Learn DU

SOLVED PYQ

Paper: Intermediate Statistics for Economics

Course: B. A.(Hons.) Economics SEM-1

Year: 2023-24

Unique Paper Code : 2272101103

Paper : DSC-3 : Intermediate Statistics for Economics | Semester–1

Course : B.A. (Hons.) Economics I Year | Dec./Jan.

Duration : 3 Hours | aximum Marks : 90 | 2023–24

PART A

(*Question No. 1 compulsory. Attempt any two questions from 2, 3 and 4*)

Q. 1. (a) The first four deviations from the mean in a sample of 5 reaction times are 0.3, 0.9, 1.0 and 1.3 Answer the questions that follow

(i) What is the deviation from the sample mean for the fifth observation?

(ii) Calculate the sample standard deviation. If each observation is multiplied by 2. What is the new variance?

(iii) What is the degree of freedom for sample standard deviation and why is it not equal to the number of observations in the sample? 1+2+2

Ans. (i)

$$(x_1 - \bar{x}) = 0.3, \quad (x_2 - \bar{x}) = 0.9, \quad (x_3 - \bar{x}) = 1.0$$

$$(x_4 - \bar{x}) = 1.3$$

$$\sum_{i=1}^{5} (x_i - \bar{x}) = 0$$

$$\Rightarrow 0.3 + 0.9 + 1 + 1.3 + (x_5 - \bar{x}) = 0$$

$$\Rightarrow \quad (x_5 - \bar{x}) = -3.50$$

(ii) Sample S.D. $= \sqrt{\dfrac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\dfrac{(0.3)^2 + (0.9)^2 + (1)^2 + (1.3)^2 + (-2.5)^2}{4}}$

$$= \sqrt{3.96} = 1.99$$

$$\text{Var}(aX) = 4\,V(X) = 4 \times 3.96 = 15.84$$

(iii) d.o.f = $(n-1)$ and not n. Although S^2 is based ion the n quantities $(x_1 - \bar{x})$, $(x_2 - \bar{x})$, $(x_3 - \bar{x})$ ... $(x_n - \bar{x})$

$\Rightarrow \sum_{i=1}^{n} (x_i - \bar{x}) = 0$, specifiying the values of any $(n-1)$ of the quantities determine the meanaining valuea. Moreover division by n underestimates σ^2 as $\Sigma(x_i - \bar{x})^2 <$ $\Sigma(\quad - \mu)$, ∴ reducing the denominator corrects for this.

(b) Consider the following data and answer the questions that follow:

Class Interval	0–5	5–10	10–15	15–20	20–30	30–40	40–60	60–90
Relative Frequency	0.177	0.166	0.175	0.136	0.194	0.078	0.044	0.030

(i) Identify the class interval in which the sample median would lie.

(ii) Calculate the density for each class interval

(iii) What proportion of observations are between 25 and 45? 1+2+2

Ans.

C.I.	Re. Frequency	C.f.	Class width	Density
0 – 5	0.177	0.177	5	0.0354
5 – 10	0.166	0.343	5	0.0332
10 – 15	0.175	0.518	5	0.035
15 – 20	0.136	0.654	5	0.0272
20 – 30	0.194	0.848	10	0.0194
30 – 40	0.078	0.926	10	0.0078
40 – 60	0.044	0.970	20	0.0022
60 – 90	0.030	1	30	0.0010

(i) Around 52% of observations are covered upto CI 10–15. Thus median should lie in the CI 10–15.

(ii)

$$\text{Rel. freq.} = (\text{Class width}) \times (\text{density})$$

$$\Rightarrow \quad \text{Density} = \frac{\text{Rel. frequency}}{\text{Class width}}$$

(iii) Splitting 20–30 and 40–60 into 20–25 and 25–30. 40–45, 45–50. 50–55 and 55–60, proportion of observations between 20–25 = 0.097

Prob. of obs. between 40–45 = 0.011

Prob. of obs. between 25–45 = 0.097 + 0.078 + 0.011

= 0.186 ≈ 18.6%

Q..2. (a) The students A_1, A_2 and A_3 go to college on any given day with probability $P(A_i)$, where $i = 1, 2$ and 3. Suppose that the event of A_1 going to college is independent of A_2 going to college on any given day, $P(A_1 \cap A_2 \cap A_3) = 0.04$, $P(A_3 / A_1, \cap A_2) = 0.25$, and $P(A_2) = 4P(A_1)$.

(i) If the probability of all three students not coming to college on any given day is 0.06, what is the probability that at least one of them will come to college on that day?

(ii) Evaluate $P(A_1 \cup A_2)$ and interpret it

(iii) If A_2 has come to college on any given day, what is the probability that A_1 and A_3 will also come to college on that day? 1+2+2

Ans. (a)

$$P(A_1 \cap A_2 \cap A_3) = 0.04$$

$$P(A_3/A_1 \cap A_2) = 0.25$$

$$P(A_2) = 4\,P(A_1)$$

$$P(A_3/A_1 \cap A_2) = \frac{P(A_3 \cap A_1 \cap A_2)}{P(A_1 \cap A_2)} = \frac{0.04}{P(A_1 \cap A_2)} = 0.25$$

$$\Rightarrow \quad P(A_1 \cap A_2) = \frac{0.04}{0.25} = 0.16$$

$$P(A_1) \cdot P(A_2) = 0.16$$

$$\Rightarrow \quad P(A_1) \cdot 4P(A_1) = 0.16$$

$$\Rightarrow \quad 4(P(A_1))^2 = 0.16$$

$$\Rightarrow \quad (P(A_1))^2 = 0.04$$

$$\Rightarrow \quad P(A_1) = 0.20$$

(i) $P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3) = 0.06$

P(at least one of them will come to college) =

$= 1 - P$(none will come to college) $= 1 - 0.06 =$ **0.94**

(ii) $P(A_2) = 4 \cdot P(A_1) = 0.80$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = 0.20 + 0.80 - 0.16$$

$$= \mathbf{0.84}$$

This is the probability that either only student 1, only student 2 or both student 1 and 2 will go to college or at least one of A_1 and A_2 will go to college.

