DAVID WAFULA MARUTI

Q1

SCT - 221 - D1 - 0045/2023

CAT 1

STA 2200 : PROBABILITY AND STATISTICS II

a)

i) Continous random Variables

This is a variable whose value can take on any value within a given range or interval unlike discrete random variables which can only take on speficic, seperate values

example

Height :- The height of a student in a class is a continous random variable it could be 175.2cm, 175.234cm and so on

ii) Expectation

Also expected value of a random variable is the average value you would expect to obtain if you repeated the experiment many times. Its a weighted average

example.

if yo have a fair six sided die, the expectation is 3.5. each outcome (1, 2, 3, 4, 5, 6) has a probability of $\frac{1}{6}$, and sum of $(1\frac{1}{6}) + (2\frac{1}{6}) + (3\frac{1}{6}) + (4\frac{1}{6}) + (5\frac{1}{6}) + (6\frac{1}{6}) = 3.5$.

b) Define the experiment

- Toss a fair coin once

Define the sample space

- The possible outcomes are

S = Heads (H) or Tails (T)

Define the random variable

This is a function that assigns a real number to each outcome in the sample space.

let X be a random variable such that

$X(H) = 1, \ X(T) = 0$

This means

If the coin lands on Heads, $X = 1$

if the coin lands on Tails, $X = 0$

Probability of Distribution of X

Since the coin is fair

$P(X=1) = P(H) = 0.5$

$P(X=0) = P(T) = 0.5$

x	pobability
0	0.5
1	0.5

∴ The outcome of the coin toss can be descenbed by a random Vanable X that maps each possible outcome (head/tail) to a number (1/0). since the result is not known in advance and follows a probability distribution, X is a random Vanable.

c) Define the random vanable X
let X be the amount of money won (or lost) based on die outcome

outcome Roll	outcome	x payoff
1	win 2 \$	X = 2
6	win 1 \$	x = 1
2-5	loose 1 \$	x = -1

Determine the probability distribution

$P(X=2) = \frac{1}{6}$ (Probability of rolling a 1.)

$P(X=1) = \frac{1}{6}$ (" " " 6)

$P(X=-1) = \frac{1}{6}$ (Probability of rolling a 2, 3, 4, or 5 $= \frac{4}{6} = \frac{2}{3}$

Calculate the expected value $E(x)$

⇒ Sum of each outcome multiplied by its probability

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

$$E(x) = \frac{2}{6} + \frac{1}{6} - \frac{4}{6} = \frac{-1}{6} \simeq -0.167$$

Expected Value

$$E(x) = -\frac{1}{6} \simeq -0.167$$

This means that on average, you will expect to loose $\frac{1}{6}$\$ per toss of the die.

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

Find the mean and median distribution. Also find $P(X > 0.5)$ and $Var(X)$

1. Mean ($E[x]$)

before we check that $f(x)$ is a valid pdf

Verify that

$$\int_0^1 f(x)dx = 1$$

$$E[x] = \int_0^1 x(3.6x - 2.4x^2)dx = \int_0^1 (3.6x^2 - 2.4x^3)dx$$

$$E[x] = [1.2x^3 - 0.6x^4]_0^1$$

$$E[x] = (1.2(1)^3 - 0.6(1)^4) - (1.2(0)^3 - 0.6(0)^4) = 1.2 - 0.6 = 0.6$$

$$\text{mean} E([x]) = 0.6$$

2. Median

The median m is the value such that $P(X \leq m) = 0.5$

$$\int_0^m (3.6x - 2.4x^2)dx = 0.5$$

$$[1.2x^2 - 0.8x^3]_0^m = 0.5$$

solve for m

$$1.2m^2 - 0.8m^3 = 0.5$$

Solving using numerical methods

if $m = 0.5$: $1.2(0.5)^2 - 0.8(0.5)^3 = 0.3 - 0.1 = 0.2$ (too low)

if $m = 0.7$: $1.2(0.7)^2 - 0.8(0.7)^3 \approx 0.588 - 0.2744 \approx 0.3136$ (too low)

if $m = 0.8$: $1.2(0.8)^2 - 0.8(0.8)^3 \approx 0.768 - 0.4096 \approx 0.3584$ (too low)

the median is ≈ 0.7

3. $P(X > 0.5)$

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

$$[1.2x^2 - 0.8x^3]_{0.5}^1 = (1.2 - 0.8) - (1.2(0.5)^2 - 0.8(0.5)^3) = 0.4 - (0.3 - 0.1) = 0.2$$

$$P(X > 0.5) = 0.2$$

4. Variance

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

Find $E(X^2)$

$$E[X^2] = \int_0^1 x^2(3.6x - 2.4x^2)dx = \int_0^1 (3.6x^3 - 2.4x^4)dx = [0.9x^4 - 0.48x^5]_0^1 = 0.9 - 0.48 = 0.42$$

