JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

SCHOOL OF COMPUTING AND INFORMATICS (SODEL)

BACHELOR OF SCIENCE INFORMATION TECHNOLOGY (BSC. IT)

PROBABILITY AND STATISTICS II

STA 2200

CAT I

MWAURA STELLA WAMBUI

SCT221-D1-0071/2023S

DON DR. HENRY KISSINGER

20TH JULY, 2025

(a) Giving relevant Examples, Explain what You understand by the Following terms:-

(i) Continous random variable (2 Marks)

This is a variable that can take any value within a given range or interval. Eg the height OF students in a class it can assume any value within a certain range.

(ii) Expectation- (1 Mark)

It's an expected value OF a random variable is the long-term average value it would take over many repetitions OF the experiment; it's usually calculated as the sum OF all possible values weighted by their probabilities.

(b) Given a Fair coin, Perform an experiment using the coin, and hence show that the outcome OF your experiment is a random variable. (4 Marks)

Possible outcomes:- Heads (H) or Tails (T).

Random variable $= X$

$$X = \begin{cases} 1 \text{ if Heads} \\ 0 \text{ if Tails} \end{cases}$$

(c) A fair six-sided die is tossed. You win \$2 if the result is a "1", You win \$1 if the result is a "6" but otherwise you lose \$1. Using this information, define the random variable X, the probability distribution and Find the expected value. (5 Marks)

Outcome	Probability	Winnings (x)
1	1/6	2
6	1/6	1
2,3,4,5	4/6	-1

$E(x) = \left(2 \times \frac{1}{6}\right) + \left(1 \times \frac{1}{6}\right) + \left((-1) \times \frac{4}{6}\right)$

$\Rightarrow \frac{2}{6} + \frac{1}{6} - \frac{4}{6} = -\frac{1}{6}$

$E(x) \approx -0.1667$

(d) The random variable X has probability density function

$$f(x) = \begin{cases} (3.6x - 2.4x^2) & 0 < Y < 1 \\ 0 & \text{Otherwise} \end{cases}$$

Find the <u>Mean</u> and <u>Median</u> of the distribution. Also Find $P(X > 0.5)$ and $Var(X)$. (8 Marks)

$f(x) = 3.6x - 2.4x^2$ for $0 < x < 1$

Find the Mean $E(X)$

$$E(X) = \int_0^1 x f(x)\,dx = \int_0^1 x(3.6x - 2.4x^2)\,dx$$

$$dx = \int_0^1 (3.6x^2 - 2.4x^3)\,dx$$

$$E(X) = \left[1.2x^3 - 0.6x^4\right]_0^1 = 1.2 - 0.6$$

$E(X) = 0.6$

Find the Median M

$P(x \le M) = \int_0^M f(x)\,dx = 0.5$

$F(M) = \int_0^M (3.6x - 2.4x^2)\,dx = \left[1.8x^2 - 0.8x^3\right]_0^M$

$\Rightarrow 1.8M^2 - 0.8M^3 = 0.5 \qquad = \dfrac{0.8M^3 - 1.8M^2 + 0.5 = 0}{0.1}$

$\Rightarrow 8M^3 - 18M^2 + 5 = 0$ if $M = 0.2$

$8(0.2)^3 - 18(0.2)^2 + 5 = (8 \times 0.008) - (18 \times 0.04) + 5 = 0.064 - 0.72 + 5$