(iii) $P(A_1 \cap A_3 / A_2) = \dfrac{P(A_1 \cap A_3 \cap A_2)}{P(A_2)} = \dfrac{0.04}{0.80} = 0.05$

(b) A bookstore purchases three copies of a book at ₹ 6.00 each and sells each at ₹ 12.00 each. Unsold copies are returned for ₹ 2.00. The PMF of X is given as follows:

X	0	1	2	3
$P(x)$	0.1	0.2	0.2	0.5

Find the

(i) PMF of the net revenue function Y.

(ii) Expected value and variance of X.

(iii) Expected value and variance of the net revenue Y.

(iv) Find $P(Y \geq 8)$ **10**

Ans. Net Revenue $= 12X - 18 + 2(3 - X) = 10X - 12$

Net Revenue $= \{10X - 12, \quad x = 0, 1, 2, 3$

(i) $Y =$ Net Revenue

x	$p(x)$	$h(x)$
0	0.1	12
1	0.2	– 2
2	0.2	8
3	0.5	18

(ii) $\in(X) = 0(0.1) + 1(0.2) + 2(0.2) + 3(0.5) = 2.1$

$$\in(X^2) = 0 + 0.2 + 0.8 + 4.5 = 5.5$$

$$V(X) = 5.5 - (2.1)^2 = 1.09$$

(iii)

$$\in h(x) = 10\,E(X) - 12 \Rightarrow 10 \times 2.1 - 12 = 9$$

$$V\,h(X) = V(Y) = V(10X - 12) = 100\,V(X)$$

$$= 100 \times 1.09 = 109$$

(iv) $P(Y \geq 8)$

X	Y	$f(Y)$
0	– 12	0.1
1	– 2	0.2
2	8	0.2
3	18	0.5

$\Rightarrow \quad P(Y \geq 8) = 0.7$

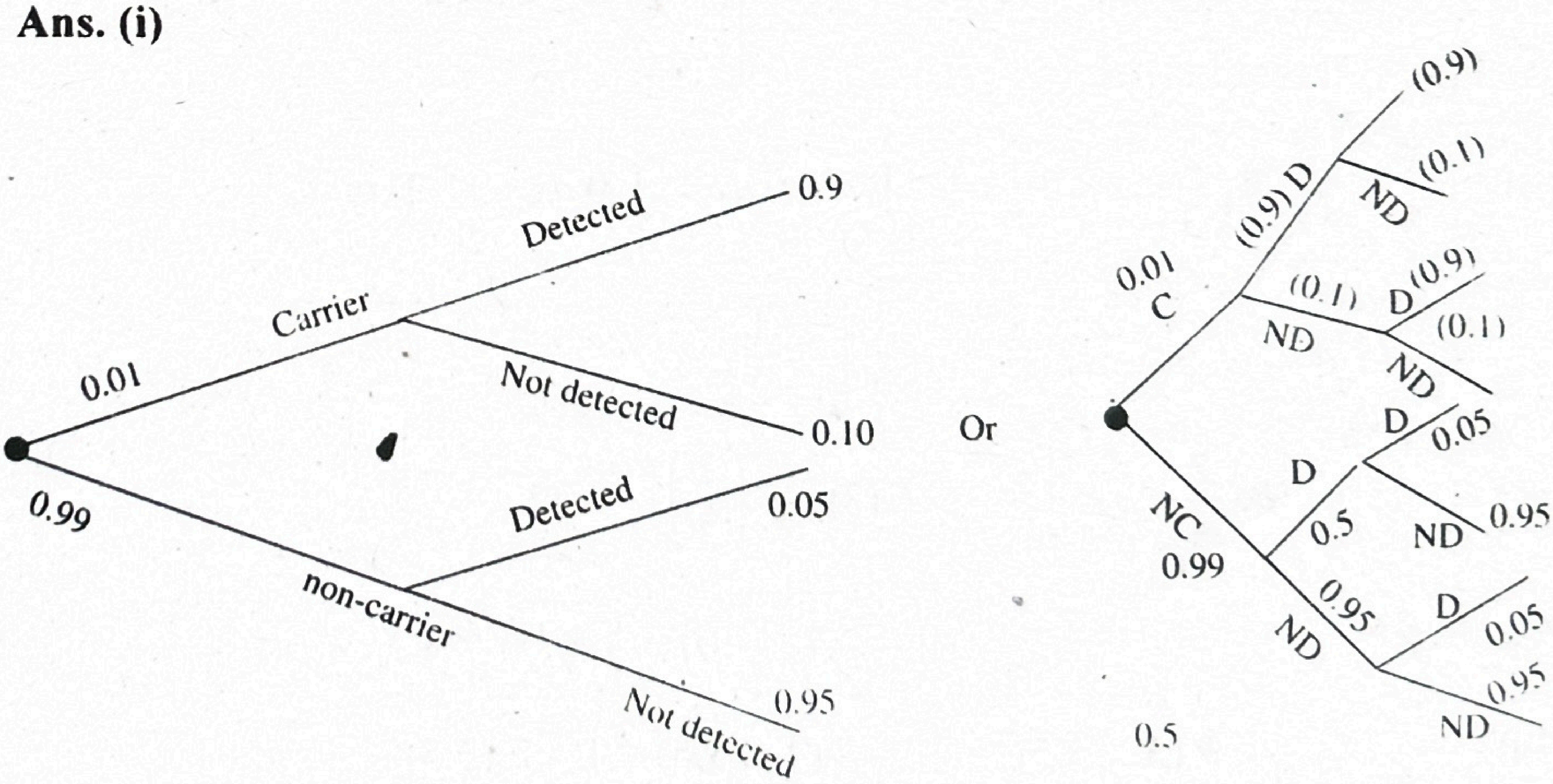
Q. 3. (a) One per cent of all individuals in a certain population are carriers of a particular disease. A diagnostic test for this disease has a 90 per cent detection rate for carriers and a 5% detection rate for noncarriers. Suppose the test is applied independently to two different blood samples from the same randomly selected individual.