$$\text{Variance } (Var(x)) = 0.06$$

Question Two

Q2 Solution

a) let x be no of passengers who show up.
tickets sold = 110
following binomial distribution

$$X \sim \text{Binomial}(n = 110, p = 0.95)$$

where $n = 110$ (number of trials
$p = 0.95$ probability of success

$\therefore$ $P(X > 100) = 1 - P(X \leq 100)$ use normal approximation to the distribution (binomial)

$$\mu = np = 110 \times 0.95 = 104.5$$
$$\sigma = \sqrt{np(1-p)} = \sqrt{110 \times 0.95 \times 0.05} \approx 3.24$$

$$X \sim N(\mu = 104.5, \sigma = 3.24)$$

Continuity Correction calculate
$P(X \leq 100.5)$
Standardize values

$$z = \frac{X - \mu}{\sigma} \quad \text{for } X = 100.5$$

$$z = \frac{100.5 - 104.5}{3.24} \approx -1.23$$ use z table to find probability

Cummulative probability for $Z = -1.23$
$P(Z \leq -1.23) \approx 0.1093$
$P(X > 100) = 1 - P(X \leq 100.5) = 1 - 0.1093 = 0.8907$

$= 89.07\%$ probability of pass showing up than seats available.

Solution

Q2

b) From Poisson Distribution

$P(X=K) = \lambda Ke - \lambda K!$

$P(X=K) = K\lambda Ke - \lambda$

where KK is the number of claims
$\lambda\lambda$ is the mean number of claims.

$\Rightarrow P(X=2) = 3 \times P(X=4)$

$P(X=2) = 3 \times P(X=4)$

using poisson distribution $P(X=K)\,P(X=K)$

$\lambda 2e - \lambda 2! = 3 \times \lambda 4e - \lambda 4!$

$2!\, \lambda 2e - \lambda = 3 \times 4!\, \lambda 4e - \lambda$

Simplify

$\lambda 2e - \lambda 2 = 3 \times \lambda 4e - \lambda 24$

$2\lambda 2e - \lambda = 3 \times 24 \lambda 4e - \lambda$

Cancel common factors $e - \lambda e - \lambda$

$= \lambda 22 = 3 \times \lambda \, \cancel{4(24)}^{424}$ $= \lambda 22 = \lambda 48$

$2\lambda 2 = 3 \times 24 \lambda 4$ $\quad 2\lambda 2 = 3 \times 24 \lambda 4 = 2\lambda 2 = 8\lambda 4$

$= \lambda 22 = \lambda 48$ multiply by 8 both sides

$2\lambda 2 = 8\lambda 4$

$4\lambda 2 = \lambda 4$ Rearrange

$4\lambda 2 = \lambda 4$

$\lambda 4 - 4\lambda 2 = 0$ Factorizing

$\lambda 4 - 4\lambda 2 = 0$

$\lambda 2(\lambda 2 - 4) = 0$ Solutions $\lambda 2 = 0$ or $\lambda 2 = 4$

$\lambda 2(\lambda 2 - 4) = 0$ $\quad \lambda 2 = 0$ or $\lambda 2 = 4$

Since $\lambda = 0\ \lambda = 0$ we ignore

$\lambda 2 = 4 \Rightarrow \lambda = 2$ calculate standard deviation

$\lambda 2 = 4 \Rightarrow \lambda = 2$

$\lambda\lambda$ is equals to the variance and S.D is the squaroot of the variance

$\sigma = \lambda = 2$

$\sigma = \lambda$

$= 2$

the S.D of the number of claims is 22 or appx 1.41.

Q2

c)

let $X \sim Binomial(n = 800, p = 0.04)$

X number of defective fans   poisson approx is appropriate when

n is large

p is small

$\lambda = np$ is moderate

$\Rightarrow n = 800$

$p = 0.04$

$\lambda = np = 800 \times 0.04 = 32$

$X \sim Poission(\lambda = 32)$

Poisson distribution mass function

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

$$P(X = 75) = \frac{e^{-32} \, 32^{75}}{75!}$$ the value is small $\approx 3.93 \times 10^{-11}$

$\therefore$ Very unlikely to have defective fans

Q2 let

d) Probability of accidents any day $p = 0.005$

Numbe of days $n = 400$

Using binomial distribution

$X \sim Binomial(n = 400, p = 0.005)$

n is large and p is small we use poisson approximation

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

$\approx X \sim Posson(\lambda = 2)$

(i) Probability of accident on only one day $P(X = 1)$

using the poisson PMF

$$P(X = 1) = \frac{e^{-2} \cdot 2^1}{1!}$$

$$P(X = 1) \approx 0.2707$$

Q2

d Solution

ii) Probability of accidents on at most two days $P(x \le 2)$

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

where

$P(x=0) = \frac{e^{-2} \cdot 2^0}{0!}$ $P(x=1) = \frac{e^{-2} \cdot 2^1}{1!}$ $P(x=2) = \frac{e^{-2} \cdot 2^2}{2!}$

