Probability

G V V Sharma*

Contents

*The author is with the Department of Electrical Engineering, Indian Institute of Technology, Hyderabad 502285 India e-mail: gadepall@iith.ac.in.

***Abstract*—This book provides solved examples on Probability**

1 Axioms

1.1. The probability that a given positive integer lying between 1 and 100 (both inclusive) is NOT divisible by 2,3 or 5 is ...
Solution: Table 1.1.1 summarizes the given information.

Event	Definition	Probability
A	$n \equiv 0 \pmod 2$	$\frac{50}{100}$
B	$n \equiv 0 \pmod 3$	$\frac{33}{100}$
C	$n \equiv 0 \pmod 5$	$\frac{20}{100}$
AB	$n \equiv 0 \pmod 6$	$\frac{16}{100}$
BC	$n \equiv 0 \pmod{15}$	$\frac{6}{100}$
AC	$n \equiv 0 \pmod{10}$	$\frac{10}{100}$
ABC	$n \equiv 0 \pmod{30}$	$\frac{3}{100}$

TABLE 1.1.1: $1 \le n \le 100$

$$\because \Pr(A+B+C) = \Pr(A) + \Pr(B) + \Pr(C) - \Pr(AB) - \Pr(BC) - \Pr(AC) + \Pr(ABC) \quad (1.1.1)$$

Substituting from Table 1.1.1 in (1.1.1),

$$\Pr(A+B+C) = \frac{50}{100} + \frac{33}{100} + \frac{20}{100} - \frac{16}{100} - \frac{6}{100} - \frac{10}{100} + \frac{3}{100} = \frac{74}{100} \quad (1.1.2)$$

Thus, the required probability is

$$1 - \Pr(A+B+C) = \frac{26}{100} \quad (1.1.3)$$

1.2. P and Q are considering to apply for a job. The probability that P applies for the job is $\frac{1}{4}$, the probability that P applies for the job given that Q applies for the job is $\frac{1}{2}$, and the probability that Q applies for the job given that P applies for the job is $\frac{1}{3}$. Then the probability that P does not apply for the job given that Q does not apply for the job is

a) $\frac{4}{5}$ b) $\frac{5}{6}$ c) $\frac{7}{8}$ d) $\frac{11}{12}$

Solution: The given information can be expressed as

$$\Pr(P) = \frac{1}{4} \quad (1.2.1)$$

$$\Pr(P|Q) = \frac{1}{2} = \frac{\Pr(PQ)}{\Pr(Q)} \quad (1.2.2)$$

$$\Pr(Q|P) = \frac{1}{3} = \frac{\Pr(PQ)}{\Pr(P)} \quad (1.2.3)$$

which yields

$$\Pr(PQ) = \frac{1}{3} \times \frac{1}{4} = \frac{1}{12}$$
$$\Pr(Q) = \frac{\frac{1}{12}}{\frac{1}{2}} = \frac{1}{6} \quad (1.2.4)$$

The objective is to find

$$\Pr(P'|Q') \quad (1.2.5)$$

(1.2.1) can be expressed as

$$\Pr(P'|Q') = \frac{\Pr(P'Q')}{\Pr(Q')} \quad (1.2.6)$$
$$= \frac{\Pr(1-(P+Q)')}{\Pr(Q')} \quad (1.2.7)$$
$$= \frac{1-\Pr(P)-\Pr(Q)+\Pr(PQ)}{1-\Pr(Q)} \quad (1.2.8)$$

Substituting from (1.2.4) and (1.2.1) in (1.2.8),

$$\Pr(P'|Q') = \frac{4}{5} \quad (1.2.9)$$

1.3. Out of 6 unbiased coins, 5 are tossed independently and they all result in heads. If the 6th coin is now independently tossed, the probability of getting head is:
(a) 1
(b) 0
(c) $\frac{1}{2}$
(d) $\frac{1}{6}$

Solution: Define a random variable $X = \{0, 1\}$ denoting the outcome of the toss of 6th coin with $X = 0$ and $X = 1$ representing tails and head respectively.Therefore,

$$\Pr(X=0) + \Pr(X=1) = 1 \quad (1.3.1)$$
$$\Pr(X=1) = \frac{1}{2} \quad (1.3.2)$$

Hence the correct answer is option (c).

1.4. Two students are solving the same problem independently,if the probability of first one solves the problem is $\frac{3}{5}$ and the probability that the second one solves the problem is $\frac{4}{5}$,what is the probability that atleast one of them solves the problem?
a) $\frac{17}{25}$
b) $\frac{19}{25}$
c) $\frac{21}{25}$
d) $\frac{23}{25}$

Solution: Let X,Y be two events representing solving the problem by students A,B respectively.

Given

$$\Pr(X) = \frac{3}{5} \tag{1.4.1}$$

$$\Pr(Y) = \frac{4}{5} \tag{1.4.2}$$

Since students solve the problem independently,So events X and Y are independent, For independent events

$$\Pr(XY) = \Pr(X) \times \Pr(Y) \tag{1.4.3}$$

from (1.4.1) and (1.4.2)

$$\Pr(XY) = \frac{3}{5} \times \frac{4}{5} \tag{1.4.4}$$

$$\Pr(XY) = \frac{12}{25} \tag{1.4.5}$$

Now we have to find probability of solving the problem by atleast one of them i.e $\Pr(X + Y)$. As,

$$\Pr(X + Y) = \Pr(X) + \Pr(Y) - \Pr(XY) \tag{1.4.6}$$

from (1.4.1), (1.4.2), (1.4.5)

$$\Pr(X + Y) = \frac{3}{5} + \frac{4}{5} - \frac{12}{25} \tag{1.4.7}$$

$$\Pr(X + Y) = \frac{23}{25} \tag{1.4.8}$$

Hence the required probability is $\frac{23}{25}$

1.5. Three types of components are used in electrical circuits 1, 2, 3 as shown below in the figure
Solution: For q_1, the truth table Multiplying

A	B	C	$(AB) + C$
1	1	0	1
1	1	1	1
0	1	1	1
0	0	1	1
1	0	1	1

TABLE 1.5.1: Circuit 1 working

and adding probability for each case of q_1 gives us the value of q_1 as

$$q_1 = p^3 - 2p^2 + 1 \tag{1.5.1}$$

For q_2, the truth table Multiplying and adding probability for each case of q_2 gives us the

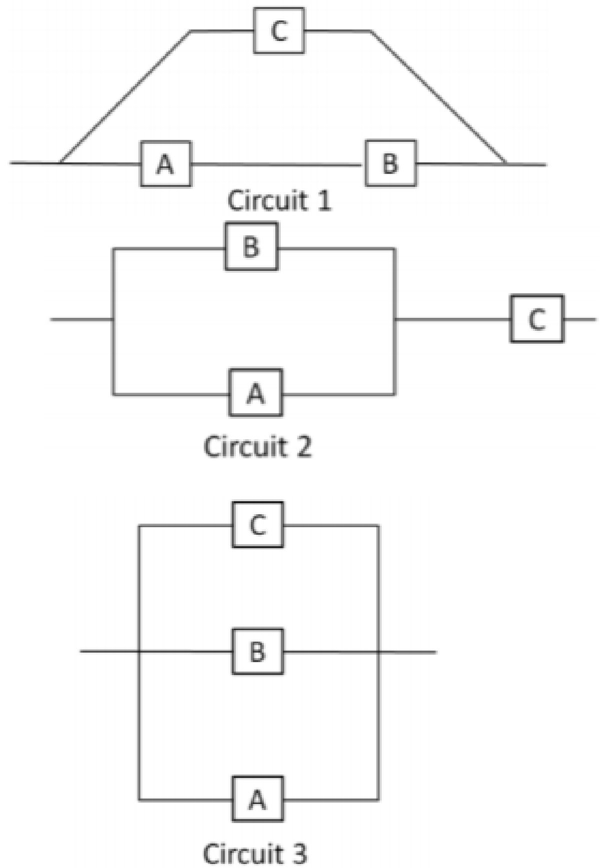


Fig. 1.5.1: Figure

A	B	C	$(A + B)C$
1	1	1	1
1	0	1	1
0	1	1	1

TABLE 1.5.2: Circuit 2 working

value of q_2 as

$$q_2 = p^3 - p^2 - p + 1 \tag{1.5.2}$$

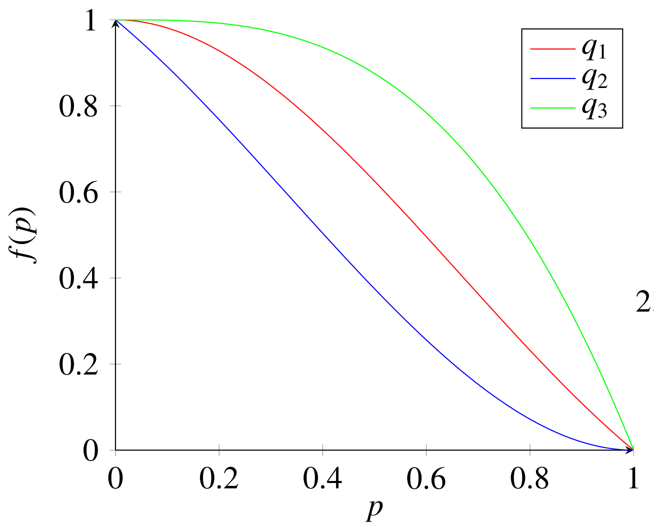
For q_3, the truth table Multiplying and adding

A	B	C	$A + B + C$
1	0	0	1
0	1	0	1
0	0	1	1
1	1	0	1
1	0	1	1
0	1	1	1
1	1	1	1

TABLE 1.5.3: Circuit 3 working

probability for each case of q_3 gives us the value of q_3 as

$$q_3 = 1 - p^3 \tag{1.5.3}$$



$$\therefore q_3 > q_1 > q_2 \tag{1.5.4}$$

Hence **Option 1** is correct

2 Elementary Probability

2.1. There are 3 red socks, 4 green socks and 3 blue socks.You choose 2 socks.The probability that they are of the same colour is

a) $\frac{1}{5}$ b) $\frac{7}{30}$ c) $\frac{1}{4}$ d) $\frac{4}{15}$

Solution: Let $X_i \in \{1, 2, 3\}$ represent the i^{th} draw, where 1, 2, 3 correspond to the colour of socks drawn as Red, Blue and Green respectively

TABLE 2.1.1

	$X_1 = 1$	$X_1 = 2$	$X_1 = 3$
$X_2 = 1$	6/90	12/90	9/90
$X_2 = 2$	12/90	12/90	12/90
$X_2 = 3$	9/90	12/90	6/90

TABLE 2.1.1 represents all the possibilities of choosing socks one by one.
The probability that the two socks drawn are of the same colour(by substituting values from table 2.1.1)

$$= \Pr(X_1 = X_2) \tag{2.1.1}$$

$$= \sum_{i=1}^{3} \Pr(X_2 = i|X_1 = i)\Pr(X_1 = i) \tag{2.1.2}$$

$$= \frac{6}{90} + \frac{12}{90} + \frac{6}{90} \tag{2.1.3}$$

$$= \frac{4}{15} \tag{2.1.4}$$

So the correct option is (D)

2.2. A box contains 40 numbered red balls and 60 numbered black balls. From the box, balls are drawn one by one at random without replacement till all the balls are drawn. The probability that the last ball drawn is black equals . . . Now, this problem is equivalent to the problem where we have to arrange 40 distinct R's and 60 distinct B's such that, a B should come at last. So, the desired probability is given by

$$\frac{(\text{placing a B at last}) \times (\text{arranging other letters})}{\text{arranging 100 letters}}$$

$$= \frac{60 \times 99!}{100!} = \frac{3}{5} \tag{2.2.1}$$

2.3. An experiment consists of two papers.paper1 and paper2.The probability of failing in paper 1 is .3 and that in paper 2 is .2.Given that a student has failed in paper 2,the probability of failing in paper 1 is .6.The probability of student failing in both is

a) .5

b) .18

c) .12

d) .06

Solution: Table 2.3.1 summarises the given information. The desired probability is

$$\Pr(X = 0, Y = 0) = \Pr(X = 0|Y = 0)\Pr(Y = 0) \tag{2.3.1}$$

$$= .12 \tag{2.3.2}$$

2.4. An urn contains 5 red balls and 5 black balls.In

	Description	Probability
0	failure	$\Pr(X = 0) = 0.3$
1	success	$\Pr(Y = 0) = 0.2$
X	Paper 1	$\Pr(X = 0\|Y = 0) = 0.6$
Y	Paper 2	

TABLE 2.3.1: Description

the first draw, one ball is picked at random and discarded without noticing its colour.The probability to get a red ball in the second draw is

a) $\frac{1}{2}$ b) $\frac{4}{9}$ c) $\frac{5}{9}$ d) $\frac{6}{9}$

Solution: Let $X_i \in \{0, 1\}$ represent the i^{th} draw where 1 denotes a red ball being drawn.

	$X_1 = 0$	$X_1 = 1$
$X_2 = 0$	4/18	5/18
$X_2 = 1$	5/18	4/18

TABLE 2.4.1: The probabilities of all possible cases when two balls are drawn one by one from the urn.

From Table 2.4.1,

$$\Pr(X_2 = 1) = \Pr(X_2 = 1, X_1 = 0) + \Pr(X_2 = 1, X_1 = 1) \quad (2.4.1)$$
$$= \frac{5}{18} + \frac{4}{18} \quad (2.4.2)$$
$$= \frac{1}{2} \quad (2.4.3)$$

The required option is (A).

2.5. A sample of size $n = 2$ is drawn from a population of size $N = 4$ using probability proportional to size without replacement scheme , Where the probabilities proportional to size are The probability of inclusion of unit (1) in

i:	1	2	3	4
p_i	0.4	0.2	0.2	0.2

Table : Probability vs Size

the sample is

a) 0.4 b) 0.6 c) 0.7 d) 0.75

Solution: Let $P_i(j)$ represent the probability for selecting unit (j) as second unit after selecting unit (i)

$$P_i(j) = \frac{p_j}{1 - p_i} \quad (2.5.1)$$

Let $\Pr(i, j)$ be probability of selecting sample {i,j} ,using (2.5.1) is

$$\Pr(i, j) = P_i(j) + P_j(i) \quad (2.5.2)$$
$$= \left(p_i \times \frac{p_j}{1 - p_i}\right) + \left(p_j \times \frac{p_i}{1 - p_j}\right) \quad (2.5.3)$$

Total samples(Size $n = 2$)are Let P_i be

Case	1	2	3	4	5	6
Sample(size $n = 2$)	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)

TABLE 2.5.1: list of samples

the probability of inclusion of unit (i) in the sample(size $n = 2$),Now i will calculate P_1 ,Favourable cases for inclusion of unit(1) are case (1,2,3),So

$$P_1 = \Pr(1, 2) + \Pr(1, 3) + \Pr(1, 4) \quad (2.5.4)$$

using (2.5.3) and p_i from question ,

$$P_1 = \frac{7}{30} + \frac{7}{30} + \frac{7}{30} \quad (2.5.5)$$
$$= 0.7 \quad (2.5.6)$$

Therefore Option (3) is correct.

2.6. There are two boxes. Box-1 contains 2 red balls and 4 green balls. Box-2 contains 4 red balls and 2 green balls. A box is selected at random and a ball is chosen randomly from the selected box. If the ball turns out to be red, what is the probability that Box-1 had been selected?

Solution: Box-1 has 2 red balls and 4 green balls.

Box-2 has 4 red balls and 2 green balls.