(i) Draw the tree diagram for the question.

(ii) What is the probability that both tests yield the same result?

(iii) If both tests are positive, what is the probability that the selected individual is a carrier? **1+2+2**

Ans. (i)



(ii) Let events B, C and BP denote the events that both tests yield the same result, tested person being a carrier and both tests yield a positive result.

$$P(B) = P(B/C) \cdot P(C) + P(B/C') \cdot P(C')$$

$$= (0.01)^2\,(0.1^2 + 0.9^2) + 0.99\,(0.05^2 + 0.95^2)$$

$$= \mathbf{0.9042}$$

(iii)

$$P(BP) = P(BP/C) \cdot P(C) + P(BP/C'') \cdot P(C')$$

$$= (0.9)^2\,(0.01) + (0.05)^2\,(0.99) = 0.0106$$

$$P(C/BP) = \frac{P(BP/C)\cdot P(C)}{P(BP)} = \frac{(0.9)^2\,(0.01)}{0.0106} = \mathbf{0.7660}$$

(b) Let X be a random variable with cdf

$$F(x) = \begin{cases} 0, & x < 2 \\ \dfrac{(x-2)}{2}, & 2 \le x \le 4 \\ 1, & x \ge 4 \end{cases}$$

(i) Find the pdf of X.

(ii) Ftnd $P(2/3 < X < 3)$.

(iii) Find $P(X > 3.5)$.

(iv) Find the 60th percentile.

(v) Find $P(X = 3)$. **2×5=10**

Ans.

$$F(x) = \begin{cases} 0, & x < 2 \\ \dfrac{(x-2)}{2}, & 2 \le x \le 4 \\ 1, & x \ge 4 \end{cases}$$

(i)

$$f(x) = F^{1}(x) = \begin{cases} 0, & x < 2 \\ (1/2) & 2 \le x \le 4 \\ 0 & x > 4 \end{cases}$$

(ii)

$$P\left(\frac{2}{3} < x < 3\right) = \int_{2/3}^{3} f(x)\,dx = \int_{2/3}^{2} 0\,dx + \int_{2}^{3} 1/2\,dx$$

$$= \left[\frac{x}{2}\right]_2^3 = \frac{1}{2}$$

or

$$P\left(\frac{2}{3} < x < 3\right) = F(3) - F\left(\frac{2}{3}\right)$$

$$= \left(\frac{3-2}{2}\right) - (0) = \frac{1}{2}$$

(iii)

$$P(X > 3.5) = 1 - P(X \le 3.5) = 1 - F(3.5)$$

$$= 1 - \left(\frac{3.5-2}{2}\right) = \frac{0.5}{2} = 0.25$$

or

$$1 - \int_{0}^{3.5} f(x)\,dx = 1 - \left[\int_{0}^{2} f(x)\,dx + \int_{2}^{3.5} f(x)\,dx\right]$$

$$= 1 - \left[0 + \left[\frac{x}{2}\right]_2^{3.5}\right] = \frac{0.5}{2} = 0.25$$

(iv) Let 60^{th} percentile be η (0.6)

$$F[\eta(p)] = p$$

$$F(\eta(0.6)) = 0.6$$

$$\Rightarrow \frac{\eta - 2}{2} = 0.6 \Rightarrow \eta = 3.2$$

(v) $P(X = 3) = 0$ because for a continuous dist. $P(X = C) = 0$ for any constant C.

Q. 4. (a) Suppose that Delhi witnesses only two types of days, rainy or sunny. The probability that a rainy day is followed by a rainy day is 0.8 and the probability that a sunny day is followed by a rainy day is 0.6. Find the probability that a rainy day is followed by:

- **(i) A rainy day, a sunny day, and another rainy day;**
- **(ii) Two sunny days and then a rainy day;**
- **(iii) No sunny day for three consecutive days;**
- **(iv) Rain two days later** **1+1+1+2**

Ans. (a)

		Today: Rainy	Today: Sunny
Tomorrow	Rainy	0.8	0.6
	Sunny	0.2	0.4

(i) P (2^{nd} day is rainy/1^{st} day is rainy) × P (3^{rd} day is sunny/2^{nd} day is rainy) × P (4^{th} day is rainy/3^{rd} day is sunny).