$= 0.1353 + 0.2707 + 0.2707 = 0.6767$

Q2 Solution

e) Mean μ of the distribution

$$\mu = \sum [x \cdot P(x)]$$

$\mu = \sum [x_i \times P(x_i)] = (20 \times 0.15) + 30 \times 0.10) + (40 \times 0.05) + (50 \times 0.20) + (60 \times 0.10) + (70 \times 0.10) + (80 \times 0.30) = 54$

Calculate the variance σ^2 of the claim size

$$\sigma^2 = \sum [(x_i - \mu)^2 \times P(x_i)]$$

$= (20-54)^2 \times 0.15 + (30-54)^2 \times 0.10 + (40-54)^2 \times 0.05 + (50-54)^2 \times 0.20 + (60-54)^2 \times 0.10 + (70-54)^2 \times 0.10 + (80-54)^2 \times 0.30 = 424$

Standard deviation σ

$\sigma = \sqrt{\sigma^2} = \sqrt{424} \approx 20.59$

Determine the range within one standard deviation of the mean

lower bound $= \mu - \sigma = 54 - 20.59 = 33.41$

upper bound $= \mu + \sigma = 54 + 20.59 = 74.59$

Identify claim sizes within the range $[33.41, 74.59]$

these are 40, 50, 60, 70

Total probability of claims within range

$P(33.41 < X \le 74.59) = P(x=40) + P(x=50) + P(x=60) + P(x=70)$

$= 0.05 + 0.20 + 0.10 + 0.10 = 0.45$

$0.45 \times 100\% = 45\%$

Q2

f) Using geometric distribution parameter $p = 0.05$. The expected value $E(x)$ is given by $1/p$

i) $E(x) = 1/p = 1/0.05 = 20$

ii) Probability mass function (PMF) of a geometric distribution is given by $P(x = x) = (1-p)^{(x-1)} \times p$ where x is the number of trials until the first success finding of a defective bulb.

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

$= 0.045125$

The third test to find a defective bulb, the prob. is appox 0.045125.

iii) The cummulative distribution function CDF of a geometric distribution is given by $F(x) = 1 - (1-p)^x$. This represents the probability of finding the first defective bulb within the first x tests

⇒ Calculate the probability of finding the first defective bulb within the first 5 tests is

$F(5) = 1 - (1-0.05)^5 = 1 - (0.95)^5 = 1 - 0.77338 \approx$

$= 0.22622$

iv) The probability of not finding a defective bulb within the first 10 tests is given by the complement rule

$P(x > 10) = (1-p)^{10}$

$P(X \geq 10) = (1-0.05)^{10} = (0.95)^{10} \simeq 0.5987$

$= 0.5987$.

Q2

g Negative Binomial Distribution

Models the number of trials needed to achieve a fixed number of successes.

The PMF is given by

$$P(X = k) = C(k-1, r-1) \times p^r \times (1-p)^{k-r}$$

where

X is the random variable representing the number of trials.

k is the number of trials until r successes are achieved

r is number of successes

lets assume $P = 0.2 (20\%)$

P is the probability of success on a single trial (0.2)

The negative binomial distribution describes the probability of k

i) failures before r success. in this case

$r = 3$ (number of successes, resolved issues)

$k = 7$ (Number of failures before 3rd success, 10 calls total)

$(x-1 \text{ choose } r-1)$ is the binomial co-efficient, calculated as

$$\frac{(x-1)!}{((r-1)! \times (x-r)!)}$$

$(10-1) \text{ choose } 3-1) = (9 \text{ Choose } 2) = 36$ substitute

$P(X = 10) = 36 \times (0.2)^3 \times (0.8)^7$

calculate probability

$P(X = 10) \approx 36 \times 0.008 \times 0.2097152 \approx 0.0604$

$=$ Approximately 0.0604, assuming $P = 0.2$

ii) Probability that the call centre resolves 3 customers issues within the first 12 calls ?

assume probability for $x = 3, 4, 5, \ldots, 12$

CDF for the negative binomial distribution is not easy to express using statistical

iii) The variance of a negative binomial distribution is given by

$$\text{Variance} = r \times \frac{(1-p)}{p^2} \quad \text{where } r = 3, \; p = 0.2$$

$$\text{Variance} = 3 \times \frac{(1-0.2)}{(0.2)^2}$$

$$= 3 \times \frac{0.8}{0.04}$$

$= 60$ assuming $P = 0.2$