Let $B \in \{1, 2\}$ represent a random variable where 1 represents selecting box-1 and 2 represents selecting box-2. From Baye's theorem

$$\Pr(R = 1) = \Pr(R = 1|B = 1) \times \Pr(B = 1) + \Pr(R = 1|B = 2) \times \Pr(B = 2) \quad (2.6.1)$$

Event	definition	value
$\Pr(B = 1)$	Probability of selecting Box-1	$\frac{1}{2}$
$\Pr(B = 2)$	Probability of selecting Box-2	$\frac{1}{2}$
$\Pr(R = 1\|B = 1)$	Probability of drawing red ball from Box-1	$\frac{1}{3}$
$\Pr(G = 1\|B = 1)$	Probability of drawing green ball from Box-1	$\frac{2}{3}$
$\Pr(R = 1\|B = 2)$	Probability of drawing red ball from Box-2	$\frac{2}{3}$
$\Pr(G = 1\|B = 2)$	Probability of drawing green ball from Box-2	$\frac{1}{3}$

TABLE 2.6.1: Table 1

Substiting values from table (2.6.1) in (2.6.1)

$$\Pr(R = 1) = \frac{1}{2} \tag{2.6.2}$$

$$\Pr((R = 1)(B = 1)) = \Pr(R = 1|B = 1) \times \Pr(B = 1) \tag{2.6.3}$$

$$= \frac{1}{6} \tag{2.6.4}$$

We need to find $\Pr(B = 1|R = 1)$

$$\Pr(B = 1|R = 1) = \frac{\Pr((R = 1)(B = 1))}{\Pr(R = 1)} \tag{2.6.5}$$

$$= \frac{1}{3} \tag{2.6.6}$$

$\therefore$ The desired probability that box-1 is selected $= \frac{1}{3}$

2.7. An urn has 3 red and 6 black balls. Balls are drawn at random one by one without replacement. The probability that second red ball appears on fifth draw is:

a) $\frac{1}{9!}$

b) $\frac{4!}{9!}$

c) $4\left(\frac{6!4!}{9!}\right)$

d) $\frac{6!4!}{9!}$

Solution: To obtain a second red ball at the fifth draw, the first 4 trials should involve drawing only 1 red ball out of the 3 and 3 black balls out of the 6. Probability of this happening:

$$\frac{{}^3C_1{}^6C_3}{{}^9C_4} \tag{2.7.1}$$

The probability of the fifth ball turning out to be red is:

$$\frac{{}^2C_1}{{}^5C_1} \tag{2.7.2}$$

By Multiplication rule, total probability:

$$\frac{{}^3C_1{}^6C_3{}^2C_1}{{}^5C_1{}^9C_4} = \frac{3! \times 6! \times 2! \times 4! \times 4! \times 5!}{2! \times 3! \times 3! \times 5! \times 9!} \tag{2.7.3}$$

$$= 4\left(\frac{4!6!}{9!}\right) \tag{2.7.4}$$

2.8. An unbalanced dice (with 6 faces, numbered from 1 to 6) is thrown. The probability that the face value is odd is 90% of the probability that the face value is even. The probability of getting any even numbered face is the same. If the probability that the face is even given that it is greater than 3 is 0.75, which one of the following options is closest to the probability that the face value exceeds 3?

(A) 0.453
(B) 0.468
(C) 0.485
(D) 0.492

3 Random Variables

3.1. Consider the function *f(x)* defined as $f(x) = ce^{-x^4}$, $x \in R$. For what value of c is f a probability density function?

a) $\frac{2}{\Gamma(1/4)}$

b) $\frac{4}{\Gamma(1/4)}$

c) $\frac{3}{\Gamma(1/3)}$

d) $\frac{1}{4\Gamma(4)}$

Solution:

3.2. Let X and Y be i.i.d random variables uniformly distributed on (0,4).Then $\Pr(X > Y|X < 2Y)$ is

a) 1/3
b) 5/6
c) 1/4
d) 2/3

Solution:
The PDF is given by

$$f_X(x) = f_Y(x) = \begin{cases} \frac{1}{4}, & \text{if } 0 < x < 4 \\ 0, & \text{otherwise} \end{cases}$$

The CDF is given by

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

$$F_X(x) = F_Y(x) = \begin{cases} 0, & x \le 0 \\ \frac{x}{4}, & \text{if } 0 < x < 4 \\ 1, & x \ge 4 \end{cases}$$

Using definition of conditional probability

$$\Pr(X > Y|X < 2Y) = \frac{\Pr(Y < X < 2Y)}{\Pr(X < 2Y)} \tag{3.2.1}$$

Now finding $\Pr(X < 2Y)$

$$\Pr(X < 2y) = F_X(2y) \tag{3.2.2}$$

$$\implies \Pr(X < 2Y) = \int_{-\infty}^{\infty} f_Y(x) \times F_X(2x)dx \tag{3.2.3}$$

$$\implies \Pr(X < 2Y) = \int_0^2 \frac{x}{8}dx + \int_2^4 \frac{1}{4}dx \tag{3.2.4}$$

$$\implies \Pr(X < 2Y) = \frac{3}{4} = 0.75 \tag{3.2.5}$$

Now to find $\Pr(Y < X < 2Y)$

$$\Pr(y < X < 2y) = F_X(2y) - F_X(y) \tag{3.2.6}$$

$$\implies \Pr(Y < X < 2Y) \tag{3.2.7}$$

$$= \int_{-\infty}^{\infty} f_Y(x)(F_X(2x) - F_X(x))dx$$

$$\implies \int_0^2 \frac{1}{4}\left(\frac{x}{2} - \frac{x}{4}\right)dx + \int_2^4 \frac{1}{4}\left(1 - \frac{x}{4}\right)dx \tag{3.2.8}$$

$$\implies \Pr(Y < X < 2Y) = \frac{1}{4} = 0.25 \tag{3.2.9}$$

Now using (3.2.1),(3.2.5) and (3.2.9)

$$\Pr(X > Y|X < 2Y) = \frac{1/4}{3/4} = \frac{1}{3} \tag{3.2.10}$$

Hence final solution is option 1) or 1/3

3.3. Suppose X is a positive random variable with the following probability density function,

$$f(x) = (\alpha x^{\alpha-1} + \beta x^{\beta-1})e^{-x^\alpha - x^\beta}; x > 0$$

for $\alpha > 0, \beta > 0$. Then the hazard function of X for some choices of α and β can be

a) an increasing function.
b) a decreasing function.
c) a constant function.
d) a non monotonic function

Solution:
CDF of X,

$$F(x) = \int_{-\infty}^{x} f(t)dt \tag{3.3.1}$$

$$= \int_0^x f(t)dt \qquad \text{as } x > 0 \tag{3.3.2}$$

$$= \int_{-\infty}^{t} \left((\alpha t^{\alpha-1} + \beta t^{\beta-1}) \times e^{-t^\alpha - t^\beta}\right) dt \tag{3.3.3}$$

$$= -e^{-t^\alpha - t^\beta}\Big|_0^x \tag{3.3.4}$$

$$= 1 - e^{-x^\alpha - x^\beta} \tag{3.3.5}$$

Hazard function,

$$h(x) = \frac{f(x)}{1 - F(x)} \tag{3.3.6}$$

$$= \alpha x^{\alpha-1} + \beta x^{\beta-1} \tag{3.3.7}$$

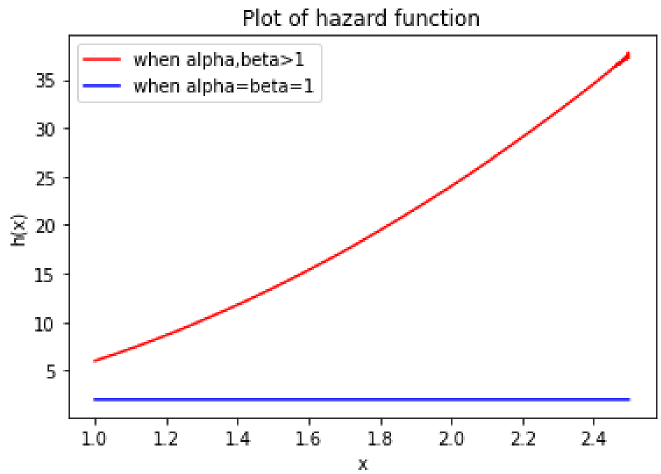
$$h'(x) = \alpha(\alpha - 1)x^{\alpha-2} + \beta(\beta - 1)x^{\beta-2} \tag{3.3.8}$$

$$h'(x) = \begin{cases} 0 & \alpha = \beta = 1 \\ > 0 & \text{otherwise} \end{cases} \tag{3.3.9}$$

Thus $h(x)$ can be either constant function or an increasing function.
From the above figure, it is verified that $h(x)$ can be either constant function or an increasing function.
Correct options are 1,3.

3.4. Suppose the random variable X has the following probability density funtion

$$f(x) = \begin{cases} \alpha\,(x - \mu)^{\alpha-1}\, e^{-(x-\mu)^\alpha}; & x > \mu \\ 0 & x \le \mu \end{cases}$$



where $\alpha > 0, -\infty < \mu < \infty$. Which of the following are correct? The hazard function of X is

a) an increasing function for all $\alpha > 0$
b) a decreasing function for all $\alpha > 0$
c) an increasing function for some $\alpha > 0$
d) a decreasing function for some $\alpha > 0$

Solution:
For the random variable X, the CDF is

$$F(x) = solutions/2017/june/118/figures/\int_0^x f(y)dy \tag{3.4.1}$$

$$= solutions/2017/june/118/figures/\int_0^{\mu} 0dy + soluti \ldots -\mu)^{\alpha-1} e^{-(y-\mu)^{\alpha}} \tag{3.4.2}$$

$$= 0 - e^{-(y-\mu)^{\alpha}}\Big|_{\mu}^{x} \tag{3.4.3}$$

$$= 1 - e^{-(x-\mu)^{\alpha}} \tag{3.4.4}$$

For X, the hazard function $H(y)$ is defined as

$$H(y) = \frac{f(y)}{1 - F(y)}$$

$$\implies H(y) = \begin{cases} \frac{\alpha(y-\mu)^{\alpha-1} e^{-(y-\mu)^{\alpha}}}{1-\left(1-e^{-(y-\mu)^{\alpha}}\right)}; & y > \mu \\ 0 & y \le \mu \end{cases}$$

$$= \begin{cases} \alpha\,(y-\mu)^{\alpha-1}; & y > \mu \\ 0 & y \le \mu \end{cases}$$

Differentiating $H(y)$ w.r.t. y

$$H'(y) = \begin{cases} \alpha\,(\alpha-1)\,(y-\mu)^{\alpha-2}; & y > \mu \\ 0 & y \le \mu \end{cases}$$

When $y \le \mu$ then $H'(y)$ is 0. When $y > \mu$ then $(y-\mu)^{\alpha-2}$ is positive. This implies that the sign for $H'(y)$ for $y > \mu$ is decided by the sign of $\alpha\,(\alpha-1)$.

$$\alpha\,(1-\alpha) < 0 \implies 0 < \alpha < 1$$
$$\alpha\,(1-\alpha) > 0 \implies \alpha > 1 \quad (\text{ignoring } \alpha < 0)$$

$\therefore$ The Hazard function of X is decreasing when $\alpha \in (0,1)$ and increasing when $\alpha \in (1,\infty)$ **Solution:** Options 3, 4

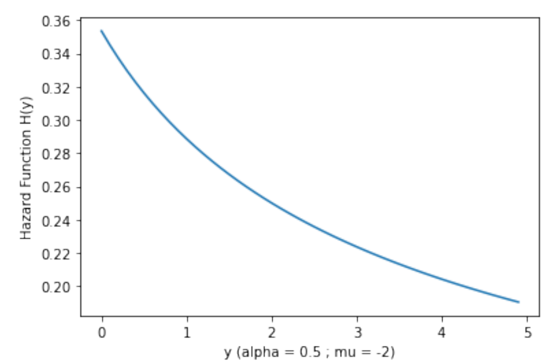


Fig. 3.4.1: Decreasing Hazard Function

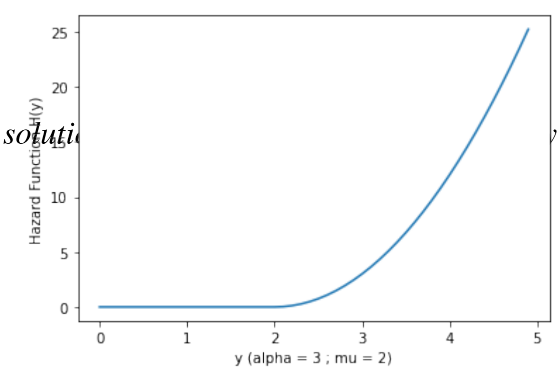


Fig. 3.4.2: Increasing Hazard Function

3.5. Let X be a random variable with a certain non-degenerate distribution. Then identify the correct statements

a) If X has an exponential distribution then $median\,(X) < E\,(X)$
b) If X has a uniform distribution on an interval $[a,b]$, then $E\,(X) < median\,(X)$

c) If X has a Binomial distribution then $V(X) < E(X)$
d) If X has a normal distribution, then $E(X) < V(X)$

Solution: Expected value($E(X)$): It is nothing but weighted average Median($median(X)$): It is the value separating the higher half from the lower half of a data sample
Variance($V(X)$): It is the expectation of the squared deviation of a random variable from its mean

a) Let's consider X has an exponential distribution.

$$X \sim Exp(\lambda) \tag{3.5.1}$$

where λ is rate parameter.
Probability function of exponential distribution,

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{3.5.2}$$

The expected value of $X \sim Exp(\lambda)$,

$$E(X) = \frac{1}{\lambda} \tag{3.5.3}$$

The median of $X \sim Exp(\lambda)$,

$$median(X) = \frac{\ln 2}{\lambda} \tag{3.5.4}$$

$$\ln 2 < 1 \tag{3.5.5}$$
$$\frac{\ln 2}{\lambda} < \frac{1}{\lambda} \tag{3.5.6}$$
$$median(X) < E(X) \tag{3.5.7}$$

Hence, option 1 is correct.

b) Let's consider X has a uniform distribution in interval $[a, b]$,

$$X \sim U(a, b) \tag{3.5.8}$$

where, a = lower limit
b = upper limit
Probability function of uniform distribution,

$$f_X(k) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x < a, x > b \end{cases} \tag{3.5.9}$$

The expected value of $X \sim U(a, b)$,

$$E(X) = \frac{1}{2}(a + b) \tag{3.5.10}$$

The median of $X \sim U(a, b)$,

$$median(X) = \frac{1}{2}(a + b) \tag{3.5.11}$$

$$E(X) = median(X) \tag{3.5.12}$$

Hence, option 2 is incorrect.

c) Let's consider X has a binomial distribution,

$$X \sim B(n, p) \tag{3.5.13}$$

where, n = no. of trails
p = success parameter
Probability function of binomial distribution,

$$f_X(k) = \begin{cases} {}^nC_k p^k (1-p)^{n-k} & 0 \leq k \leq n \\ 0 & otherwise \end{cases} \tag{3.5.14}$$

The expected value of $X \sim B(n, p)$,

$$E(X) = np \tag{3.5.15}$$

The variance of $X \sim B(n, p)$,

$$V(X) = \sigma^2 = np(1-p) \tag{3.5.16}$$

$$1 - p \leq 1 \tag{3.5.17}$$
$$np(1-p) \leq np \tag{3.5.18}$$
$$V(X) \leq E(X) \tag{3.5.19}$$

Hence, option 3 is incorrect.

d) Let's consider X has a normal distribution,

$$X \sim N\left(\mu, \sigma^2\right) \tag{3.5.20}$$

where, μ = mean of distribution
σ^2 = variance
Probability function of normal distribution,

$$f_X(k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{2\sigma}\right)^2} \tag{3.5.21}$$

The expected value of $X \sim N\left(\mu, \sigma^2\right)$,

$$E(X) = \mu \tag{3.5.22}$$

The variance of $X \sim N\left(\mu, \sigma^2\right)$,

$$V(X) = \sigma^2 \tag{3.5.23}$$

$E(X)$ and $V(X)$ are user defined. So, they can take any value.
Hence, option 4 is incorrect.

