2016

Student ID: ....

Question 5. DISCRETE RANDOM VARIABLES [15 marks]

(a) [4 marks] The discrete random variable, X, takes on the values 0, 1, 2, 3, 4 with the following probabilities: $P(X=0)=\frac{1}{16}$, $P(X=1)=\frac{1}{16}$, $P(X=2)=\frac{1}{8}$, $P(X=3)=\frac{1}{8}$, $P(X=4)=\frac{5}{8}$.

(i) Find the probability that X is at least 1.

$$P(X \geq 1) = \frac{1}{16}+\frac{1}{8}+\frac{1}{8}+\frac{5}{8} = \frac{15}{16}$$

(ii) Find $E(X)$

$$E(X) = \sum_{\text{all } x} x P(X=x) = 0\left(\frac{1}{16}\right)+1\left(\frac{1}{16}\right)+2\left(\frac{1}{8}\right)+3\left(\frac{1}{8}\right)+4\left(\frac{5}{8}\right)$$

$$= \frac{51}{16}$$

(iii) Find $Var(X)$

$$E(X^2) = \sum_{\text{all } x} x^2 P(X=x) = 0^2\left(\frac{1}{16}\right)+1^2\left(\frac{1}{16}\right)+2^2\left(\frac{1}{8}\right)+3^2\left(\frac{1}{8}\right)+4^2\left(\frac{5}{8}\right)$$

$$= \frac{187}{16}$$

$$Var(X) = \frac{187}{16} - \left(\frac{51}{16}\right)^2$$

(b) [3 marks]

Let the random variable, X, count the number of nanotubes in a batch of 50 that have structural defects. What are the assumptions needed for X to be well modelled as a binomial random variable? Are these likely to hold?

- counting no. of defects } GOOD FIT TO BINOMIAL
- fixed no. trials (50) }
- binary (defect or not) – REASONABLE
- const. prob of success – MAYBE

Possible Problem

- independence – might have clusters of defect tubes

(c) [3 marks]

Let the random variable, X, be a Poisson random variable with parameter 5. What is the probability $P(X = 5)$?

$$P(x=5) = \frac{e^{-5} 5^{5}}{5!}$$

(d) [3 marks] If X is a binomial random variable with parameters $n = 8$ and $p = 0.5$, what are the mean and variance of Y and Z where $Y = -3X + 6$ and $Z = 5(5 + X)$?

$E(X) = np = 4 \quad Var(X) = np(1-p) = 2$

$E(Y) = -3(4) + 6 = -6$

$Var(Y) = (-3)^2 2 = 18$

$E(Z) = 25 + 5(4) = 45$

$Var(Z) = 5^2(2) = 50$

(e) [2 marks] The Poisson random variable, X ($X \sim \text{Po}(\mu)$) is designed to model the number of active transmitters operating in the 2 MHz to 4 MHz range. What is the distribution of the number of active transmitters operating in the 1.5 MHz to 2.5 MHz range? Explain your assumptions.

$Po\left(\frac{\mu}{2}\right)$

transmitters arrive in the 1.5MHz – 4MHz band singly, independently and at constant rate.

Question 6. CONTINUOUS RANDOM VARIABLES [15 marks]

(a) [8 marks] Let the random variable X have the probability density function, $f(x)$, where:

$$f(x) = A + x, \quad 0 \leq x \leq 1, \tag{1}$$

and $f(x) = 0$ for $x < 0$ and $x > 1$. Answer the following questions:

(i) Find A.

$$\int_0^1 A + x \, dx = \left[Ax + \frac{x^2}{2}\right]_0^1 = A + \frac{1}{2} = 1 \implies A = 1/2$$

(ii) Find $P(X > 0.8)$.

$$\int_{0.8}^1 \frac{1}{2} + x \, dx = \left[\frac{x}{2} + \frac{x^2}{2}\right]_{0.8}^1 = \left(\frac{1}{2} + \frac{1}{2}\right) - \left(\frac{0.8}{2} + \frac{0.64}{2}\right)$$

$$= 1 - 0.4 - 0.32$$

$$= 0.28$$

(iii) Find $E(X)$.

$$\int_0^1 x\left(\frac{1}{2} + x\right) dx = \int_0^1 \frac{x}{2} + x^2 \, dx = \left[\frac{x^2}{4} + \frac{x^3}{3}\right]_0^1$$

$$= \frac{1}{4} + \frac{1}{3} = \frac{7}{12}$$

(b) [4 marks]

Let the random variable X have the cumulative distribution function, $F(x)$, where:

$$F(x) = 1 - \exp(-x/2), \quad x \geq 0, \tag{2}$$

reads as $e^{-x/2}$

and $F(x) = 0$ for $x < 0$. Answer the following questions:

(i) Find the probability density function of X.

$$f(x) = F'(x) = \begin{cases} \frac{1}{2} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

(ii) Find $P(2 < X < 3)$.

$$F(3) - F(2) = \left(1 - e^{-3/2}\right) - \left(1 - e^{-2/2}\right)$$

$$= e^{-1} - e^{-3/2}$$

(c) [3 marks] The continuous gamma random variable X has the probability density function:

$$f(x) = \frac{1}{r!} x^r \exp(-x), \quad x \geq 0, \tag{3}$$

and $f(x) = 0$ for $x < 0$, where r is a positive integer. Find the variance of X using the following integral result:

$$\int_0^{\infty} x^n \exp(-x) dx = n!$$

where $n!$ is the factorial of n which is a positive integer.

$$E(X) = \int_0^{\infty} x \frac{1}{r!} x^r e^{-x} dx = \frac{1}{r!} \int_0^{\infty} x^{r+1} e^{-x} dx = \frac{(r+1)!}{r!} = r+1$$

$$E(X^2) = \int_0^{\infty} x^2 \frac{1}{r!} x^r e^{-x} dx = \frac{1}{r!} \int_0^{\infty} x^{r+2} e^{-x} dx = \frac{(r+2)!}{r!} = (r+1)(r+2)$$

$$Var(X) = (r+1)(r+2) - (r+1)^2 = (r+1)\left[r+2-(r+1)\right]$$

$$= r+1$$

Question 7. CONFIDENCE INTERVALS [9 marks]

(a) [6 marks] Let X be a binomial random variable with parameters n and p.

(i) Prove that $E(\hat{p}) = p$, where $\hat{p} = X/n$ is the proportion of successes in the n trials.

$$E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{np}{n} = p$$

(ii) Prove that $Var(\hat{p}) = \frac{p(1-p)}{n}$.

$$Var(\hat{p}) = Var\left(\frac{X}{n}\right) = \frac{1}{n^2}Var(X) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

(iii) Explain the significance of $E(\hat{p}) = p$ and $Var(\hat{p}) = \frac{p(1-p)}{n}$ in terms of estimation.

$$\left.\begin{array}{l} \text{unbiased } (E(\hat{p}) = p) \\ Var(\hat{p}) \to 0 \text{ as } n \to \infty \end{array}\right\} \text{together} \Rightarrow \text{consistent}$$

(b) [2 marks] Give the form of a 95% confidence interval for p based on the sample proportion $\hat{p} = X/n$.

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

(c) [1 mark] Explain why the results in Question 7(b) above show that the empirical cumulative distribution function (ECDF) is a reasonable estimator of the true CDF of a random variable.

$F(x) = P(X \leq x)$

If you take a SRS and count the no. of values less than or equal to x then

$$ECDF(x) = \frac{\text{no. values} \leq x}{n} = \hat{p} \quad \text{and}$$

$$E[\hat{p}] = p = P(X \leq x)$$

End of Questions