$$= 0.8 \times 0.2 \times 0.6$$

$$= 0.096$$

(ii) P (2^{nd} day is sunny/1^{st} day is rainy) × P (3^{rd} day is sunny/2^{nd} day is sunny) × P (4^{th} day is rainy/3^{rd} day is sunny)

$$= 0.2 \times 0.4 \times 0.6 = 0.048$$

(iii) No sunny day for 3 consecutive days = P (2^{nd} day is rainy/1^{st} day is rainy) × P (3^{rd} day is rainy/2^{nd} day is rainy) × P (4^{th} day is rainy/3^{rd} day is rainy).

$$= 0.8 \times 0.8 \times 0.8 = 0.512$$

(iv) Rain two days late

= P (2^{nd} day is sunny/1^{st} day is rainy) × P (3^{rd} day is sunny/2^{nd} day is sunny) × P (4^{th} day is rainy/3^{rd} day is sunny)

$$= 0.2 \times 0.4 \times 0.6 = 0.048$$

(b) (i) The time T, in days, required for the completion of a contracted project is a random variable with PDF

$$f_T(t) = 0.1\, e^{-0.1t} \text{ for } t > 0 \text{ and } 0 \text{ otherwise}$$

Suppose the contracted project must be completed in 15 days. If $T < 15$ there is a cost of ₹ $5(15 - T)$ and if $T > 15$ there is a cost of ₹ $10(T - 15)$. Find the expected value of the cost.

(ii) Consider the experiment where product items are being inspected for the presence of a particular defect until the first defective product item is found. Let X denote the total number of items inspected. Suppose a product item is defective with probability $p, p > 0$, independently of other product items. Find $E(X)$ in terms of p when pmf is given as $p(x) = (1 - p)^{x-1} p$. **5+5**

Ans. Let $h(t) = \begin{cases} 5(15-t), & t < 15 \\ 10(t-15), & t > 15 \end{cases}$

$$E(h(t)) = \int_{-\infty}^{+\infty} h(t) f_T \, dt = \int_0^{15} h(t) f_T \, dt + \int_{15}^{\infty} h(t) f_T \, dt$$

$$= \int_0^{15} 5(15-t)\, 0.1 e^{-0.1t} dt + \int_{15}^{\infty} 10(t-15)\, 0.1 e^{-0.1t} \, dt$$

$$= 75 \int_0^{15} 0.1 e^{-0.1t} dt + 5 \int_0^{15} 0.1 t e^{-0.1t} dt + 10 \int_{15}^{\infty} 0.1 t e^{-0.1t} dt$$

$$+ 150 \int_{15}^{\infty} 0.1 e^{-0.1t} \, dt$$

$$= 75\left[-e^{-0.1t}\right]_0^{15} + 5\left[-t e^{-0.1t}\right]_0^{15} - 5\left[\frac{1}{0.1} e^{-0.1t}\right]_0^{15}$$

$$+ 10\left[-t e^{-0.1t}\right]_{15}^{\infty} - 10\left[\frac{1}{0.1} e^{-0.1t}\right]_{15}^{\infty} + 150\left[-e^{-0.1t}\right]_{15}^{\infty}$$

$$= 58.46$$

(ii) $p(x) = \begin{cases} (1-p)^{x-1} p & x = 1, 2, 3 \ldots \\ 0 & \text{otherwise} \end{cases}$

$$\in (X) = \sum_{x=1}^{\infty} x(1-p)^{x-1} p$$

$$= p[1 + 2(1-p) + 3(1-p)^2 + \ldots]$$

$$= p\left[\frac{1}{1-(1-p)} + \frac{1-p}{(1-1+p)^2}\right] \qquad \left[\because \frac{a}{1-r} + \frac{dr}{(1-r)^2}\right]$$

$$= p\left[\frac{1}{p} + \frac{(1-p)}{p^2}\right] = 1 + \frac{1-p}{p} = \frac{1}{p}$$

Part B

(Attempt any two questions from 5, 6 and 7)

Q. 5. (a) Suppose the useful lifetime, in years, of a personal computer (PC) is exponentially distributed with parameter λ = 0-25. A student entering a four-year undergraduate program inherits a two-year-old PC from his sister who just graduated. Find the probability the useful life time of the PC the student inherited will last at least until the student graduates. **5**

Ans.

$$\lambda = 0.25, \quad X \sim \text{exp.}(\lambda)$$

$$f(x;\lambda) = \begin{cases} 0.25\, e^{-0.25x} & x \geq o \\ 0 & \text{otherwise} \end{cases}$$

Given the memory less property of exponential dist.

$$P(X \geq 4) = 1 - F(4; 0.25]$$
$$= 1 - [1 - e^{-0.25 \times 4}] = e^{-1}$$
$$= \frac{1}{e}$$

(b) In a shipment of 10,000 of a certain type of electronic component, 300 are defective. Suppose that 50 components are selected at random for inspection, and let X denote the number of defective components found.

(i) Find $F(X < 1)$ what distribution would you use?