3.6. A fair coin is tossed repeatedly. Let X be the number of tails before the first heads occurs. Let Y denote the number of tails between the first and second heads. Let $X + Y = N$. Then which of the following are true?

a) X and Y are independent random variables with

$$\Pr(X = k) = \Pr(Y = k) = \begin{cases} 2^{-(k+1)} & k = 0, 1, 2 \dots \\ 0 & otherwise \end{cases} \quad (3.6.1)$$

b) N has a probability mass function given by

$$\Pr(N = k) = \begin{cases} (k-1)2^{-k} & k = 2, 3, 4 \dots \\ 0 & otherwise \end{cases} \quad (3.6.2)$$

c) Given $N = n$, the conditional distribution of X and Y are independent

d) Given $N = n$

$$\Pr(X = k) = \begin{cases} \frac{1}{n+1} & n = 0, 1, 2 \dots \\ 0 & otherwise \end{cases} \quad (3.6.3)$$

3.7. Assume that $X \sim Binomial(n, p)$ for some $n \geq 1$ and $0 < p < 1$ and $Y \sim poisson(\lambda)$for some $\lambda > 0$.Suppose E[X]=E[Y].Then

a) $var(X) = Var(Y)$
b) $var(X) < Var(Y)$
c) $var(Y) < Var(X)$
d) $Var(X)$ may be larger or smaller than $Var(Y)$ depending on the values of n,p and λ

Solution: For the random variable

$$X \sim Binomial(n, p)$$

As we know,

$$E[X] = np \quad (3.7.1)$$
$$Var(X) = np(1-p) \quad (3.7.2)$$

for the random variable

$$Y \sim poisson(\lambda)$$

As we know,

$$E[Y] = \lambda \quad (3.7.3)$$
$$Var(Y) = \lambda \quad (3.7.4)$$

given that,

$$E[X] = E[Y] \quad (3.7.5)$$
$$np = \lambda \quad (3.7.6)$$

from (3.7.2),

$$Var(X) = np(1-p) \quad (3.7.7)$$

using (3.7.6),

$$Var(X) = \lambda(1-p) \quad (3.7.8)$$

using (3.7.4),

TABLE 3.7.1: Mean and Variance for random variables X and Y

	X	Y
E	λ	λ
var	$\lambda(1-p)$	λ

$$Var(X) = Var(Y)(1-p) \quad (3.7.9)$$
$$\frac{Var(X)}{Var(Y)} = 1 - p \quad (3.7.10)$$

as,

$$1 - p < 1 \quad (3.7.11)$$
$$\frac{Var(X)}{Var(Y)} < 1 \quad (3.7.12)$$
$$Var(X) < Var(Y) \quad (3.7.13)$$

$\therefore Var(Y) > Var(X)$, independent of n,p and λ.

a) $var(X) = Var(Y)$
using TABLE 3.7.1,

$$\lambda(1-p) = \lambda \quad (3.7.14)$$
$$1 - p = 1 \quad (3.7.15)$$
$$p = 0 \quad (3.7.16)$$

which is wrong as per the question($0 < p < 1$). hence the option is incorrect.

b) $var(X) < Var(Y)$
using TABLE 3.7.1,

$$\lambda(1-p) < \lambda \quad (3.7.17)$$
$$1 - p < 1 \quad (3.7.18)$$
$$p > 0 \quad (3.7.19)$$

which is true as per the question($0 < p < 1$). hence the option is correct.

c) $var(Y) < Var(X)$
using TABLE 3.7.1,

$$\lambda(1-p) > \lambda \quad (3.7.20)$$
$$1 - p > 1 \quad (3.7.21)$$
$$p < 0 \quad (3.7.22)$$

which is wrong as per the question($0 < p < 1$). hence the option is incorrect.

d) $Var(X)$ may be larger or smaller than $Var(Y)$ depending on the values of n,p and λ.
Wrong, since we have shown that irrespective of the values of lambda, n, and p, $var(y) > var(x)$

3.8. Let X be a non-negative integer valued random variable with probability mass function $f(x)$ satisfying $(x+1)f(x+1) = (\alpha + \beta x)f(x),\ x = 0, 1, 2, ...;$ $\beta \neq 1$. You may assume that $E(X)$ and $Var(X)$ exist. Then which of the following statements are true?

a) $E(X) = \dfrac{\alpha}{1-\beta}$

b) $E(X) = \dfrac{\alpha^2}{(1-\beta)(1+\alpha)}$

c) $Var(X) = \dfrac{\alpha^2}{(1-\beta)^2}$

d) $Var(X) = \dfrac{\alpha}{(1-\beta)^2}$

Solution: For a discrete random variable X with P.D.F. $f(x)$ and which can take values from a set $\mathbb{S}$,

$$E(X) = \sum_{x \in \mathbb{S}} x f(x) \tag{3.8.1}$$

And,

$$E(X^2) = \sum_{x \in \mathbb{S}} x^2 f(x) \tag{3.8.2}$$

Also, as $f(x)$ is the P.D.F.,

$$\sum_{x \in \mathbb{S}} f(x) = 1 \tag{3.8.3}$$

Given, for $x \in \mathbb{S} = \{0, 1, 2, ...n\}$,

$$(x+1)f(x+1) = (\alpha + \beta x)f(x) \tag{3.8.4}$$

Summing both sides for $x \in \mathbb{S}$ we get,

$$\sum_{x=0}^{n}(x+1)f(x+1) = \sum_{x=0}^{n}(\alpha + \beta x)f(x) \tag{3.8.5}$$

Replacing $x+1$ with x in L.H.S. we get,

$$\sum_{x=1}^{n+1} x f(x) = \sum_{x=0}^{n}(\alpha + \beta x)f(x) \tag{3.8.6}$$

Rewriting LHS, we get,

$$\sum_{x=0}^{n} x f(x) + (n+1)f(n+1) = \sum_{x=0}^{n}(\alpha + \beta x)f(x) \tag{3.8.7}$$

But as $x \in \{0, 1, 2...n\}$, $f(n+1) = 0$. So the equation becomes

$$\sum_{x=0}^{n} x f(x) = \alpha \sum_{x=0}^{n} f(x) + \beta \sum_{x=0}^{n} x f(x) \tag{3.8.8}$$

Using (3.8.1) and (3.8.3), we get,

$$E(X) = \alpha(1) + \beta E(X) \tag{3.8.9}$$

So,

$$E(X) = \frac{\alpha}{1-\beta} \tag{3.8.10}$$

Now in (3.8.4), multiplying both sides by $(x+1)$, we get,

$$(x+1)^2 f(x+1) = (\alpha + \beta x)(x+1)f(x) \tag{3.8.11}$$

Summing both sides for $x \in \mathbb{S}$ we get,

$$\sum_{x=0}^{n}(x+1)^2 f(x+1) = \sum_{x=0}^{n}(\alpha + \beta x)(x+1)f(x) \tag{3.8.12}$$

Replacing $x+1$ with x in L.H.S. we get,

$$\sum_{x=1}^{n+1} x^2 f(x) = \sum_{x=0}^{n}(\beta x^2 f(x) + (\alpha + \beta)x f(x) + \alpha f(x)) \tag{3.8.13}$$

Rewriting LHS similarly as before, we get,

$$\sum_{x=0}^{n} x^2 f(x) = \beta \sum_{x=0}^{n} x^2 f(x) + (\alpha + \beta) \sum_{x=0}^{n} x f(x) + \alpha \sum_{x=0}^{n} f(x) \tag{3.8.14}$$

Using (3.8.1), (3.8.2) and (3.8.3), we get,

$$E(X^2) = \beta E(X^2) + (\alpha + \beta)E(X) + \alpha(1) \tag{3.8.15}$$

Using (3.8.10)

$$E(X^2)(1-\beta) = \frac{\alpha(\alpha + \beta)}{1-\beta} + \alpha \tag{3.8.16}$$

So,

$$E(X^2) = \frac{\alpha^2 + \alpha}{(1-\beta)^2} \tag{3.8.17}$$

Now,

$$Var(X) = E(X^2) - (E(X))^2 \tag{3.8.18}$$

Using (3.8.10) and (3.8.17),

$$Var(X) = \frac{\alpha^2 + \alpha}{(1-\beta)^2} - \frac{\alpha^2}{(1-\beta)^2} \tag{3.8.19}$$

So,

$$Var(X) = \frac{\alpha}{(1-\beta)^2} \tag{3.8.20}$$

So, options 1 and 4 are correct.

3.9. Let X be a random variable with probability density function,

$$f(x) = \alpha(x-\mu)^{\alpha-1} e^{-(x-\mu)^\alpha} \tag{3.9.1}$$

such that $-\infty < \mu < \infty$; $\alpha > 0$; $x > \mu$, The hazard function is:

a) constant for all α
b) an increasing function for some α
c) independent of α
d) independent of μ when $\alpha = 1$

Solution: Given PDF of X,

$$f(x) = \alpha(x-\mu)^{\alpha-1} e^{-(x-\mu)^\alpha} \tag{3.9.2}$$

Important property(using in (3.9.8) as $x > \mu$): Given $x - y > 0$ and $-\infty < y < \infty$, then

$$\lim_{x\to-\infty} x - y = 0 \tag{3.9.3}$$

CDF of X,

$$F(x) = \int_{-\infty}^{x} f(x)\,dx \tag{3.9.4}$$

$$= \int_{-\infty}^{x} \alpha(x-\mu)^{\alpha-1} e^{-(x-\mu)^\alpha} dx \tag{3.9.5}$$

$$= \int_{-\infty}^{x} e^{-(x-\mu)^\alpha} d(x-\mu)^\alpha \tag{3.9.6}$$

$$= \left[\frac{e^{-(x-\mu)^\alpha}}{-1}\right]_{-\infty}^{x} \tag{3.9.7}$$

$$= -e^{-(x-\mu)^\alpha} - \lim_{x\to-\infty} \frac{e^{-(x-\mu)^\alpha}}{-1} \tag{3.9.8}$$

$$= -e^{-(x-\mu)^\alpha} + e^{-(0)^\alpha} \tag{3.9.9}$$

$$F(x) = 1 - e^{-(x-\mu)^\alpha} \tag{3.9.10}$$

Hazard function $\beta(x)$,(using (3.9.2) and (3.9.10))

$$\beta(x) = \frac{f(x)}{1 - F(x)} \tag{3.9.11}$$

$$= \frac{\alpha(x-\mu)^{\alpha-1} e^{-(x-\mu)^\alpha}}{1 - (1 - e^{-(x-\mu)^\alpha})} \tag{3.9.12}$$

$$= \frac{\alpha(x-\mu)^{\alpha-1} e^{-(x-\mu)^\alpha}}{e^{-(x-\mu)^\alpha}} \tag{3.9.13}$$

$$\beta(x) = \alpha(x-\mu)^{\alpha-1} \tag{3.9.14}$$

a) $\beta(x)$ is not constant for all α
b) $\beta(x) = \alpha(x-\mu)^{\alpha-1}$ is an increasing function for $\alpha < 0$ *or* $\alpha > 1$ as given $x - \mu > 0$ for all x.

Proof: Using first derivative test, A function is increasing iff its first derivative is positive for all x.

$$\frac{d}{dx}\beta(x) = \frac{d}{dx}\alpha(x-\mu)^{\alpha-1} \tag{3.9.15}$$

$$= \alpha(\alpha-1)(x-\mu)^{\alpha-2} \tag{3.9.16}$$

For (3.9.16) to be positive, (As given $x - \mu > 0$)

$$\alpha(\alpha-1)(x-\mu)^{\alpha-2} > 0 \tag{3.9.17}$$

$$\alpha(\alpha-1) > 0 \tag{3.9.18}$$

$$\implies \alpha \in (-\infty, 0) \cup (1, \infty) \tag{3.9.19}$$

$\therefore \beta(x)$ an increasing function for some α

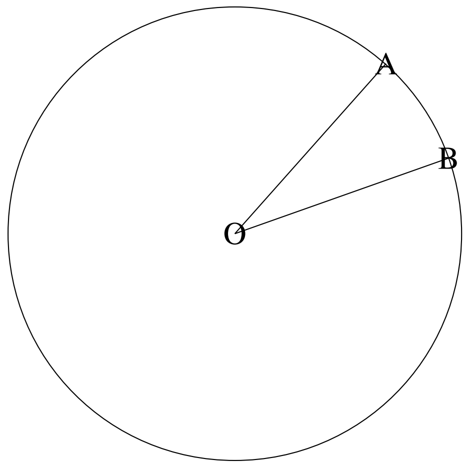
c) $\beta(x)$ is dependent of α
d) when $\alpha = 1$,

$$\beta(x) = \alpha(x-\mu)^0 = \alpha \tag{3.9.20}$$

Therefore the hazard function is independent of μ when $\alpha = 1$.

ANSWER: (2) and (4)

3.10. A point is chosen at random from a circular disc shown below. What is the probability that the point lies in the sector OAB?



(where $\angle$AOB = x radians)

a) $\frac{2x}{\pi}$
b) $\frac{x}{\pi}$
c) $\frac{x}{2\pi}$
d) $\frac{x}{4\pi}$

Solution:

Let $X \in \{0, 1\}$ be a random variable such that X=0 means we choose a point lying in sector OAB and X=1 means that we choose a point lying outside sector OAB and inside the circle.

Area of a sector subtending an angle θ at the centre of circle with radius a is given by :

$$A = \frac{1}{2}a^2\theta \tag{3.10.1}$$

where θ is in radians.

Let the radius of circle shown in figure be r. It is given that sector OAB subtends an angle of x radians at the centre of the circle. Probability that the chosen point lies in sector OAB is:

$$\Pr(X = 0) = \frac{\text{Area of sector OAB}}{\text{Area of circle}} \tag{3.10.2}$$

$$= \frac{\frac{1}{2}r^2x}{\pi r^2} \tag{3.10.3}$$

$$= \frac{x}{2\pi} \tag{3.10.4}$$

$\therefore$The correct answer is **option (3)** $\frac{x}{2\pi}$.

ALTERNATE SOLUTION

The joint pdf is given by:

$$f_{r\theta}(r,\theta) = \begin{cases} \frac{r}{\pi R^2} & \text{if } 0 < \text{r} < \text{R} \text{ , } 0 < \theta < 2\pi \\ 0 & \text{otherwise} \end{cases} \tag{3.10.5}$$

Let A $\equiv$ (R, θ_2) and B $\equiv$ (R, θ_1).
Hence,

$$(\theta_2 - \theta_1) = x \tag{3.10.6}$$

We want $\theta \in (\theta_1 , \theta_2)$ and r $\in$ (0,R) for point to lie in the sector. Let the point to be chosen be (r, θ).

