 \text{ and } F_X(x) = 1 - e^{-x/4} \; ; \; 0 \leq x < \infty$$

Find the expected lifetime of the equipment with its variance.

Solution From the definition of its CDF, we can see that X is a random variable that takes only nonnegative values. Thus,

(a) The expected lifetime of the equipment is given by

$$E[X] = \int_0^\infty P[X > x]dx = \int_0^\infty [1 - F_X(x)]dx$$

$$= \int_0^\infty e^{-x/4}dx = 4$$

$E[X^2] =$ 32 (calculated using Integration by parts similar to the solution below)

Thus, the second moment of X is given by

$$E[X^2] = \int_{-\infty}^\infty x^2 f_X(x)dx = \frac{1}{5}\int_0^\infty x^2 e^{-x/5}dx$$

Let $u = x^2 \Rightarrow du = 2xdx$, and let $dv = e^{-x/5}dx \Rightarrow v = -5e^{-x/5}$. Thus,

$$E[X^2] = \left\{-\frac{5x^2e^{-x/5}}{5}\right\}_0^\infty + 10\int_0^\infty \frac{xe^{-x/5}}{5}dx$$

$$= 0 + 2\int_0^\infty xe^{-x/5}dx = 2\int_0^\infty xe^{-x/5}dx$$

Let $u = x \Rightarrow du = dx$, and let $dv = e^{-x/5}dx \Rightarrow v = -5e^{-x/5}$. Then we have that

$$E[X^2] = 2\{-5xe^{-x/5}\}_0^\infty + 10\int_0^\infty e^{-x/5}dx = 0 + 10[-5e^{-x/5}]_0^\infty = 50$$

Finally, the Variance of X is given by

$\sigma^2_X = E[X^2] - \{E[X]\}^2 = 32 - 16 = \mathbf{16}$

6. (a). Let the experiment be tossing three virtual coins on a computer. Each coin can assume any of the two states such as ‘H’or ‘T’ as an outcome. If the outcomes are independent, find the PMF of number of tails obtained.

Solution: Let X(n) denote the number of tails that appear in three tosses. Then X is a binomial random variable with n=3 and p= ½. Thus the PMF of X(n) is

$P_{X(3)}(x) = {}^3C_x\,(½)^x\,(½)^{3-x} = {}^3C_x\,(½)^3 = {}^3C_x\,(⅛)$
$P_{X(3)}(0) = {}^3C_0\,(⅛) = ⅛$
$P_{X(3)}(1) = {}^3C_1\,(⅛) = ⅜$
$P_{X(3)}(2) = {}^3C_2\,(⅛) = ⅜$
$P_{X(3)}(3) = {}^3C_3\,(⅛) = ⅛$

(b). List two discrete and one continuous probability distribution. Give their PDF/ PMF with their mean and variance.

Solution: Any 2 PMF and 1 PDF from the table below:

Random Variable	PMF	PDF	CDF	Mean	Variance
Bernoulli	$p_X(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \end{cases}$	–	$F_X(x) = \begin{cases} 0 & x<0 \\ 1-p & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$	p	$p(1-p)$
Binomial	$p_{X(n)}(x) = \binom{n}{x} p^x (1-p)^{n-x}$ where $x = 0, 1, 2, \ldots, n$	–	$F_{X(n)}(x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$ for $0 \le x < n$ and $F_{X(n)}(x) = 1$ for $x \ge n$	np	$np(1-p)$
Geometric	$p_X(x) = p(1-p)^{x-1}$ where $x = 1, 2, 3, \ldots$	–	$F_X(x) = 1-(1-p)^x \quad x \ge 1$	$1/p$	$\frac{1-p}{p^2}$
Pascal-k	$p_{X_k}(n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}$ where $k = 1, 2, \ldots; n = k, k+1, \ldots$	– –	$F_{X_k}(x) = \sum_{n=k}^{x} \binom{n-1}{k-1} p^k (1-p)^{n-k}$ for $x = k, k+1, \ldots$	k/p	$\frac{k(1-p)}{p^2}$
Hyper-geometric	$p_{K_n}(k) = \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}}$ where $k = 0, 1, \ldots, \min(n, N_1)$	–	$F_{K_n}(k) = \sum_{j=0}^{k} \left\{ \frac{\binom{N_1}{j}\binom{N-N_1}{n-j}}{\binom{N}{n}} \right\}$ for $k = 0, 1, \ldots, \min(n, N_1)$ and $F_{K_n}(k) = 1$ for $k > \min(n, N_1)$	np $p = \frac{N_1}{N}$	$\frac{n(N-n)p(1-p)}{N-1}$
Poisson	$p_K(k) = \frac{\lambda^k e^{-\lambda}}{k!}$ where $k = 0, 1, 2, \ldots$	–	$F_K(k) = \sum_{r=0}^{k} \frac{\lambda^r}{r!} e^{-\lambda}$ where $k \ge 0$	λ	λ
Exponential	–	$f_X(x) = \lambda e^{-\lambda x}$ where $x \ge 0$ and $\lambda > 0$	$F_X(x) = 1 - e^{-\lambda x}$ where $x \ge 0$	$1/\lambda$	$1/\lambda^2$
Erlang-k	–	$f_{X_k(x)} = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}$ where $k = 1, 2, \ldots,$ and $x \ge 0$	$F_{X_k}(x) = 1 - \sum_{r=0}^{k-1} \frac{(\lambda x)^r}{r!} e^{-\lambda x}$ where $x \ge 0$	k/λ	k/λ^2
Uniform	–	$f_X(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$	$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x \ge b \end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
Normal	–	$f_X(x) = \frac{e^{-(x-\mu_X)/2\sigma_X^2}}{\sqrt{2\pi\sigma_X^2}}$ where $-\infty < x < \infty$	$F_X(x) = \Phi\left(\frac{x-\mu_X}{\sigma_X}\right)$ where $-\infty < x < \infty$	μ_X	σ_X^2