(ii) Find $E(X)$ and $V(X)$

(iii) Find an approximation to the probability $P(X \leq 3)$, and compare it with the exact probability found in part (i) Which distribution would you use to find approximate probability? Give reasons for your answer. **3+3+4**

Ans. $n = 50; \quad N = 10{,}000; \quad M = 300; \quad N - M = 9700$

(i) $P(X \leq 3)$ = Hypergeometric distribution is the exact distribution

$$= P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$\frac{300_{c_0} \cdot 9700_{c_{50}} + 300_{c_1}\, 9700_{c_{49}} + 300_{c_2}\, 9700_{c_{48}} + 300_{c_3}\, 9700_{c_{47}}}{1000_{c_{50}}}$$

(ii) $$\in(X) = \frac{n \cdot M}{N}\ 15; \quad V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \frac{M}{N}\left(1 - \frac{M}{N}\right)$$

$$= 1.447$$

(iii) Approximate Prob. Dist. is binomial distribution provided sampling is done without replacement and $\frac{n}{N}$ is at most 0.05

$$\frac{n}{N} = \frac{50}{10{,}000} = 0.005 < 0.05$$

$$p = \frac{300}{10{,}000} = 0.03; \quad q = 0.97$$

$$P(X \leq 3) = F(3;\quad 50,\quad 0.03,\quad 0.97)$$

(c) A candy company distributes boxes of chocolates with a mixture of creams, toffees, and cordials. Suppose that the weight of each box is 1 kilogram, but the individual weights of the creams, toffees, and cordials vary from box to box. For a randomly selected box, let X and Y represent the weights of the creams and the toffees, respectively, and suppose that the joint density function of these variables is

$$f(X, Y) = \frac{1}{5}(2x + 3y),\ f\text{ or } 0\quad x \leq 1; 0 \leq y \leq 1, 0 \leq x + y \leq 1 = 0, \text{ otherwise}$$

(i) Find the probability that in a given box cordials account for more than 1/2 of the weight. What is the probability that the weight of the toffees in a box is less than 1/8 of a kilogram if it is known that creams constitute 3/4 of the weight?

(ii) Are X and Y independent? Explain using appropriate statistical measures.

5×2=10

Ans. (i)

$$P\left(X + Y \leq \frac{1}{2}\right) = \int_0^{0.5}\int_0^{0.5-x} \frac{2}{5}(2x + 3y)\, dy\, dx$$

$$= \int_0^{0.5}\left[\frac{4}{5}x\,[y]_0^{0.5-x} + \frac{6}{5}\left[\frac{y^2}{2}\right]_0^{0.5-x}\right]$$

$$= \left[\int_0^{1/2} \frac{4}{5}x(0.5 - x) + \frac{3}{5}(0.5 - x)^2\, dx\right]$$

$$= \int_0^{1/2}\left(\frac{4}{10}x - \frac{4}{5}x^2 + \frac{3}{20} + \frac{3}{5}x^2 - \frac{3x}{5}\right)dx$$

$$= \frac{4}{10}\left[\frac{x^2}{2}\right]_0^{1/2} - \frac{4}{5}\left[\frac{x^3}{3}\right]_0^{1/2} + \frac{3}{20}[x]_0^{1/2} + \frac{3}{5}\left(\frac{x^3}{3}\right)$$

$$- \frac{3}{5}\left[\frac{x^2}{2}\right]_0^{1/2} = \frac{1}{24} = 0.0416$$

$$\Rightarrow P\left(Y < \frac{1}{8} \middle/ x = \frac{3}{4}\right)$$

$$f(Y/X) = \frac{f(Y, X)}{f_x(X)} = \frac{\frac{2}{5}(2x + 3y)}{\frac{4x}{5} + \frac{3}{5}} = \frac{2(2x + 3y)}{4x + 3}$$

$$P\left(Y < \frac{1}{8} \middle/ X = \frac{3}{4}\right) = \int_0^{1/r} f(Y/X)\, dy$$

$$= \frac{2}{4x+3}\left[2\,xy + \frac{3y^2}{2}\right]_0^{1/8} = \frac{9}{128} = \mathbf{0.0703}$$

(ii)
$$f(Y/X) = \frac{2}{4x+3}(2x+3y)$$

$$f(Y) = \int_0^1 \frac{2}{5}(2x+3y)\,dx = \frac{4}{5}\left[\frac{x^2}{2}\right]_0^1 + \frac{6}{5}Y\,[x]_0^1$$

$$= \frac{2+6y}{5}$$

$f(Y/X) \neq f(Y)$. Thus, X and Y are not independent.

Q. 6. (a) An examination is frequently regarded as being good (in the sense of determining a valid grade spread for those taking it) if the lest scores of those taking the examination can be approximated by a normal density function. The instructor often uses the test scores to estimate the normal parameters p and o and then assigns the letter grade A to those whose test score is greater than $\mu + \sigma \cdot B$ to those whose score is between μ and $\mu + \sigma$, C to those whose score is between $\mu - \sigma$ and μ, D to those whose score is between $\mu - 2\sigma$ and $\mu - \sigma$, and F to those getting a score below $\mu - 2\sigma$. Determine the % of students who receive grade A, B, C and D. **5**

Ans. (a) Let X be students score on the test

$$P(A) = P(X > \mu + \sigma) = P\left(\frac{X - v}{\sigma} \geq 1\right) = 1 - P\left(\frac{x-\mu}{\sigma} < 1\right)$$

$$P(A) = 1 - \Phi(1) = 1 - 0.84134 = 0.15866$$
$$P(B) = P(\mu < X < \mu + \sigma) = \Phi(1) - \Xi(0) = 0.34134$$
$$P(C) = P(\mu - \sigma < X < \mu) = \Phi(0) - \Xi(-1) = 0.34134$$
$$P(D) = P(\mu - 2\sigma < X < \mu - \sigma) = \Phi(-1) - \Xi(-2) = 0.13591$$

% of student who receive grades A, B, C and D are 15.86, 34.13, 34.13, and 13.59 respectively.

(b) A typesetting agency used by a scientific journal employs two typesetters. Let X_1 and X_2 denote the number of errors committed by typesetter 1 and 2, respectively, when asked to typeset an article. Suppose that X_1 and X_2 are Poisson random variables with expected values 2.6 and 3.8, respectively.