So, Required probability is:

$$\Pr(\theta_1 < \theta < \theta_2, 0 < r < R) = \int_{\theta_1}^{\theta_2}\int_0^R \frac{r}{\pi R^2}\,dr\,d\theta \tag{3.10.7}$$

$$= \int_{\theta_1}^{\theta_2} \frac{1}{\pi R^2} \frac{r^2}{2}\bigg|_0^R \tag{3.10.8}$$

$$= \int_{\theta_1}^{\theta_2} \frac{R^2}{2\pi R^2}\, d\theta \tag{3.10.9}$$

$$= \int_{\theta_1}^{\theta_2} \frac{1}{2\pi}\, d\theta \tag{3.10.10}$$

$$= \frac{\theta}{2\pi}\bigg|_{\theta_1}^{\theta_2} \tag{3.10.11}$$

$$= \frac{\theta_2 - \theta_1}{2\pi} \tag{3.10.12}$$

$$= \frac{x}{2\pi} \tag{3.10.13}$$

$\therefore$The correct answer is **option (3)** $\frac{x}{2\pi}$.

3.11. Let X and Y be independent random variables each following a uniform distribution on $(0, 1)$.Let $W = XI_{\{Y \le X^2\}}$,where I_A denotes the indicator function of set A.Then which of the following statements are true?

a) The cumulative distribution function of W is given by

$$F_W(t) = t^2 I_{\{0 \le t \le 1\}} + I_{\{t > 1\}} \tag{3.11.1}$$

b) $P[W > 0] = \frac{1}{3}$

c) The cumulative distribution function of W is continuous

d) The cumulative distribution function of W is given by

$$F_W(t) = \left(\frac{2 + t^3}{3}\right) I_{\{0 \le t \le 1\}} + I_{\{t > 1\}} \tag{3.11.2}$$

Solution:

Given X and Y are two independent random variables.

Given $W = XI_{\{Y \le X^2\}}$

$X \in (0, 1)$, $Y \in (0, 1)$, $W \in [0, 1)$

a) We need to find CDF of W

i) The PDF for X is

$$p_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & otherwise \end{cases} \tag{3.11.3}$$

ii) The CDF for X is

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ x & 0 < x < 1 \\ 1 & otherwise \end{cases} \tag{3.11.4}$$

iii) The PDF for Y is

$$p_Y(y) = \begin{cases} 1 & 0 < y < 1 \\ 0 & otherwise \end{cases} \tag{3.11.5}$$

iv) The CDF for Y is

$$F_Y(y) = \begin{cases} 0 & y \le 0 \\ y & 0 < y < 1 \\ 1 & otherwise \end{cases} \tag{3.11.6}$$

v) $I_{\{Y \le X^2\}}$ is defined as follows

$$I_{\{Y \le X^2\}} = \begin{cases} 1 & y \le x^2 \\ 0 & otherwise \end{cases} \tag{3.11.7}$$

vi) W is defined as follows

$$W = \begin{cases} x & y \le x^2 \\ 0 & otherwise \end{cases} \tag{3.11.8}$$

From (3.11.8)

$$p_W(W = 0) = \Pr(I_{\{Y \le X^2\}} = 0) \tag{3.11.9}$$

$$= \Pr(x^2 < y) \tag{3.11.10}$$

vii) Let $Z = X^2 - Y$ be a random variable where $Z \in (-1, 1)$

$$F_{X^2}(u) = \Pr(X^2 \le u) \tag{3.11.11}$$

$$= \Pr(X \le \sqrt{u}) \tag{3.11.12}$$

$$= F_X(\sqrt{u}) \tag{3.11.13}$$

A) From (3.11.4),The CDF for X^2 is

$$F_{X^2}(u) = \begin{cases} 0 & u \le 0 \\ \sqrt{u} & 0 < u < 1 \\ 1 & otherwise \end{cases} \tag{3.11.14}$$

B) The PDF for X^2 is

$$p_{X^2}(u) = \begin{cases} \frac{1}{2\sqrt{u}} & 0 < u < 1 \\ 0 & otherwise \end{cases} \tag{3.11.15}$$

$$F_{\{-Y\}}(v) = \Pr(-Y \le v) \tag{3.11.16}$$

$$= \Pr(Y \ge -v) \tag{3.11.17}$$

$$= 1 - F_Y(-v) \tag{3.11.18}$$

C) From (3.11.6),The CDF for $(-Y)$ is

$$F_{\{-Y\}}(v) = \begin{cases} 0 & v \le -1 \\ 1 + v & -1 < v < 0 \\ 1 & otherwise \end{cases} \tag{3.11.19}$$

D) The PDF for $(-Y)$ is

$$p_{\{-Y\}}(v) = \begin{cases} 1 & -1 < v < 0 \\ 0 & otherwise \end{cases} \tag{3.11.20}$$

E) $Z = X^2 - Y \implies z = u + v$
Using convolution

$$p_Z(z) = \int_{-\infty}^{\infty} p_{X^2}(z - v) p_{\{-Y\}}(v) \mathrm{dv} \tag{3.11.21}$$

Solving (3.11.21) using (3.11.20),(3.11.15) for $z \in (-1, 1)$, we get PDF of Z as follows

$$p_Z(z) = \begin{cases} \sqrt{z + 1} & -1 < z \le 0 \\ 1 - \sqrt{z} & 0 < z < 1 \\ 0 & otherwise \end{cases} \tag{3.11.22}$$

F) CDF of Z as follows

$$F_Z(z) = \begin{cases} \frac{2}{3}(z+1)^{\frac{3}{2}} & -1 < z \leq 0 \\ z - \frac{2}{3}z^{\frac{3}{2}} & 0 < z < 1 \\ 1 & otherwise \end{cases} \tag{3.11.23}$$

viii) using (3.11.23) to find $p_W(W=0)$

$$p_W(W=0) = \Pr(x^2 < y) \tag{3.11.24}$$

$$= F_z(0) \tag{3.11.25}$$

$$= \frac{2}{3} \tag{3.11.26}$$

ix) $W = t \implies X = t$ where $t \in (0,1)$

$$p_W(t) = \int_{-\infty}^{\infty} p_X(t) I_{\{y \leq t^2\}} \mathrm{dy} \tag{3.11.27}$$

$$0 < y < 1 \tag{3.11.28}$$

$$0 < y \leq t^2 \tag{3.11.29}$$

For $0 < t < 1$,

$$p_W(t) = \int_{0}^{t^2} p_X(t) I_{\{y \leq t^2\}} \mathrm{dy} \tag{3.11.30}$$

$$= t^2 \tag{3.11.31}$$

x) $\therefore$ PDF of W is as follows

$$p_W(t) = \begin{cases} \frac{2}{3} & t = 0 \\ t^2 & 0 < t < 1 \\ 0 & otherwise \end{cases} \tag{3.11.32}$$

xi) The CDF of W is as follows:

$$F_W(t) = \begin{cases} 0 & t < 0 \\ \frac{2+t^3}{3} & 0 \leq t \leq 1 \\ 1 & otherwise \end{cases} \tag{3.11.33}$$

b) We need to find $P[W > 0]$

$$\Pr(W > 0) = 1 - F_W(0) \tag{3.11.34}$$

$$= \frac{1}{3} \tag{3.11.35}$$

$$\therefore \Pr(W > 0) = \frac{1}{3} \tag{3.11.36}$$

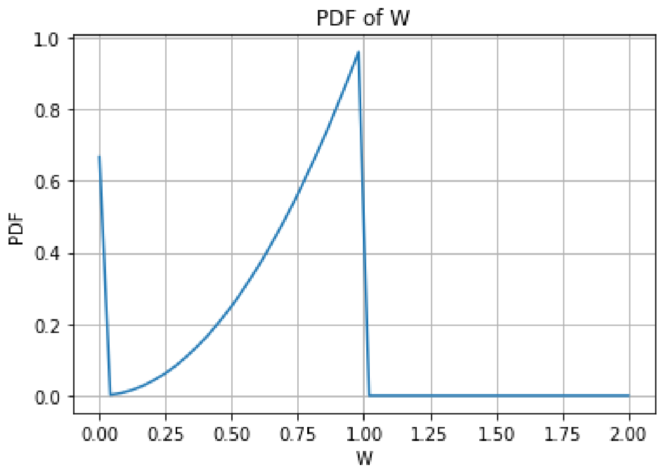
c) CDF of W is discontinuous at $W = 0$.
$\therefore$ option 3 is incorrect.

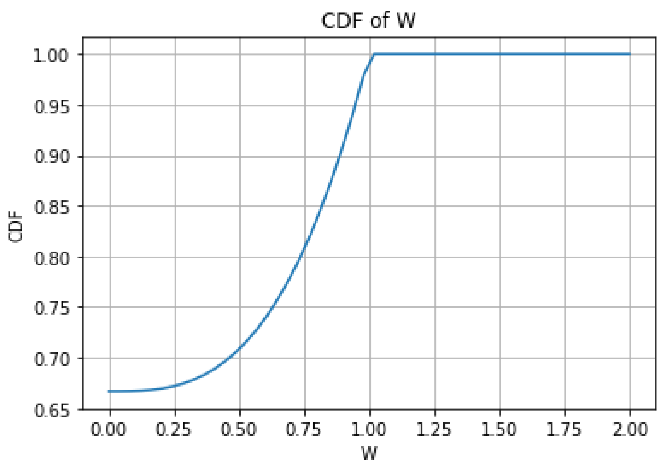
d) The CDF in (3.11.33) can be written as

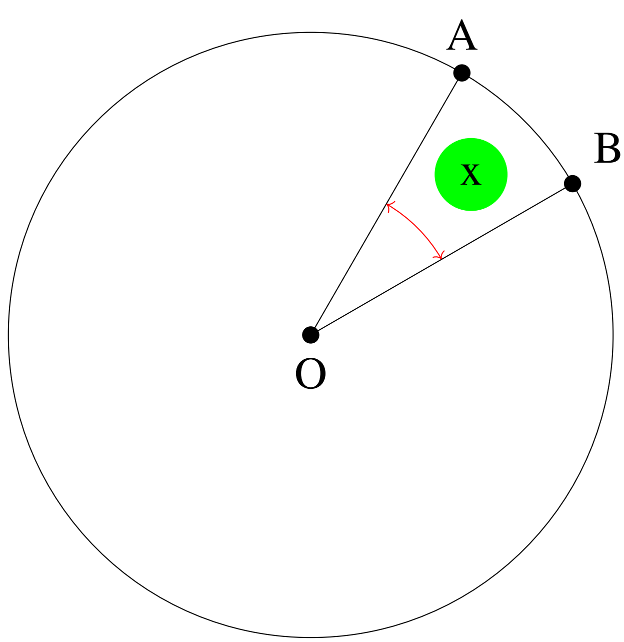
$$F_W(t) = \left(\frac{2+t^3}{3}\right) I_{\{0 \leq t \leq 1\}} + I_{\{t > 1\}} \tag{3.11.37}$$

$\therefore$ option 2 and 4 are correct.

3.12. A point is chosen at random from a circular disc shown below. What is the probability that the point lies in the sector OAB?







(where ∠AOB = x radians)

a) $\frac{2x}{\pi}$
b) $\frac{x}{\pi}$
c) $\frac{x}{2\pi}$
d) $\frac{x}{4\pi}$

Solution:

Let $X \in \{0, 1\}$ be a random variable such that X=0 means we choose a point lying in sector OAB and X=1 means that we choose a point lying outside sector OAB and inside the circle. Area of a sector subtending an angle θ at the centre of circle with radius a is given by :

$$A = \frac{1}{2}a^2\theta \tag{3.12.1}$$

where θ is in radians.
Let the radius of circle shown in figure be r. It is given that sector OAB subtends an angle of x radians at the centre of the circle. Probability that the chosen point lies in sector OAB is:

$$\Pr(X = 0) = \frac{\text{Area of sector OAB}}{\text{Area of circle}} \tag{3.12.2}$$

$$= \frac{\frac{1}{2}r^2x}{\pi r^2} \tag{3.12.3}$$

$$= \frac{x}{2\pi} \tag{3.12.4}$$

∴The correct answer is **option (3)** $\frac{x}{2\pi}$. **alternate solution** The joint pdf is given by:

$$f_{r\theta}(r,\theta) = \begin{cases} \frac{r}{\pi R^2} & \text{if } 0 < r < R \text{ , } 0 < \theta < 2\pi \\ 0 & \text{otherwise} \end{cases} \tag{3.12.5}$$

Let A ≡ (R, θ_2) and B ≡ (R, θ_1). Hence,

$$(\theta_2 - \theta_1) = x \tag{3.12.6}$$

We want $\theta \in (\theta_1 , \theta_2)$ and r ∈ (0,R) for point to lie in the sector. Let the point to be chosen be (r, θ).

So, Required probability is:

$$\Pr(\theta_1 < \theta < \theta_2, 0 < r < R)$$

$$= \int_{\theta_1}^{\theta_2} \int_0^R \frac{r}{\pi R^2} \, dr \, d\theta \tag{3.12.7}$$

$$= \int_{\theta_1}^{\theta_2} \frac{1}{\pi R^2} \frac{r^2}{2} \bigg|_0^R \tag{3.12.8}$$

$$= \int_{\theta_1}^{\theta_2} \frac{R^2}{2\pi R^2} \, d\theta \tag{3.12.9}$$

$$= \int_{\theta_1}^{\theta_2} \frac{1}{2\pi} \, d\theta \tag{3.12.10}$$

$$= \frac{\theta}{2\pi} \bigg|_{\theta_1}^{\theta_2} \tag{3.12.11}$$

$$= \frac{\theta_2 - \theta_1}{2\pi} \tag{3.12.12}$$

$$= \frac{x}{2\pi} \tag{3.12.13}$$

∴The correct answer is **option (3)** $\frac{x}{2\pi}$.

4 Transformation of Variables

4.1. Let X be a random variable with pdf

$$f_X(x) = \begin{cases} \frac{2x}{\pi^2} & 0 < x < \pi \\ 0 & \text{otherwise} \end{cases} \tag{4.1.1}$$

Let Y = $\sin X$, then for $0 < y < 1$, the pdf of Y is given by,

(A) $\frac{2\pi}{\sqrt{1-y^2}}$

(B) $\frac{\pi}{2}\sqrt{1-y^2}$

(C) $\frac{2}{\pi}\sqrt{1-y^2}$

(D) $\frac{2}{\pi\sqrt{1-y^2}}$

Solution: From the given information,

$$F_X(x) = \Pr(X \le x) \quad (4.1.2)$$
$$= \begin{cases} 0 & x \le 0 \\ \frac{x^2}{\pi^2} & 0 < x < \pi \\ 1 & x \ge \pi \end{cases} \quad (4.1.3)$$

after integration. Consequently,

$$F_Y(y) = \Pr(Y \le y) \quad (4.1.4)$$
$$= \Pr(\sin X \le y) \quad (4.1.5)$$

From Fig. 4.1.1,

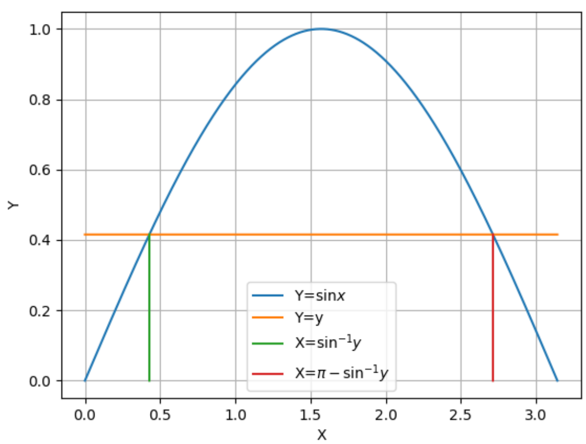


Fig. 4.1.1: $Y = \sin X$ plot

$$\sin X \le y$$
$$\implies \left\{X \le \sin^{-1} y \cup X \ge \pi - \sin^{-1} y\right\} \quad (4.1.6)$$

$$\implies F_Y(y) = \Pr\left(X \le \sin^{-1} y\right) + \Pr\left(X \ge \pi - \sin^{-1} y\right) \quad (4.1.7)$$
$$= F_X(\sin^{-1} y) + 1 - \Pr\left(X \le \pi - \sin^{-1} y\right) \quad (4.1.8)$$

$$\implies F_Y(y) = 1 + F_X(\sin^{-1} y) - F_X(\pi - \sin^{-1} y) \quad (4.1.9)$$

Substituting from (4.1.3) in (4.1.9)

$$F_Y(y) = \frac{\left(\sin^{-1} y\right)^2}{\pi^2} + 1 - \frac{\left(\pi - \sin^{-1} y\right)^2}{\pi^2} \quad (4.1.10)$$
$$= \frac{2\sin^{-1} y}{\pi} \quad (4.1.11)$$

$$\therefore f_Y(y) = \frac{dF_Y(y)}{dy} \quad (4.1.12)$$
$$= \frac{2}{\pi\sqrt{1-y^2}} \quad (4.1.13)$$

Hence, option(D) is correct.