$\simeq 4.344 > 0$

$\Rightarrow 8M^3 - 18M^2 + 5 = 0$ if $M = 0.1$

$(8 \times 0.001) - (18 \times 0.01) + 5 = 0.008 - 0.18 + 5$

$\simeq 4.828 > 0$

$\Rightarrow$ if $M = 0.5$

$(8 \times 0.125) - 18 \times 0.25 + 5 = 1 - 4.5 + 5$

$= 1.5 > 0$

$\Rightarrow$ if $M = 0.8$

$(8 \times 0.512) - (18 \times 0.64) + 5 = 4.096 - 11.52 + 5$

$= -2.424 < 0$

Median $\therefore$ $M \simeq 0.615$

(iii) $P(x > 0.5)$

$$P(x > 0.5) = \int_{0.5}^{1} f(x)\,dx$$

$$= \int_{0.5}^{1} (3.6x - 2.4x^2)\,dx$$

$$= 1.8x^2 - 0.8x^3 \Big|_{0.5}^{1}$$

$x = 1$

$$\Rightarrow 1.8(1)^2 - 0.8(1)^3 = 1.8 - 0.8 = 1.0$$

$x = 0.5$

$$(1.8(0.5))^2 - (0.8 \times 0.5)^3) = 0.45 - 0.1 = 0.35$$

$$P(x > 0.5) = 1.0 - 0.35$$

$$\simeq 0.65$$

(iv) Variance $Var(x)$

$$Var(x) = E[x^2] - (E[x])^2$$

$$E[x^2] = \int_0^1 x^2 f(x)\,dx = \int_0^1 x^2(3.6x - 2.4x^2)\,dx$$

$$= \int_0^1 (3.6x^3 - 2.4x^4)\,dx \Rightarrow 3.6 \times \left(\frac{x^4}{4}\right) - 2.4 \times \left(\frac{x^5}{5}\right)\Big|_0^1$$

$$= 3.6 \times 1/4 - 2.4 \times 1/5 = 0.9 - 0.48 = 0.42$$

$$Var(x) = 0.42 - (0.6)^2 = 0.42 - 0.36$$

$$\simeq 0.06$$

QUESTION TWO (30 MARKS)

(a) A commuter plane has 100 seats. The probability that any particular passenger will not show up for a flight is 0.05, independent of other passengers. The airline sells 110 tickets for the flight. Calculate the probability that more passengers show up for the flight than there are seats available (3 Marks)

No of Tickets sold: 110	Probability Passenger does not show up: 0.05
Seats Available: 100	Probability Passenger shows up: 0.95

Number of passengers who show up: $X \sim \text{Binomial}(n = 110, P = 0.95)$

$$P(\text{More than 100 show up}) = P(X > 100)$$

Mean: $M = nP = 110 \times 0.95 = 104.5$

Variance: $\sigma^2 = nP(1-P) = 110 \times 0.95 \times 0.05 = 5.225$

Standard deviation: $\sigma \simeq \sqrt{5.225} \simeq 2.29$

$$P(X > 100) \simeq P(X \geq 101) \simeq P(Y > 100.5)$$

$$Z = \frac{100.5 - 104.5}{2.29} \simeq \frac{-4}{2.29} \simeq -1.75$$

$$P(X > 100) \simeq P(Z > -1.75) = 1 - \Phi(-1.75) = \Phi(1.75)$$

$$\simeq 0.9599$$

$$\simeq 0.96$$

(b.) An actuary has discovered that IT policyholders are three times as likely to file two claims as to file four claims. If the number of claims filed by IT Policy Holders has a Poisson Distribution. What is the standard deviation of the number of claims filed? (3 Marks)

Given:- $P(X = 2) = 3 \times P(X = 4)$

$X \sim \text{Poisson}(\lambda)$

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\frac{\lambda^2 e^{-\lambda}}{2!} = 3 \times \frac{\lambda^4 e^{-\lambda}}{4}$$

$$\frac{\lambda^2}{2} = 3 \times \frac{\lambda^4}{24} \Rightarrow 4\lambda^2 = \lambda^4 \Rightarrow \lambda^2 = 4 \Rightarrow \lambda = 2$$

$$6 = \sqrt{\lambda} = \sqrt{2} \simeq 1.414$$

$$\simeq 1.414$$

6.) Suppose that a sample of $n = 800$ laptop fans of the same type are obtained at random from an ongoing production process in which 4% of all such fans produced are defective. Using the Poisson ~~distribution~~ approximation to the binomial, what is the probability that in such a sample, exactly 75 fans will be defective? (3 Marks)

$n = 800$ $\quad X = \text{Binomial}(n = 800, P = 0.04)$

$P = 0.04$ $\quad \lambda = nP = 800 \times 0.04 = 32$

$$P(x = 75) \simeq \frac{e^{-\lambda}\lambda^{75}}{75!} \Rightarrow \frac{32^{75}e^{-32}}{75!}$$

$$\Rightarrow \frac{1.36 \times 10^{112} \times 1.266 \times 10^{-14}}{6.72 \times 10^{109}}$$

$$\Rightarrow P(x = 75) \simeq 2.56 \times 10^{-12}$$

$\lambda = 32.$

d) At a certain industrial facility, accidents occur infrequently. It's known that the probability of an accident on any given day is 0.005 and the accidents are indipendent of each other. For a given period of 400 days, what is the probability that there:

(i) Will be an accident on only one day? (2 Marks)

$P = 0.005 \quad n = 400$

$\lambda = nP = 400 \times 0.005 = 2$

$$P(x=1) = \frac{2^1 e^{-2}}{1!} = 2e^{-2} \simeq 0.2707$$

$\simeq 0.2707$

(ii) Are at most two days with an accident? (2 Marks)