(i) What is the variance of X_1 and of X_2?

(ii) Suppose that typesetter 1 handles 60% of the articles. Find the probability that the next article will have no errors.

(iii) If an article has no typesetting errors, what is the probability it was typeset by the second typesetter? **2+3+5**

Ans. $X_1 \sim$ Poi (2.6) ; $X_2 \sim$ Poi (3.8)

(i) $V(X_1) = 2.6;\quad V(X_2) = 3.8$

(ii) Let T_1 and T_2 be the event that the article is handled by typesetter 1 and 2, respectively. Then

$$P\,(\text{No error}/T_1) = e^{-\lambda}\frac{\lambda_1^0}{0!} = e^{-2.6}.$$

$$P\,(\text{No error}/T_2) = e^{-\lambda_2}\frac{\lambda_2^0}{0!}\,e^{-3.8}$$

Thus, $$P\,(\text{No error}) = P\,(\text{No error}/T_1) + P\,(\text{No error}/T_2).\,P\,(T_2)$$

$$P\,(\text{no error}) = 0.6 \times e^{-2.6} + 0.4 \times e^{-3.8}$$

(iii) P (article was typeset by typesetter $2/X = 0$)

$$= \frac{e^{-3.8} \times 0.4}{P\,(\text{No error})} = \frac{e^{-3.8} \times 0.4}{0.6 \times e^{-2.6} + 0.4 \times e^{-3.8}}$$

(c) A coin is tossed twice. Let A denote the number of heads on the first toss and B the total number of heads on the 2 tosses. If the coin is unbalanced and a head has a 30% chance of occurring,

(i) Find the joint probability distribution of A and B.

(ii) Check whether A and B arc dependent, using appropriate statistical measures. **2×5=10**

Ans. A takes values 0, 1

B takes values 0, 1, and 2.

$$P\,(A = 0,\,B = 0) = 0.7 \times 0.7 = 0.49$$
$$P\,(A = 0,\,B = 1) = 0.7 \times 0.3 = 0.21$$
$$P\,(A = 1,\,B = 1) = 0.3 \times 0.7 = 0.21$$
$$P\,(A = 1,\,B = 2) = 0.3 \times 0.3 = 0.09$$

(ii) $$f_a\,(A) =$$
$$P\,(A = 0) = 0.7, \quad P\,(A = 1) = 0.3$$
$$f_b\,(B) =$$
$$P\,(B = 0) = 0.7 \times 0.7 = 0.49$$
$$P\,(B = 1) = 0.7 \times 0.3 + 0.3 \times 0.7 = 0.48$$
$$P\,(B = 2) = 0.3 \times 0.3 = 0.09$$
$$P\,(A = 0,\,B = 0) = 0.49 \neq P\,(A = 0) \times P\,(B = 0)$$

Thus, A and B are not independent.

Q. 7. (a) An engineer at a construction firm has a subcontract for the electrical work in the construction of a new office building. From past experience with this electrical subcontractor, the engineer knows that each light switch that is installed will be faulty with probability $p = 0.002$ independent of the other switches installed. The building will have $n = 1500$ light switches in it. Let X be the number of faulty light switches in the building. Find approximate probability $P(4 \leq X \leq 8)$. Which distribution do you use and why? **5**

Ans. $p = 0.002, \quad n = 1500; \quad q = 0.998$

$X \sim$ Bin (1500; 0.002). Since n is large ($n > 50$) and p is small ($np < 5$)

Dist. of X can be approximated by Poisson ($\lambda = np$)

$$P(4 \leq X \leq 8) = F(8) - F(3)$$
$$= 0.996 - 0.647 = 0.349$$

(b) The yield strength (ksi) for a particular type of steel is normally distributed with $\mu = 43$ and $\sigma = 4.5$.

(i) What is the 25th percentile of the distribution of this steel strength?

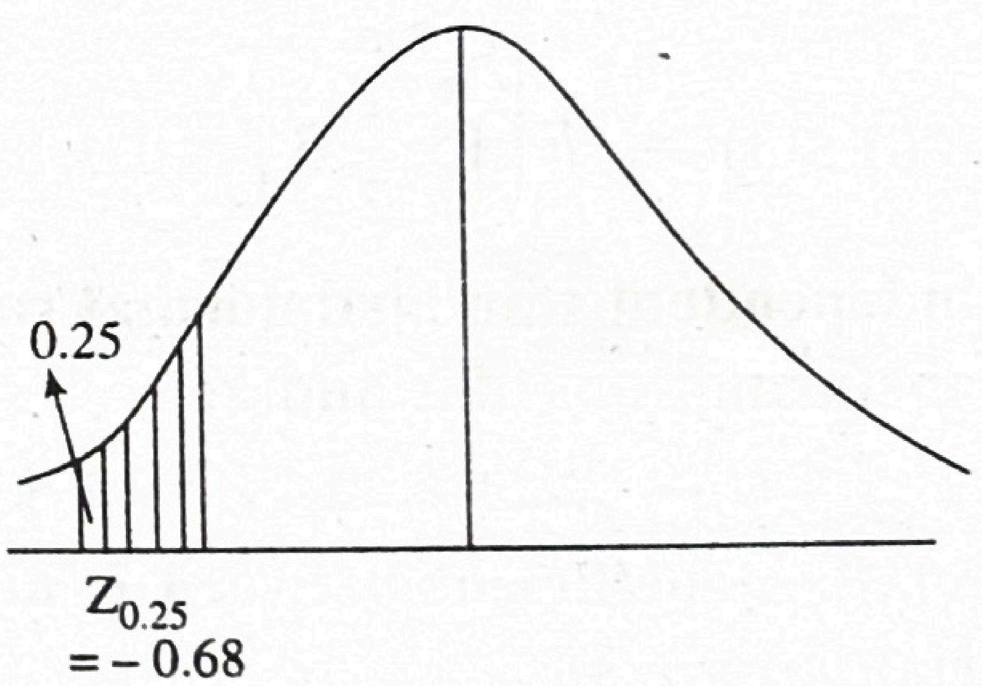
(ii) What strength value separates the strongest 10% from the others?