5 Independent Random Variables

5.1. Let $X \in \{0, 1\}$ and $Y \in \{0, 1\}$ be two independent binary random variables. If $\Pr(X = 0) =$ p and $\Pr(Y = 0) =$ q, then $\Pr(X + Y \geqslant 1))$ is equal to
  a) $pq + (1-p)(1-q)$
  b) pq
  c) $p(1-q)$
  d) $1 - pq$

Solution:

5.2. Two independent random variables X and Y are uniformly distributed in the interval $[-1, 1]$. The probability that $\max(X, Y)$ is less than $\frac{1}{2}$ is

a) $\frac{3}{4}$ b) $\frac{9}{16}$ c) $\frac{1}{4}$ d) $\frac{2}{3}$

Solution:

5.3. Suppose that $X_1, X_2, X_3, ..., X_{10}$ are i.i.d, N(0,1). Which of the following statements are correct ?
  (A) $\Pr(X_1 > X_2 + X_3 + ... + X_{10}) = \frac{1}{2}$
  (B) $\Pr(X_1 > X_2X_3...X_{10}) = \frac{1}{2}$
  (C) $\Pr(\sin X_1 > \sin X_2 + \sin X_3 + ... + \sin X_{10}) = \frac{1}{2}$
  (D) $\Pr(\sin X_1 > \sin(X_2 + X_3 + \ldots + X_{10})) = \frac{1}{2}$

Solution:

Lemma 5.1. *If $X \sim \mathcal{N}(0, 1)$ then $Y = -X$ also follows standard normal distribution.*

Proof.

$$
\begin{aligned}
P(Y \leq u) &= P(-X \leq u) && (5.3.1)\\
&= P(X > -u) && (5.3.2)\\
&= 1 - P(X \leq -u) && (5.3.3)\\
&= 1 - (1 - P(X \leq u) && (5.3.4)\\
&= P(X \leq u) && (5.3.5)
\end{aligned}
$$

As the distribution is symmetric,

$$P(X \leq -u) = P(X \geq u) = 1 - P(X \leq u) \tag{5.3.6}$$

□

Lemma 5.2. *If n is an even number and $g(x)$ is an odd function, then,*

a)

$$\Pr\left(g(X_1) > \sum_{k=2}^{n} g(X_k)\right) = \Pr\left(g(X_1) < \sum_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \tag{5.3.7}$$

b)

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \Pr\left(g(X_1) < \prod_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \tag{5.3.8}$$

Proof. a)

$$\Pr\left(g(X_1) > \sum_{k=2}^{n} g(X_k)\right) = \Pr\left(g(-X_1) < \sum_{k=2}^{n} g(-X_k)\right) = \Pr\left(g(X_1) < \sum_{k=2}^{n} g(X_k)\right) \tag{5.3.9}$$

As the cases

$$g(X_1) > \sum_{k=2}^{n} g(X_k) \tag{5.3.10}$$

and

$$g(X_1) < \sum_{k=2}^{n} g(X_k) \tag{5.3.11}$$

are complementary to each other,

$$\Pr\left(g(X_1) > \sum_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \tag{5.3.12}$$

b) Similarly,

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \Pr\left(g(-X_1) < \prod_{k=2}^{n} g(-X_k)\right) = \Pr\left(g(X_1) < \prod_{k=2}^{n} g(X_k)\right) \tag{5.3.13}$$

As they follow the same distribution, the above expression is true. Thus we have

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \Pr\left(g(X_1) < \prod_{k=2}^{n} g(X_k)\right) \tag{5.3.14}$$

As the cases

$$g(X_1) > \prod_{k=2}^{n} g(X_k) \tag{5.3.15}$$

and

$$g(X_1) < \prod_{k=2}^{n} g(X_k) \tag{5.3.16}$$

are complementary to each other and from (5.3.7) we have

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \tag{5.3.17}$$

(A) From (5.3.12) , taking $g(x) = x$,

$$\Pr(X_1 > X_2 + ... + X_{10}) = \frac{1}{2} \tag{5.3.18}$$

(B) From (5.3.17) taking $g(x) = x$

$$\Pr(X_1 > X_2X_3...X_{10}) = \frac{1}{2} \tag{5.3.19}$$

(C) From (5.3.12) taking $g(x) = \sin x$

$$\Pr(\sin X_1 > \sin X_2 + ... + \sin X_{10}) = \frac{1}{2} \tag{5.3.20}$$

(D)

$$\begin{aligned}&\Pr(\sin X_1 > \sin(X_2 + ... + X_{10})) \\ &= \Pr(\sin(-X_1) < \sin(-X_2 - ... - X_{10})) \\ &= \Pr(\sin X_1 < \sin(X_2 + ... + X_{10}))\end{aligned} \tag{5.3.21}$$

As they follow the same distribution, the above expression is true. Thus we have

$$\begin{aligned}&\Pr(\sin X_1 > \sin(X_2 + ... + X_{10})) \\ &= \Pr(\sin X_1 < \sin(X_2 + ... + X_{10}))\end{aligned} \tag{5.3.22}$$

Also, as X_1 is a continuous random variable

$$\Pr(\sin X_1 = \sin(X_2 + ... + X_{10})) = 0 \tag{5.3.23}$$

As the cases

$$X_1 > X_2 + ... + X_{10} \tag{5.3.24}$$

and

$$X_1 < X_2 + ... + X_{10} \tag{5.3.25}$$

are complementary to each other

$$\Pr(\sin X_1 > \sin(X_2 + ... + X_{10})) = \frac{1}{2} \tag{5.3.26}$$

□

5.4. Which of the following conditions imply independence of the random variables X and Y ?

a) $\Pr(X > a|Y > a) = \Pr(X > a)\ \forall\ a \in \mathbb{R}$

b) $\Pr(X > a|Y < b) = \Pr(X > a)\ \forall\ a,\ b \in \mathbb{R}$

c) X and Y are uncorrelated.

d) $E[(X-a)(Y-b)] = E(X-a)\ E(Y-b)\ \forall\ a,\ b \in \mathbb{R}$

Solution:

Definition 1. *Two random variables X and Y are independent when the joint probability distribution of random variables is product of their individual probability distributions i.e for all sets A,B*

$$\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B) \tag{5.4.1}$$

Alternatively,

$$F_{X,Y}(a,b) = F_X(a)\,F_Y(b) \tag{5.4.2}$$

Lemma 5.3. *From* (5.4.2)*, it follows that*

$$\implies f_{X,Y}(a,b) = f_X(a)\,f_Y(b) \tag{5.4.3}$$

Proof. From (5.4.2),

$$\frac{\partial^2 F_{X,Y}(a,b)}{\partial b \partial a} = \frac{\partial F_X(a)}{\partial a}\frac{\partial F_Y(b)}{\partial b} \tag{5.4.4}$$

yielding (5.4.3). □

a) From the given information,

$$\Pr(X > a, Y > a) = \Pr(X > a)\Pr(Y > a) \tag{5.4.5}$$

$$= [1 - F_X(a)]\,[1 - F_Y(a)] \tag{5.4.6}$$

$$\begin{aligned}&\because \Pr(X > a) - \Pr(Y < a) \\ &= \Pr(X > a, Y > a) + \Pr(X > a, Y < a) \\ &- \Pr(X > a, Y < a) - \Pr(X < a, Y < a), \\ &\Pr(X > a, Y > a) = 1 - F_X(a) - F_Y(a) \\ &\qquad + F_{X,Y}(a,a)\end{aligned} \tag{5.4.7}$$

which, upon substituting from (5.4.6) yields

$$\implies F_{X,Y}(a,a) = F_X(a)\,F_Y(a) \tag{5.4.8}$$

which is a special case of (5.4.1) for $b = a$. The spectrum of conditions for independence is hence underrepresented. Thus, the given condition does not imply independence of X and Y.
Option 1 is incorrect.

b) From Bayes theorem,

$$\Pr(X > a|Y < b) = \Pr(X > a) \tag{5.4.9}$$

$$\begin{aligned}\implies &\Pr(X > a, Y < b) \\ &= \Pr(X > a)\Pr(Y < b)\end{aligned} \tag{5.4.10}$$

for all $a, b \in R$.

$$\because F_Y(b) = \Pr(X > a, Y < b) + \Pr(X < a, Y < b), \quad (5.4.11)$$

$$\begin{aligned} \Pr(X > a, Y < b) &= F_Y(b) - F_{X,Y}(a,b) \\ \implies F_Y(b) - F_{X,Y}(a,b) &= (1 - F_X(a)) F_Y(b) \\ \text{or, } F_{X,Y}(a,b) &= F_X(a) F_Y(b) \quad (5.4.12) \end{aligned}$$

upon substituting from (5.4.10) and simplifying. Thus, X and Y are independent.
Option 2 is correct.

c) We prove through a counterexample.

Definition 2. *Two random variables X and Y are uncorrelated if their covariance is zero, i.e.,*

$$cov(X,Y) = E(XY) - E(X)E(Y) = 0 \quad (5.4.13)$$

Let $X \sim U[-1,1]$ be a uniformly distributed random variable such that

$$f_X(x) = \begin{cases} \frac{1}{2} & -1 \le x \le 1 \\ 0 & otherwise \end{cases} \quad (5.4.14)$$

$$\mathrm{E}(X) = \int_{-1}^{1} x f(x)\, dx = 0 \quad (5.4.15)$$

Let

$$Y = X^2. \quad (5.4.16)$$

so that X and Y are dependent. Then,

$$\begin{aligned} \mathrm{cov}(X,Y) &= \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) \quad (5.4.17) \\ &= \mathrm{E}\left(X^3\right) - 0 \times \mathrm{E}(Y) \quad (5.4.18) \\ &= \int_{-1}^{1} x^3 f(x)\, dx = 0 \quad (5.4.19) \end{aligned}$$

X and Y are uncorrelated but not independent.
Option 3 is incorrect

d) Given that,

$$\mathrm{E}((X-a)(Y-b)) = \mathrm{E}(X-a)\mathrm{E}(Y-b) \quad (5.4.20)$$

$$\implies \mathrm{cov}(X-a, Y-b) = \mathrm{E}((X-a)(Y-b)) - \mathrm{E}(X-a)\mathrm{E}(Y-b) \quad (5.4.21)$$

$$\text{or, } \mathrm{cov}(X-a, Y-b) = 0 = \mathrm{cov}(X,Y) \quad (5.4.22)$$

From option 3, it follows that X and Y are not necessarily independent.
Option 4 is incorrect.

5.5. Let X and Y be two independent and identically distributed (I.I.D) random variables uniformly distributed in (0,1). Let $Z = max(X,Y)$ and $W = min(X,Y)$, then the probability that $[Z - W > \frac{1}{2}]$ is

(A) $\frac{1}{2}$

(B) $\frac{3}{4}$

(C) $\frac{1}{4}$

(D) $\frac{2}{3}$ **Solution:**

X and Y are two independent random variables. Let

$$f_X(x) = \Pr(X = x) \quad (5.5.1)$$
$$f_Y(y) = \Pr(Y = y) \quad (5.5.2)$$
$$f_V(v) = \Pr(V = v) \quad (5.5.3)$$

be the probability densities of random variables X ,Y and V=X-Y.
The density for X is

$$f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & otherwise \end{cases} \quad (5.5.4)$$

We have ,

$$V = X - Y \iff v = x - y \iff x = v + y \quad (5.5.5)$$

The density of X can also be represented as,

$$f_X(v+y) = \begin{cases} 1 & 0 \le v + y \le 1 \\ 0 & otherwise \end{cases} \quad (5.5.6)$$

and the density of Y is,

$$f_Y(y) = \begin{cases} 1 & 0 \le y \le 1 \\ 0 & otherwise \end{cases} \quad (5.5.7)$$

The density of V i.e. $V = X - Y$ is given by the convolution of $f_X(-v)$ with $f_Y(v)$.

$$f_V(v) = \int_{-\infty}^{\infty} f_X(v+y) f_Y(y)\, dy \tag{5.5.8}$$

From 5.5.6 and 5.5.7 we have,
The integrand is 1 when,

$$0 \le y \le 1 \tag{5.5.9}$$
$$0 \le v + y \le 1 \tag{5.5.10}$$
$$-v \le y \le 1 - v \tag{5.5.11}$$

and zero, otherwise.
Now when $-1 \le v \le 0$ we have,

$$f_V(v) = \int_{-v}^{1} dy \tag{5.5.12}$$
$$= (1 - (-v)) \tag{5.5.13}$$
$$= 1 + v \tag{5.5.14}$$

For $0 \le v \le 1$ we have,

$$f_V(v) = \int_{0}^{1-v} dy \tag{5.5.15}$$
$$= (1 - v - (0)) \tag{5.5.16}$$
$$= 1 - v \tag{5.5.17}$$

Therefore the density of V is given by

$$f_V(v) = \begin{cases} 1 + v & -1 \le v \le 0 \\ 1 - v & 0 < v \le 1 \\ 0 & otherwise \end{cases} \tag{5.5.18}$$

The plot for PDF of V can be observed at figure 5.5.1

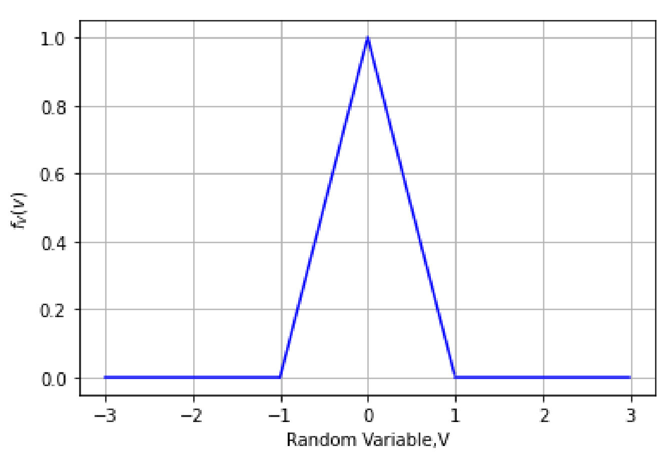


Fig. 5.5.1: The PDF of V

The CDF of V is defined as,

$$F_V(v) = \Pr(V \le v) \tag{5.5.19}$$

Now for $v \le 0$,

$$\Pr(V \le v) = \int_{-\infty}^{v} f_V(v)\, dv \tag{5.5.20}$$
$$= \int_{-1}^{v} (1 + v)\, dv \tag{5.5.21}$$
$$= \left(\frac{v^2}{2} + v\right)\Bigg|_{-1}^{v} \tag{5.5.22}$$
$$= \left(\left(\frac{v^2}{2} + v\right) - \left(\frac{1}{2} - 1\right)\right) \tag{5.5.23}$$
$$= \frac{v^2 + 2v + 1}{2} \tag{5.5.24}$$