$P(x \leq 2) = P(x=0) + P(x=1) + P(x=2)$

$P(x=0) = e^{-2} \simeq 0.1353$

$P(x=1) = 2e^{-2} \simeq 0.2707$

$$P(x=2) = \frac{2^2 e^{-2}}{2!} \simeq 0.2707$$

$P(x \leq 2) \simeq 0.1353 + 0.2707 + 0.2707 = 0.6767$

$\simeq 0.677$

e) A probability distribution of the claim sizes for an auto insurance policy is given in the table below:-

Claim size	Probability
20	0.15
30	0.10
40	0.05
50	0.20
60	0.10
70	0.10
80	0.30

What Percentage of the claims is within one standard deviation of the mean claim size? (3 Marks)

$\text{Mean} = \sum(\text{Claim size} \times \text{Probability})$

$\Rightarrow (20 \times 0.15) + (30 \times 0.10) + (40 \times 0.05) + (50 \times 0.20) + (60 \times 0.10) + (70 \times 0.10) + (80 \times 0.30)$

$\text{Mean} = 55.0$

$\text{Standard deviation} = \sqrt{\sum \text{Probability} \times (\text{Claimsize} - \text{Mean})^2}$

$\sum \text{Probability} = 60 + 90 + 80 + 500 + 360 + 490 + 1920 = 3500$

$\sigma^2 = 3500 - (55)^2$

$\Rightarrow 3500 - 3025 = 475$

$\sigma = \sqrt{475}$

$\simeq 21.79$

Range:- $\mu \pm \sigma = 55 \pm 21.79 \Rightarrow [33.21, 76.79]$

$0.05 + 0.20 + 0.10 + 0.10 = 0.45$

$\Rightarrow 45\%$

f) A manufacturing company produces light bulbs, and the probability that a random chosen light bulb is defective is $P = 0.05$. Quality control inspectors test the light bulbs one by one, and they are interested in the number of light bulbs they need to test until they find the first defective bulb (hint: use geometric distribution)

i) Given the information, what is the expected number of light bulbs the inspectors need to test until they find the first defective one? (hint: $E(x)$) (2 Marks)

$E(x) = 1/p = 1/0.05$

$\Rightarrow 20$

(ii) What is the probability that the inspectors find the first defective bulb on the 3^{rd} test? (Hint: Probability Distribution; $P(x=x)$) (2 Marks)

$P(x=3) = (1-P)^{3-1} \times P = (0.95)^2 \times 0.05$

$\Rightarrow 0.045125$

(iii) What is the probability that the inspectors find the first defective bulb within the first 5 tests? (Hint: CDF; $F(x)$) (2 Marks)

$P(x \leq 5) = 1-(1-P)^5 = 1-(0.95)^5$

$1 - 0.77378$

$\Rightarrow 0.22622$

(iv) What is the probability that the inspectors will not find a defective bulb within the first 10 tests? (hint: Complement Rule) (2 Marks)

$P(x > 10)$

$= (1-P)^{10} = (0.95)^{10}$

$\Rightarrow 0.59874$

Q) A call center finds that the probability that a customer call results in a resolution on the first attempt is $P = 0.2$. The call center wants to determine the number of calls needed to successfully resolve 3 customer issues (hint: use negative Binomial Distribution)

(i) What is the probability that the 3^{rd} customer issue is resolved on the 10^{th} call? (2 Marks)

$P(x=k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$ $\quad k=10;\ r=3$

$P(x=10) = \binom{9}{2} (0.2)^3 (0.8)^7$

$$\binom{9}{2} = \frac{9 \times 8}{2} = 36$$

$$P(x = 10) = 36 \times (0.008) \times (0.2097)$$

$$\Rightarrow 36 \times 0.0016776$$

$$\Rightarrow 0.0604$$

(ii) What is the probability that the call centre resolves 3 customer issues within the first 12 calls? (2 Marks)

$P(T \leq 12)$ $r = 3$

$$C_p = \sum_{k=3}^{12} \binom{k-1}{3-1} p^3 (1-p)^{k-3}$$

$$P(T \leq 12) \simeq 0.998$$

(iii) What is the variance in the number of calls needed to resolve 3 customer issues? (2 Marks)

$$\text{Variance} = \frac{r(1-p)}{p^2}$$

$$= \frac{3 \times 0.8}{(0.2)^2} \Rightarrow \frac{2.4}{0.04}$$

$$\Rightarrow 60$$