(iii) What is the value of c such that the interval ($43 - e$, $43 + e$) includes 99% of all strength values?

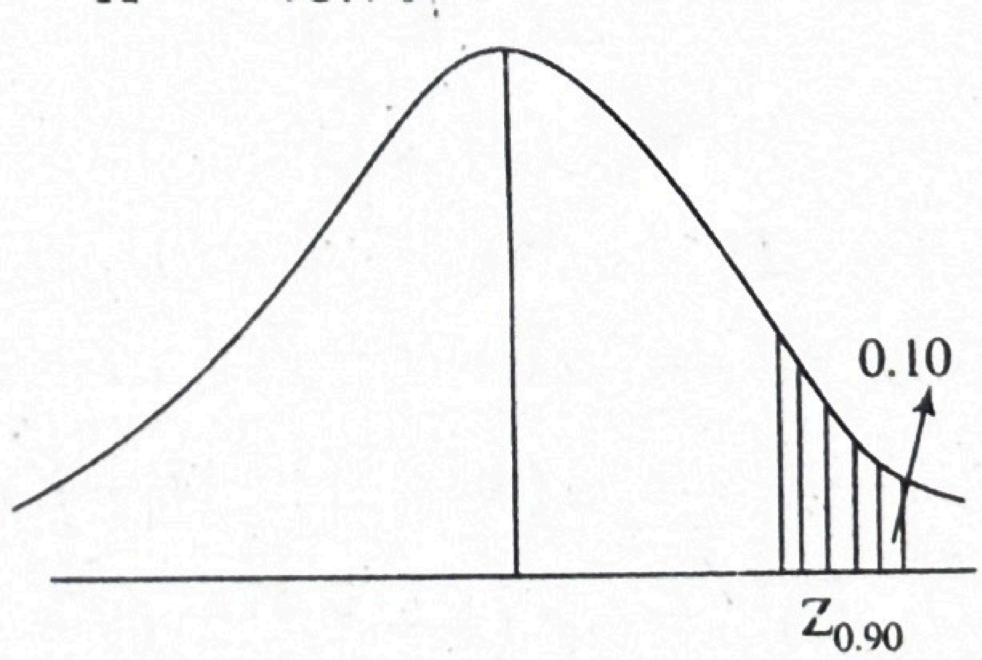
(iv) What is the probability that at most three of 15 independently selected steels have strength less than 43? **2+2+3+3**

Ans. $X \sim N(43, 4.5)$

(i) $$-0.68 = \frac{X - 43}{4.5} = 39.94$$



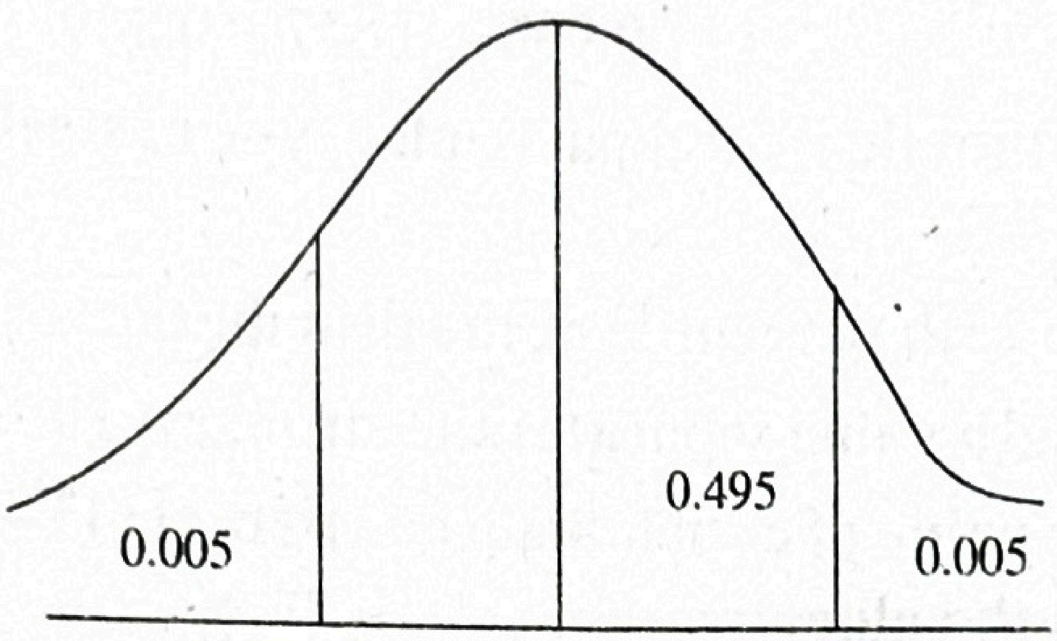
(ii) $$Z_{0.90} = 1.28$$
$$1.28 = \frac{X - 43}{4.5}$$
$$\Rightarrow \quad X = 48.76$$