Similarly for $v \le 1$,

$$\Pr(V \le v) = \int_{-\infty}^{v} f_V(v)\, dv \tag{5.5.25}$$
$$= \frac{1}{2} + \int_{0}^{v} (1 - v)\, dz \tag{5.5.26}$$
$$= \frac{-v^2 + 2v + 1}{2} \tag{5.5.27}$$

The CDF is as below:

$$F_V(v) = \begin{cases} 0 & v < -1 \\ \dfrac{v^2 + 2v + 1}{2} & v \le 0 \\ \dfrac{-v^2 + 2v + 1}{2} & v \le 1 \\ 1 & v > 1 \end{cases} \tag{5.5.28}$$

The plot for CDF of V can be observed at figure 5.5.2

We need $\Pr\left(Z - W > \frac{1}{2}\right)$ where $Z = max(X, Y)$ and $W = min(X, Y)$. Now,

$$Z - W = \begin{cases} X - Y & \text{for } X \ge Y \\ Y - X & \text{for } X < Y \end{cases} \tag{5.5.29}$$

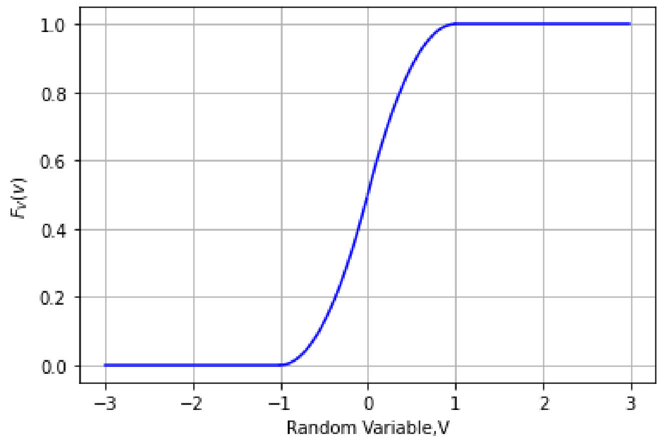


Fig. 5.5.2: The CDF of V

Therefore,

$$\Pr\left(Z - W > \frac{1}{2}\right) = \Pr\left(X - Y > \frac{1}{2}, X \geq Y\right) + \Pr\left(Y - X > \frac{1}{2}, X < Y\right) \quad (5.5.30)$$

$$= \Pr\left(X - Y > \frac{1}{2}\right) + \Pr\left(Y - X > \frac{1}{2}\right) \quad (5.5.31)$$

$$= \Pr\left(V > \frac{1}{2}\right) + \Pr\left(-V > \frac{1}{2}\right) \quad (5.5.32)$$

$$= 1 - \Pr\left(V \leq \frac{1}{2}\right) + \Pr\left(V < \frac{-1}{2}\right) \quad (5.5.33)$$

$$= 1 - F_V(\frac{1}{2}) + F_V(-\frac{1}{2}) \quad (5.5.34)$$

$$= 1 - \frac{7}{8} + \frac{1}{8} \quad (5.5.35)$$

$$= \frac{1}{4} \quad (5.5.36)$$

Hence the correct answer is option (C).

5.6. Let X and Y be independent exponential random variables. If $E[X] = 1$ and $E[Y] = \frac{1}{2}$ then $\Pr(X > 2Y | X > Y)$ is

1. $\frac{1}{2}$
2. $\frac{1}{3}$
3. $\frac{2}{3}$
4. $\frac{3}{4}$

Solution: Since X and Y are exponential random variables with means'

$$E[X] = 1 \text{ and } E[Y] = \frac{1}{2} \quad (5.6.1)$$

Marginal PDFs of X and Y are given by

$$f_X(x) = e^{-x}, x > 0 \quad (5.6.2)$$

$$f_Y(y) = 2e^{-2y}, y > 0 \quad (5.6.3)$$

CDFs for X and Y are

$$F_X(b) = \int_0^b f_X(x)\, d_x \quad (5.6.4)$$

$$= \int_0^b e^{-x}\, d_x \quad (5.6.5)$$

$$= 1 - e^{-b} \quad (5.6.6)$$

$$F_Y(b) = \int_0^b f_Y(y)\, d_y \quad (5.6.7)$$

$$= \int_0^b 2e^{-2y}\, d_y \quad (5.6.8)$$

$$= \left[-e^{-2y}\right]_0^b \quad (5.6.9)$$

$$= 1 - e^{-2b} \quad (5.6.10)$$

Now,

$$\Pr(X > 2Y | X > Y) = \frac{\Pr(X > 2Y\,,\, X > Y)}{\Pr(X > Y)} \quad (5.6.11)$$

$$= \frac{\Pr(X > 2Y)}{\Pr(X > Y)} \quad (5.6.12)$$

$$\Pr(X > Y) = \Pr(Y < X) \quad (5.6.13)$$

$$= E[F_Y(X)] \quad (5.6.14)$$

$$= \int_0^\infty F_Y(X)\, f_X(x)\, d_x \quad (5.6.15)$$

$$= \int_0^\infty (1 - e^{-2x})\, e^{-x} d_x \quad (5.6.16)$$

$$= \left[\frac{e^{-x}}{-1} - \frac{e^{-3x}}{-3}\right]_0^\infty \quad (5.6.17)$$

$$= (0 + 1) + \frac{1}{3}(0 - 1) \quad (5.6.18)$$

$$= \frac{2}{3} \quad (5.6.19)$$

$$\Pr(X > 2Y) = \Pr\left(Y < \frac{X}{2}\right) \tag{5.6.20}$$
$$= E[F_Y(X/2)] \tag{5.6.21}$$
$$= \int_0^\infty F_Y(X/2)\, f_X(x)\, d_x \tag{5.6.22}$$
$$= \int_0^\infty (1 - e^{-x})\, e^{-x} d_x \tag{5.6.23}$$
$$= \left[\frac{e^{-x}}{-1} - \frac{e^{-2x}}{-2}\right]_0^\infty \tag{5.6.24}$$
$$= (0 + 1) + \frac{1}{2}(0 - 1) \tag{5.6.25}$$
$$= \frac{1}{2} \tag{5.6.26}$$

Putting (5.6.19) and (5.6.26) in (5.6.12)

$$\Pr(X > 2Y \,|\, X > Y) = \frac{1/2}{2/3} \tag{5.6.27}$$
$$= \frac{3}{4} \tag{5.6.28}$$

$\therefore$ Option 4 is the correct answer.

5.7. A fair die is thrown two times independently. Let X, Y be the outcomes of these two throws and $Z = X+Y$. Let U be the remainder obtained when Z is divided by 6. Then which of the following statement(s) is/are true?

a) X and Z are independent
b) X and U are independent
c) Z and U are independent
d) Y and Z are not independent

Solution: Let $X \in \{1, 2, 3, 4, 5, 6\}$ represent the random variable which represents the outcome of the first throw of a dice. Similarly, $Y \in \{1, 2, 3, 4, 5, 6\}$ represents the random variable which represents the outcome of the second throw of a dice.

$$n(X = i) = 1, \quad i \in \{1, 2, 3, 4, 5, 6\} \tag{5.7.1}$$

$$\Pr(X = i) = \begin{cases} \frac{1}{6} & i \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7.2}$$

Similarly,

$$\Pr(Y = i) = \begin{cases} \frac{1}{6} & i \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7.3}$$

$$Z = X + Y \tag{5.7.4}$$
$$\text{Let } z \in \{1, 2, \dots, 11, 12\} \tag{5.7.5}$$
$$\Pr(Z = z) = \Pr(X + Y = z) \tag{5.7.6}$$
$$= \sum_{x=0}^{z} \Pr(X = x) \Pr(Y = z - x) \tag{5.7.7}$$
$$= (6 - |z - 7|) \times \frac{1}{6} \times \frac{1}{6} \tag{5.7.8}$$
$$= \frac{6 - |z - 7|}{36} \tag{5.7.9}$$

$$\Pr(Z = z) = \begin{cases} \frac{6-|z-7|}{36} & z \in \{1, 2, \dots, 11, 12\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7.10}$$

U is the remainder obtained when Z is divided by 6.

$$\text{Let } u \in \{0, 1, 2, 3, 4, 5\} \tag{5.7.11}$$

$$\Pr(U = u) = \sum_{k=0}^{2} \Pr(Z = 6k + u) \tag{5.7.12}$$

$$\Pr(U = 0) = \Pr(Z = 0) + \Pr(Z = 6) + \Pr(Z = 12) \tag{5.7.13}$$
$$= 0 + \frac{5}{36} + \frac{1}{36} = \frac{1}{6} \tag{5.7.14}$$
$$\text{for } u \in \{1, 2, 3, 4, 5\} \tag{5.7.15}$$
$$\Pr(U = u) = \Pr(Z = 0 + u) + \Pr(Z = 6 + u) \tag{5.7.16}$$
$$= \frac{6 - |u - 7|}{36} + \frac{6 - |6 + u - 7|}{36} \tag{5.7.17}$$
$$= \frac{6 - (7 - u)}{36} + \frac{6 - (u - 1)}{36} \tag{5.7.18}$$
$$= \frac{u - 1 + 7 - u}{36} = \frac{6}{36} \tag{5.7.19}$$
$$= \frac{1}{6} \tag{5.7.20}$$

$$\Pr(U = u) = \begin{cases} \frac{1}{6} & u \in \{0, 1, 2, 3, 4, 5\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7.21}$$

Now, for checking each option,

a) Checking if X and Z are independent

$$p_1 = \Pr(Z = z, X = x) \tag{5.7.22}$$
$$= \Pr(Y = z - x, X = x) \tag{5.7.23}$$
$$= \Pr(Y = z - x) \times \Pr(X = x) \tag{5.7.24}$$
$$= \begin{cases} \frac{1}{36} & z - x \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7.25}$$

$$\Pr(Z = z) \times \Pr(X = x) = \frac{6 - |z - 7|}{36} \times \frac{1}{6} \tag{5.7.26}$$
$$= \frac{6 - |z - 7|}{216} \tag{5.7.27}$$
$$\Pr(Z = z)\Pr(X = x) \neq \Pr(Z = z, X = x) \tag{5.7.28}$$

X and Z are not independent from (5.7.28) and hence option (5.7a) is false.

b) Checking if X and U are independent

$$p_2 = \Pr(U = u, X = x) \tag{5.7.29}$$

$$p_2 = \Pr((Z = u) + (Z = 6 + u) + (Z = 12 + u), X = x) \tag{5.7.30}$$

$$p_2 = \Pr((Y = u - x) + (Y = 6 + u - x) + (Y = 12 + u - x), X = x) \tag{5.7.31}$$

$$p_2 = \frac{1}{6} \times \frac{1}{6} \tag{5.7.32}$$
$$= \frac{1}{36} \tag{5.7.33}$$

$$\Pr(U = u) \times \Pr(X = x) = \frac{1}{6} \times \frac{1}{6} \tag{5.7.34}$$
$$= \frac{1}{36} \tag{5.7.35}$$
$$\Pr(U = u)\Pr(X = x) = \Pr(U = u, X = x) \tag{5.7.36}$$

X and U are independent from (5.7.36) and hence option (5.7b) is true.

c) Checking if Z and U are independent

$$p_3 = \Pr(Z = z|U = u) \tag{5.7.37}$$

$$p_3 = \begin{cases} 1 & u = 1 \text{ and } z = 7 \\ \frac{1}{2} & u = 0 \text{ and } z \in \{6, 12\} \\ \frac{1}{2} & u \in \{2, 3, 4, 5\} \text{ and } \\ & z = u \text{ or } z = 6 + u \\ 0 & \text{otherwise} \end{cases} \tag{5.7.38}$$

$$\Pr(Z = z) = \frac{6 - |z - 7|}{36} \tag{5.7.39}$$

If Z and U are independent, then

$$\Pr(Z = z|U = u) = \frac{\Pr(Z = z, U = u)}{\Pr(U = u)} \tag{5.7.40}$$
$$= \frac{\Pr(Z = z)\Pr(U = u)}{\Pr(U = u)} \tag{5.7.41}$$
$$= \Pr(Z = z) \tag{5.7.42}$$

But,

$$\Pr(Z = z|U = u) \neq \Pr(Z = z) \tag{5.7.43}$$

X and U are not independent from (5.7.43) and hence option (5.7c) is false.

d) Checking if Y and Z are independent

$$p_1 = \Pr(Z = z, Y = y) \tag{5.7.44}$$
$$= \Pr(X = z - y, Y = y) \tag{5.7.45}$$
$$= \Pr(X = z - y) \times \Pr(Y = y) \tag{5.7.46}$$
$$= \begin{cases} \frac{1}{36} & z - y \in \{1, 2, 3, 4, 5, 6\} \\ 0 & \text{otherwise} \end{cases} \tag{5.7.47}$$

$$\Pr(Z = z) \times \Pr(Y = y) = \frac{6 - |z - 7|}{36} \times \frac{1}{6} \tag{5.7.48}$$
$$= \frac{6 - |z - 7|}{216} \tag{5.7.49}$$
$$\Pr(Z = z)\Pr(Y = y) \neq \Pr(Z = z, Y = y) \tag{5.7.50}$$

X and Z are not independent from (5.7.50) and hence option (5.7d) is true.

Thus, options (5.7b) and (5.7d) are true.

5.8. Suppose X_1,X_2,X_3 and X_4 are independent and identically distributed random variables, having density function f. Then,

a) $\Pr(X_4 > Max(X_1, X_2) > X_3) = \frac{1}{6}$
b) $\Pr(X_4 > Max(X_1, X_2) > X_3) = \frac{1}{8}$
c) $\Pr(X_4 > X_3 > Max(X_1, X_2)) = \frac{1}{12}$
d) $\Pr(X_4 > X_3 > Max(X_1, X_2)) = \frac{1}{6}$

Solution: The probability density function (pdf) f(x) of a random variable X is defined as the derivative of the cdf F(x):

$$f(x) = \frac{d}{dx}F(x).$$

It is sometimes useful to consider the cdf F(x) in terms of the pdf f(x):

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

The PDF of X is,

$$F_X(x) = \int_{-\infty}^{\infty} f(x)dx \tag{5.8.1}$$

a) $\Pr(X_2 > X_1)$

$$= \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} f_X(t)\ dtdx \tag{5.8.2}$$

$$= \int_{-\infty}^{\infty} f_X(x)F_X(x)dx \tag{5.8.3}$$

$$= \frac{F_X^2(x)}{2}\Big|_{-\infty}^{\infty} \tag{5.8.4}$$

$$= \frac{1}{2}. \tag{5.8.5}$$

b) $\Pr(X_4 > Max(X_1, X_2) > X_3)$

$$= \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} f_X(t).^2C_1.$$

$$\left[\int_{-\infty}^{t} f_X(w)dw\right]\int_{-\infty}^{t} f_X(z)dzdtdx \tag{5.8.6}$$

$$= \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} 2f_X(t)F_X^2(t)dtdx \tag{5.8.7}$$

$$= \int_{-\infty}^{\infty} f_X(x).\frac{2}{3}F_X^3(x)dx \tag{5.8.8}$$

$$= \frac{2}{3}\frac{F_X^4(x)}{4}\Big|_{-\infty}^{\infty} \tag{5.8.9}$$

$$= \frac{1}{6}. \tag{5.8.10}$$

c) $\Pr(X_4 > X_3 > Max(X_1, X_2)$

$$= \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} f_X(t) \int_{-\infty}^{t} f_X(z).^2C_1.$$

$$\left[\int_{-\infty}^{t} f_X(w)dw\right]\ dzdtdx \tag{5.8.11}$$

$$= \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} f_X(t) \int_{-\infty}^{t} 2f_X(z)F_X(t)\ dzdtdx \tag{5.8.12}$$

$$= \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{x} f_X(t)F_X^2(t)dtdx \tag{5.8.13}$$

$$= \int_{-\infty}^{\infty} f_X(x).\frac{1}{3}F_X^3(x)dx \tag{5.8.14}$$

$$= \frac{1}{3}\frac{F_X^4(x)}{4}\Big|_{-\infty}^{\infty} \tag{5.8.15}$$

$$= \frac{1}{12}. \tag{5.8.16}$$

$\therefore$ $\boxed{\textbf{Option 1,3} \text{ are } \textbf{correct} \text{ answers.}}$

5.9. Consider a parallel system with two components. The lifetimes of the two components are independent and identically distributed random variables each following an exponential distribution with mean 1. The expected lifetime of the system is:

A) 1

B) $\frac{1}{2}$

C) $\frac{3}{2}$

D) 2

Solution:

Consider two random variables X and Y which represent the lifetime of the two components namely A and B.

$$X \sim Exp(\lambda_X) \tag{5.9.1}$$

$$Y \sim Exp(\lambda_Y) \tag{5.9.2}$$

Let $f_X(x)$ denote the probability distribution function for random variable X.

$$f_X(x) = \begin{cases} \lambda_X e^{-\lambda_X x} & x \geq 0 \\ 0 & otherwise \end{cases} \tag{5.9.3}$$

Let $f_Y(y)$ denote the probability distribution function for random variable Y.

$$f_Y(y) = \begin{cases} \lambda_Y e^{-\lambda_Y y} & y \geq 0 \\ 0 & otherwise \end{cases} \tag{5.9.4}$$

Let $F_X(x)$ denote the cumulative distribution function for random variable X.

$$F_X(x) = \begin{cases} 1 - e^{-\lambda_X x} & x \geq 0 \\ 0 & otherwise \end{cases} \tag{5.9.5}$$

Let $F_Y(y)$ denote the cumulative distribution function for random variable Y.

$$F_Y(y) = \begin{cases} 1 - e^{-\lambda_Y y} & y \geq 0 \\ 0 & otherwise \end{cases} \tag{5.9.6}$$

$$E(X) = \frac{1}{\lambda_X} \tag{5.9.7}$$

$$E(Y) = \frac{1}{\lambda_Y} \tag{5.9.8}$$

From 5.9.7 and 5.9.8,

$$\lambda_X = \lambda_Y = 1 \tag{5.9.9}$$

Let Z be a random variable such that $Z = max(X, Y)$

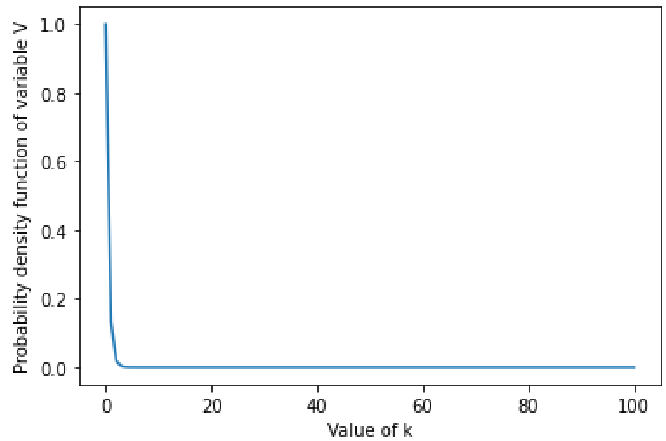


Fig. 5.9.1: Parallel system

$$P(Z \leq z) = P(max(X, Y) \leq z) \tag{5.9.10}$$
$$= P(X \leq z, Y \leq z) \tag{5.9.11}$$
$$= P(X \leq z)P(Y \leq z) \tag{5.9.12}$$
$$= (F_X(z) - F_X(0))(F_Y(z) - F_Y(0)) \tag{5.9.13}$$
$$= 1 - e^{-(\lambda_X)z} - e^{-(\lambda_Y)z} + e^{-(\lambda_X+\lambda_Y)z} \tag{5.9.14}$$

$P(Z \leq z)$ denotes the probability that the system dies in the first z seconds.

$$Expectation = \int_0^\infty z\, d(P(Z \leq z)) \tag{5.9.15}$$
$$= \int_0^\infty z(\lambda_X e^{-(\lambda_X)z} + \lambda_Y e^{-(\lambda_Y)z} - (\lambda_X + \lambda_Y)e^{-(\lambda_X+\lambda_Y)z})\, dz \tag{5.9.16}$$
$$= \frac{1}{\lambda_X} + \frac{1}{\lambda_Y} - \frac{1}{\lambda_X + \lambda_Y} \tag{5.9.17}$$

From 5.9.9,

$$Expectation = \frac{3}{2} \tag{5.9.18}$$

Therefore, option C correct.

5.10. Let X and Y be independent and identically distributed random variables such that $\Pr(X = 0) = \Pr(X = 1) = \frac{1}{2}$. Let $Z = X + Y$ and $W = |X - Y|$. Then which statement is not correct?

a) X and W are independent.

b) Y and W are independent.

c) Z and W are uncorrelated.

d) Z and W are independent.

6 Binomial Distribution

6.1. The probability that a part manufactured by a company will be defective is 0.05. If 15 such parts are selected randomly and inspected,the probability that atleast two parts will be defective is ...

Solution: The desired probabilty is

$$\Pr(X \geq 2) = 1 - \Pr(X < 2) \tag{6.1.1}$$
$$= 1 - \Pr(X = 0) - \Pr(X = 1) \tag{6.1.2}$$
$$= 1 - {}^{15}C_0 p^0 q^{15} - {}^{15}C_1 p^1 q^{14} \tag{6.1.3}$$
$$= 0.1709 \tag{6.1.4}$$

where

$$p = 0.0.5, q = 1 - p = 0.95 \tag{6.1.5}$$

and X is binomial with parameters $(15, p)$.

6.2. Let X be a binomial random variable with parameters $\left(11, \frac{1}{3}\right)$. At which value(s) of k is $\Pr(X = k)$ maximized?

a) $k = 2$
b) $k = 3$
c) $k = 4$
d) $k = 5$

Solution:
The binomial distribution is given by:

$$\Pr(X = k) = {}^{n}C_k \times p^k \times q^{n-k} \tag{6.2.1}$$

We are given $n = 11$, $p = \frac{1}{3}$ and hence $q = \frac{2}{3}$

$$\Pr(X = k) = {}^{11}C_k \left(\frac{2}{3}\right)^{11-k} \left(\frac{1}{3}\right)^{k} \tag{6.2.2}$$

$$\frac{\Pr(X = k)}{\Pr(X = k+1)} \geq 1 \tag{6.2.3}$$
$$\implies \frac{{}^{11}C_k \left(\frac{2}{3}\right)^{11-k} \left(\frac{1}{3}\right)^{k}}{{}^{11}C_{k+1} \left(\frac{2}{3}\right)^{10-k} \left(\frac{1}{3}\right)^{k+1}} \geq 1 \tag{6.2.4}$$
$$\text{or, } k \geq 3 \tag{6.2.5}$$

after some algebra. Similarly,

$$\frac{\Pr(X = k)}{\Pr(X = k-1)} \geq 1 \tag{6.2.6}$$
$$k \leq 4 \tag{6.2.7}$$

From (6.2.5) and (6.2.7), it is obvious that $\Pr(X = k)$ is maximized for $k = 3$, $k = 4$. Hence, options 2 and 3 are correct. See Fig. 6.2.1 for a graphical verification.

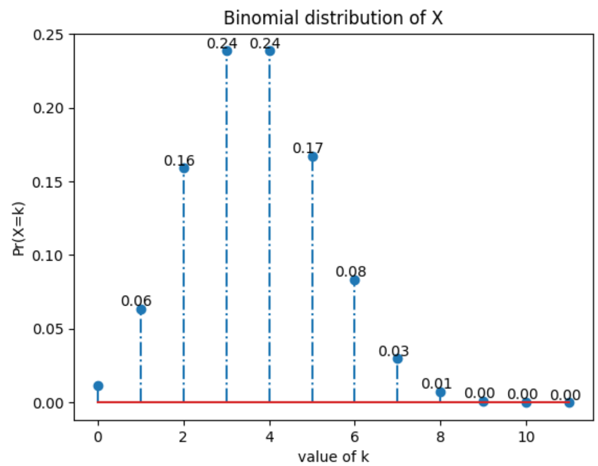


Fig. 6.2.1: Binomial Distribution of X

6.3. Suppose X_1, X_2, X_3, X_4 are i.i.d random variables taking values 1 and -1 with probability $1/2$ each. Then $E(X_1 + X_2 + X_3 + X_4)^4$ equals

a) 4 b) 76 c) 16 d) 12

Solution:

Theorem 6.1. *If $X_1, \ldots, X_n$ are i.i.d. random variables, all Bernoulli trials with success probability p, then their sum is distributed according to a binomial distribution with parameters n and p $\sum_{k=1}^{n} X_k \sim B(n, p)$*

Corollary 6.2. *For a binomial random variable X with parameters n and p and $q = 1 - p$*

$$E(X) = np \tag{6.3.1}$$
$$E\left(X^2\right) = np\,(np + q) \tag{6.3.2}$$
$$E\left(X^3\right) = np\left(n^2p^2 + 3npq - 2pq + q\right) \tag{6.3.3}$$
$$E\left(X^4\right) = np\left(n^3p^3 + 6n^2p^2q - 11np^2q + 7npq - 6pq^2 + q\right) \tag{6.3.4}$$

Given that X_i are i.i.d random variable for $i \in \{1, 2, 3, 4\}$ with,

$$\Pr(X_i = +1) = \frac{1}{2} \tag{6.3.5}$$

$$\Pr(X_i = -1) = \frac{1}{2} \tag{6.3.6}$$

Lemma 6.1. *Let random variables $M_i = \frac{X_i+1}{2}$ then, M_i are Bernoulli random variables with $p = 0.5$*

Proof.

$$\Pr(M_i = 1) = \Pr(X_i = 1) = \frac{1}{2} \tag{6.3.7}$$

$$\Pr(M_i = 0) = \Pr(X_i = -1) = \frac{1}{2} \tag{6.3.8}$$

$\therefore M_i$ are Bernoulli random variables with $p = 0.5$ □

Let assume random variable Y as

$$Y = \sum_{i=1}^{4} X_i \tag{6.3.9}$$

similarly let

$$Z = \sum_{i=1}^{4} M_i \tag{6.3.10}$$

Corollary 6.3. *Z is Binomial random variable ,*
from theorem 6.1 $Z \sim B(4, 0.5)$

$$Z = \sum_{i=1}^{4} \frac{X_i + 1}{2} \tag{6.3.11}$$

$$Z = \frac{Y}{2} + 2 \tag{6.3.12}$$

$$Y^4 = 16(Z-2)^4 \tag{6.3.13}$$

$$Y^4 = 16\left(Z^4 - 8z^3 + 24Z^2 - 32Z + 16\right) \tag{6.3.14}$$

$$\begin{aligned} E\left(Y^4\right) = &16\left(E\left(Z^4\right) - 8E\left(z^3\right)\right. \\ &\left.+ 24E\left(Z^2\right) - 32E(Z) + 16\right) \end{aligned} \tag{6.3.15}$$

From corollary 6.2 and 6.3

$$E\left(Y^4\right) = 40 \tag{6.3.16}$$

6.4. The probability that a ticketless traveler is caught during a trip is 0.1. If the traveler makes 4 trips , the probability that he/she will be caught during at least one of the trips is:

a) $1 - (0.9)^4$
b) $(1 - 0.9)^4$
c) $1 - (1 - 0.9)^4$
d) $(0.9)^4$

Solution: Let $X_i \in \{0, 1\}$ represent the ith trip where 1 denotes a ticketless traveller is caught. Given,

$$\Pr(X_i = 1) = p = 0.1 \tag{6.4.1}$$

Let,

$$X = \sum_{i=1}^{n} X_i \tag{6.4.2}$$

where n is the number of trips and X has a binomial distribution.

$$p_X(k) = \begin{cases} {}^nC_k\, p^K (1-p)^{n-k}, & 0 \le k \le n \\ 0, & otherwise \end{cases} \tag{6.4.3}$$

As he/she makes 4 trips in total, Using (6.4.1) and (6.4.3),

$$\Pr(X = 0) = p_X(0) \tag{6.4.4}$$

$$= {}^4C_0\, p^0 (1-p)^4 \tag{6.4.5}$$

$$\Pr(X = 0) = (0.9)^4 \tag{6.4.6}$$

Then probability of being caught in atleast one trip is,(Using (6.4.6))

$$\Pr(X \ge 1) = 1 - \Pr(X < 1) \tag{6.4.7}$$

$$= 1 - \Pr(X = 0) \tag{6.4.8}$$

$$= 1 - (0.9)^4 \tag{6.4.9}$$

6.5. Let X be a binomial random variable with parameters $\left(11, \frac{1}{3}\right)$. At which value(s) of k is $\Pr(X = k)$ maximized?

a) k=2
b) k=3
c) k=4
d) k=5

Solution: X has a binomial distribution :

$$\Pr(X = k) = {}^nC_k (q)^{n-k} (p)^k \tag{6.5.1}$$

Where,

- n=11
- $p = \frac{1}{3}$
- $q = 1 - p = 1 - \frac{1}{3} = \frac{2}{3}$

$$\Pr(X = k) = {}^{11}C_k \left(\frac{2}{3}\right)^{11-k} \left(\frac{1}{3}\right)^k \quad (6.5.2)$$

For Pr(X = k) to be maximized

$$\Pr(X = k) \geq \Pr(X = k+1) \quad (6.5.3)$$

$$\frac{\Pr(X = k)}{\Pr(X = k+1)} = \frac{{}^{11}C_k \left(\frac{2}{3}\right)^{11-k} \left(\frac{1}{3}\right)^k}{{}^{11}C_{k+1} \left(\frac{2}{3}\right)^{10-k} \left(\frac{1}{3}\right)^{k+1}} \geq 1 \quad (6.5.4)$$

$$\frac{2(k+1)}{11-k} \geq 1 \quad (6.5.5)$$

$$\implies k \geq 3 \quad (6.5.6)$$

$$\Pr(X = k) \geq \Pr(X = k-1) \quad (6.5.7)$$

$$\frac{\Pr(X = k)}{\Pr(X = k-1)} = \frac{{}^{11}C_k \left(\frac{2}{3}\right)^{11-k} \left(\frac{1}{3}\right)^k}{{}^{11}C_{k-1} \left(\frac{2}{3}\right)^{12-k} \left(\frac{1}{3}\right)^{k-1}} \geq 1 \quad (6.5.8)$$

$$\frac{12-k}{2k} \geq 1 \quad (6.5.9)$$

$$\implies k \leq 4 \quad (6.5.10)$$

From (6.5.6) , (6.5.10) and since k is an integer
$\Pr(X = k)$ is maximized for k=3, k=4
Thus options 2) and 3) are correct

7 Poisson Distribution

7.1. Let X be a Poisson random variable with p.m.f

$$P(X = k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!}, & k = 0, 1, 2, ...; \lambda > 0 \\ 0 & \text{otherwise} \end{cases} \quad (7.1.1)$$

If $Y = X^2 + 3$, then what is $P(Y = y)$ equal to?