(iii) $$Z_{0.005} = -2.58$$
$$X = \mu - C$$
$$X - \mu = -C$$

$$\frac{X-\mu}{\sigma} = \frac{-C}{\sigma}$$

$$-2.58 \Rightarrow \frac{-C}{\sigma} \Rightarrow \frac{C}{4.5} = 2.58 \Rightarrow C = 11.61$$



(iv)

$$n = 15; \quad p = \frac{1}{2}$$

$$P(X < 43) = P(Z < 0) = \frac{1}{2}$$

$$P(X \le 3) = F\left(\overset{n}{15}; \overset{p}{\frac{1}{2}}, \overset{x}{3}\right) = 0.018$$

(c) Let X and Y be independent standard normal random variables. Consider the following linear transformations of X and Y:

$$U = aX + b \text{ and } V = cY + d \quad a, b, c, d \in R$$

(i) Find the correlation coefficient between U and V. Justify your answer using appropriate proof (s).

(ii) Consider a random variable $W = 0.6X + 0.8Y$. Determine corr (X, W) and corr (Y, W). Compare both with corr(X, Y). **5×2=10**

Ans. $X \sim N(0, 1), \quad Y \sim N(0, 1)$

X and Y are independent $\Rightarrow$ Cov $(X, Y) = 0$

$$U = aX + b, \quad V = cY + d$$

$$E(X, Y) = e(X) . e(Y); \quad \sigma_x^2 = 1, \quad \sigma_Y^2 = 1$$

(i)

$$U = aX + b, \quad V = cY + d$$

$$\bar{U} = a\bar{X} + b \quad \bar{V} = c\bar{Y} + d$$

$$U - \bar{U} = a(X - \bar{X}); \quad (V - \bar{V}) = c(Y - \bar{Y})$$

$$(U - \bar{U})^2 = a^2 (X - \bar{X})^2; \quad (V - \bar{V})^2 = c^2 (Y - \bar{Y})^2$$

$$\text{Cov}(U, V) = \frac{1}{n}\Sigma(U - \bar{U})(V - \bar{V}) = \frac{1}{n}[\Sigma a(X - \bar{X}) c(Y - \bar{Y})]$$

$$= \frac{aC}{n}\Sigma(X-\bar{X})(Y-\bar{Y}) = aC\,\text{Cov}(X,Y)$$

$$\sigma_U^2 = \frac{1}{n}\Sigma(U-\bar{U})^2 = a^2\frac{\Sigma(X-\bar{X})^2}{n} = a^2\sigma_X^2$$

$$\sigma_V^2 = \frac{1}{n}\Sigma(V-\bar{V})^2\, c^2\frac{\Sigma(Y-\bar{Y})^2}{n} = c^2\sigma_Y^2$$

$$r(U,V) = \frac{\text{Cov}(U,V)}{\sigma_U\ \sigma_V} = \frac{aC\,\text{Cov}(X,Y)}{|a\sigma_X||c\sigma_Y|} = \frac{aC}{|ac|}r(X,Y)$$

$$\Rightarrow \quad r(aX+b, cY+d) = r(aX, cY) = \frac{ac}{|ac|}r(X,Y)$$

(ii)

$$\text{Corr}(X,Y) = 0, \quad \sigma_X^2 = 1, \quad \sigma_Y^2 = 1$$

$$\text{Cov}(X,W) = E(XW) - E(X)\cdot E(W)$$

$$E(XW) = E[X(0.6X + 0.8Y] = 0.6E(X)^2 + 0.8E(XY)$$

$$E(X)E(W) = E(X)[0.6E(X) + 0.8E(Y)]$$

$$= 0.6(E(X))^2 + 0.8E(X) - E(Y)$$

$$\text{Cov}(X,W) = 0.6\sigma_X^2 \qquad (\because E(XY) = E(X).E(Y)$$

$$\text{Corr}(X,W) = \frac{\text{Cov}(X,W)}{\sigma_X\cdot\sigma_W} = \frac{0.6\sigma_X^2}{\sigma_X\cdot\sigma_W} = 0.6\frac{\sigma_X}{\sigma_W} = 0.6$$

$$\sigma_W^2 = (0.6)^2\sigma_X^2 + (0.8)^2\sigma_Y^2 = 0.36 + 0.64 = 1$$

$$\because \quad \text{Cov}(X,Y) = 0, \quad \sigma_X^2 = 1, \quad \sigma_Y^2 = 1$$

$$\text{Corr}(X,W) = 0.6 \quad |||^{ey}\ \text{Corr}(Y,W) = 0.8$$

wehereas corr $(X, Y) = 0$

Why corr $(Y, W) = 0.8$? $\Rightarrow \text{Cov}(Y,W) = E(YW) - E(Y).E(W)$

$$E(YW) = E\{Y(0.6X + 0.8Y\} = 0.6E(XY) + 0.8E(Y^2)$$

$$E(Y).E(W) = E(Y)[0.6E(X) + 0.8E(Y)]$$

$$= 0.6E(Y).E(X) + 0.8(E(Y))^2$$

$$\text{Cov}(Y,W) = 0.8[E(Y^2) - (E(Y))^2] = 0.8\sigma_Y^2 = 0.8$$

$$\text{Corr}(Y,W) = \frac{\text{Cov}(Y,W)}{\sigma_Y\,\sigma_W} = \frac{0.8}{1} = 0.8$$

□□□