(A) $\frac{e^{-\lambda}\lambda^{\sqrt{y-3}}}{\sqrt{(y-3)!}}$, for $y = \{3, 4, 7, 12, ...\}$

(B) $\frac{e^{-\lambda}\lambda^{-\sqrt{y-3}}}{\sqrt{(3-y)!}}$, for $y = \{3, 4, 7, 12, ...\}$

(C) $\frac{e^{-\lambda}\lambda^{\sqrt{3-y}}}{\sqrt{(3-y)!}}$, for $y = \{4, 7, 12, ...\}$

(D) $\frac{e^{-\lambda}\lambda^{-\sqrt{3-y}}}{\sqrt{(3-y)!}}$, for $y = \{4, 7, 12, ...\}$

Solution:

$$Y = X^2 + 3 \quad (7.1.2)$$

$$\implies X = \sqrt{Y - 3} \quad (7.1.3)$$

Substituting $k = \sqrt{y-3}$ in (7.1.1),

$$p_Y(y) = \begin{cases} \frac{e^{-\lambda}\lambda^{\sqrt{y-3}}}{\sqrt{(y-3)!}}, & y = 3, 4, 7, 12, ... \\ 0 & \text{otherwise} \end{cases} \quad (7.1.4)$$

Hence, the correct option is (A).

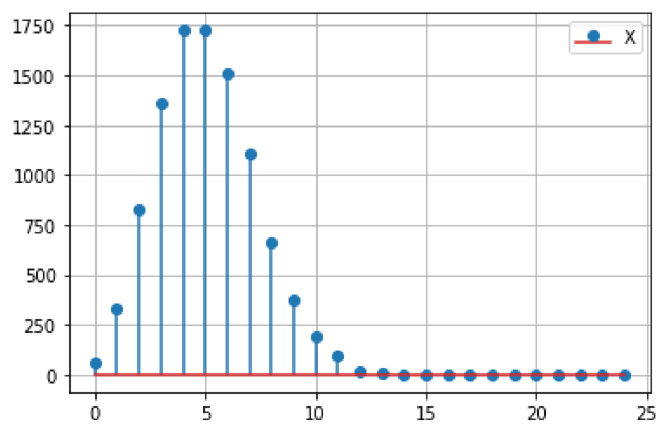


Fig. 7.1.1: Poisson stem plot for X ($\lambda = 5$)

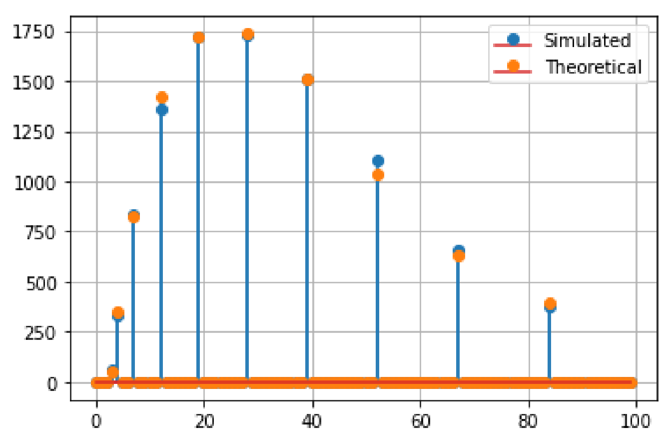


Fig. 7.1.2: Stem plot for Y (Simulated and Theoretical) ($\lambda = 5$)

7.2. For $n \geq 1$, let X_n be a Poisson random variable with mean n^2. Which of the following are equal to $\frac{1}{\sqrt{2\pi}} \int_2^\infty e^{-x^2/2}\, dx$

a) $\lim_{n\to\infty} \Pr\left(X_n > (n+1)^2\right)$
b) $\lim_{n\to\infty} \Pr\left(X_n \leqslant (n+1)^2\right)$
c) $\lim_{n\to\infty} \Pr\left(X_n < (n-1)^2\right)$
d) $\lim_{n\to\infty} \Pr\left(X_n < (n-2)^2\right)$

Solution:

Lemma 7.1. *Let Y_i be a Poisson random variable with mean 1 for $i \in \left(1, n^2\right)$ Then,*

$$\sum_{i}^{n^2} Y_i = X_n \tag{7.2.1}$$

Lemma 7.2.

$$\lim_{n\to\infty} \frac{X_n - n^2}{n} = \mathcal{N}(0,1) \tag{7.2.2}$$

Proof. Using the central limit theorem,

$$\lim_{n\to\infty} \frac{Y_1 + Y_2 + .. + Y_{n^2} - n^2}{n} = \mathcal{N}(0,1) \tag{7.2.3}$$

yielding (7.2.2). □

Lemma 7.3.

$$Q(2) = \frac{1}{\sqrt{2\pi}} \int_2^\infty e^{-x^2/2}\, dx \tag{7.2.4}$$

Lemma 7.4.

$$\Pr(X_n > k) \tag{7.2.5}$$
$$= Q\left(\frac{k - n^2}{n}\right) \tag{7.2.6}$$

Proof. For $X \sim \mathcal{N}(0,1)$, using Lemma 7.2,

$$\Pr(X_n > k) = \Pr\left(X > \frac{k - n^2}{n}\right) \tag{7.2.7}$$

yielding (7.2.6). □

Lemma 7.5.

$$Q(X) = 1 - Q(-x) \tag{7.2.8}$$

a) Using Lemma 7.4,

$$\lim_{n\to\infty} \Pr\left(X_n > (n+1)^2\right) = \lim_{n\to\infty} Q\left(\frac{(n+1)^2 - n^2}{n}\right) \tag{7.2.9}$$
$$= Q(2) \tag{7.2.10}$$

$\therefore$ **Option 1 is correct**

b) Using Lemma 7.4,

$$\lim_{n\to\infty} \Pr\left(X_n \leqslant (n+1)^2\right) = 1 - \lim_{n\to\infty} \Pr\left(X_n > (n+1)^2\right) \tag{7.2.11}$$
$$= 1 - \lim_{n\to\infty} Q\left(\frac{(n+1)^2 - n^2}{n}\right) \tag{7.2.12}$$
$$= 1 - Q(2) \tag{7.2.13}$$
$$\neq Q(2) \tag{7.2.14}$$

$\therefore$, from Lemma 7.3 **Option 2 is incorrect**

c) Using Lemma 7.4 and Lemma 7.5

$$\lim_{n\to\infty} \Pr\left(X_n < (n-1)^2\right) = 1 - \lim_{n\to\infty} \Pr\left(X_n \geq (n-1)^2\right) \tag{7.2.15}$$
$$= 1 - \lim_{n\to\infty} Q\left(\frac{(n-1)^2 - n^2}{n}\right) \tag{7.2.16}$$
$$= 1 - Q(-2) \tag{7.2.17}$$
$$= Q(2) \tag{7.2.18}$$

$\therefore$ **Option 3 is also correct**

d) Using Lemma 7.4 and Lemma 7.5,

$$\lim_{n\to\infty} \Pr\left(X_n < (n-2)^2\right) = 1 - \lim_{n\to\infty} \Pr\left(X_n \geq (n-2)^2\right) \tag{7.2.19}$$
$$= 1 - \lim_{n\to\infty} Q\left(\frac{(n-2)^2 - n^2}{n}\right) \tag{7.2.20}$$
$$= 1 - Q(-4) \tag{7.2.21}$$
$$= Q(4) < Q(2) \tag{7.2.22}$$

$\therefore$ **Option 4 is incorrect**

Solution:

Definition 3. *(Poisson Process)*
A counting process (number of arrivals from time 0 to t) is called a Poisson process with rate λ if

(i) The process has independent increments, and
(ii) The number of arrivals in any interval of length $\tau > 0$ has Poisson($\lambda\tau$) distribution.

∴ The distribution of the number of arrivals in any interval depends only on the length of the interval, and not on the exact location of the interval on the real line or on past or future arrivals.

Definition 4. *X and Y are Poisson distributions with parameters $\mu_1 = \lambda\tau_1 = 2 \times 1$ and $\mu_2 = \lambda\tau_2 = 2 \times 5$ respectively.*
The pmf (probability mass function) of a Poisson distribution with parameter μ is given by

$$p_N(n) = \frac{e^{-\mu} \cdot \mu^n}{n!} \tag{7.2.23}$$

Hence the pmfs of random variables X and Y are given by:

$$p_X(x) = \frac{e^{-2} \cdot 2^x}{x!}, \qquad \text{for } x = 0, 1, 2, \ldots \tag{7.2.24}$$

$$p_Y(y) = \frac{e^{-10} \cdot 10^y}{y!}, \qquad \text{for } y = 0, 1, 2, \ldots \tag{7.2.25}$$

Lemma 7.6. *If X and Y are two independent Poisson distributions with parameters μ_1 and μ_2 respectively, then the distribution of $X + Y$ is also Poisson with parameter $\mu_1 + \mu_2$.*

Proof. We have for $k \geq 0$, the probability mass function $p_{X+Y}(k)$ is a convolution of pmfs $p_X(x)$ and $p_Y(y)$:

$$p_{X+Y}(k) = \Pr(X + Y = k) = \Pr(Y = k - X) \tag{7.2.26}$$

$$= \sum_i \Pr(Y = k - i | X = i) \times p_X(i) \tag{7.2.27}$$

As X and Y are independent:

$$\Pr(Y = k - i \mid X = i) = \Pr(Y = k - i) = p_Y(k - i) \tag{7.2.28}$$

Simplifying (7.2.27)

$$p_{X+Y}(k) = \sum_{i=0}^{k} p_Y(k - i) \times p_X(i) \tag{7.2.29}$$

$$= \sum_{i=0}^{k} e^{-\mu_2} \frac{\mu_2^{k-i}}{(k - i)!} e^{-\mu_1} \frac{\mu_1^i}{i!} \tag{7.2.30}$$

$$= e^{-(\mu_1+\mu_2)} \frac{1}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k - i)!} \mu_1^i \mu_2^{k-i} \tag{7.2.31}$$

$$= e^{-(\mu_1+\mu_2)} \frac{1}{k!} \sum_{i=0}^{k} \binom{k}{i} \mu_1^i \mu_2^{k-i} \tag{7.2.32}$$

$$p_{X+Y}(k) = \frac{e^{-(\mu_1+\mu_2)} \cdot (\mu_1 + \mu_2)^k}{k!} \tag{7.2.33}$$

□

Lemma 7.7. *If X and Y are two independent Poisson distributions with parameters μ_1 and μ_2 respectively, then the distribution of $X - Y$ is no longer Poisson, as $X - Y$ will also attain negative values. More precisely, pmf of $X - Y$ is given by* (7.2.38)*, which is a Skellam distribution.*

Proof. We have for $k \in \mathbb{Z}$, the probability mass function $p_{X-Y}(k)$ is a convolution of pmfs $p_X(x)$ and $p_Y(y)$:

$$p_{X-Y}(k) = \Pr(X - Y = k) = \Pr(X = k + Y) \tag{7.2.34}$$

$$= \sum_i \Pr(X = k + i | Y = i) \times p_Y(i) \tag{7.2.35}$$

As X and Y are independent, and the Poisson distribution is zero for negative values of the count, the sum is only taken for those terms where $i \geq 0$ and $i + k \geq 0$.

$$p_{X-Y}(k) = \sum_{i=max(0,-k)}^{\infty} p_X(k + i) \times p_Y(i) \tag{7.2.36}$$

$$= \sum_{i=max(0,-k)}^{\infty} e^{-\mu_1} \frac{\mu_1^{k+i}}{(k + i)!} e^{-\mu_2} \frac{\mu_2^i}{i!} \tag{7.2.37}$$

$$= e^{-(\mu_1+\mu_2)} \sum_{i=max(0,-k)}^{\infty} \frac{\mu_1^{k+i} \mu_2^i}{i!(k + i)!} \tag{7.2.38}$$

Hence we can see from (7.2.38) that X - Y is not a Poisson distribution.

Note: This result is valid only for independent Poisson distributions. We may have dependent Poisson distributions whose difference is also Poisson. For example let Z be the random variable with distribution $X + Y$, where X and Y are independent Poisson distributions. From lemma (7.6), Z is Poisson. And, $Z - Y = X$ is also Poisson by definition, as Z is always at least as great as Y (*i.e.* not independent).

□

(A) From the definition of Poisson process (Definition 3), X and Y are independent. Hence option (A) is **correct.**

(B) As X and Y are independent Poisson distributions, using lemma (7.6), $X + Y$ is a Poisson with parameter $\mu_1 + \mu_2 = 12 \neq 6$. Hence option (B) is **incorrect.**

$$p_{X+Y}(k) = \frac{e^{-(12)} \cdot (12)^k}{k!} \tag{7.2.39}$$

(C) As X and Y are independent Poisson distributions, using lemma (7.7), $X - Y$ is not Poisson. Hence option (C) is **incorrect.**

(D)

$$\Pr(X = 0 \mid X + Y = 12) = \frac{\Pr(X = 0, Y = 12)}{\Pr(X + Y = 12)} \tag{7.2.40}$$

As X and Y are independent, and using (7.2.24), (7.2.25), and (7.2.39), we have:

$$\Pr(X=0 \mid X+Y=12) = \frac{\Pr(X=0) \cdot \Pr(Y=12)}{\Pr(X + Y = 12)} \tag{7.2.41}$$

$$= \frac{\dfrac{e^{-2} \times 2^0}{0!} \times \dfrac{e^{-10} \times 10^{12}}{12!}}{\dfrac{e^{-12} \times 12^{12}}{12!}} \tag{7.2.42}$$

$$= \left(\frac{5}{6}\right)^{12} \tag{7.2.43}$$

Hence option (D) is **correct.**

Ans: (A), (D)

7.3. Suppose customers arrive in a shop according to a Poisson process with rate 4 per hour. The shop opens at 10 : 00 am. If it is given that the second customer arrives at 10 : 40 am, what is the probability that no customer arrived before 10 : 30 am?

a) $\frac{1}{4}$
b) e^{-2}
c) $\frac{1}{2}$
d) $e^{\frac{1}{2}}$

Solution: We need to find

Random Variable	Time at which people arrive
X_p	$p = 10:00 - 10:30$
X_q	$q = 10:30 - 10:40$
X_r	$r = 10:00 - 10:40$
Y	$10:40$

TABLE 7.3.1: Random Variables

$$\Pr\left(X_p = 0 | Y = 2\right) \tag{7.3.1}$$

In the world where the 2^{nd} person arrives at 10 : 40 am the (7.3.1) becomes:

$$= \frac{\Pr\left(X_p = 0, X_q = 1\right)}{\Pr(X_r = 1)} \tag{7.3.2}$$

$$= \frac{\Pr\left(X_p = 0\right) \times \Pr\left(X_q = 1\right)}{\Pr(X_r = 1)} \tag{7.3.3}$$

The Poisson function distribution for time interval t and rate λ for a random variable X:

$$f_X(x;t) = \frac{(\lambda t)^x \exp(-\lambda t)}{x!}$$

For the time interval p:

$$\lambda = 4, t = 0.5, x = 0 \tag{7.3.4}$$

$$\Pr\left(X_p = 0\right) = f_X\left(0; \frac{1}{2}\right) \tag{7.3.5}$$

$$= e^{-2} \tag{7.3.6}$$

$$\tag{7.3.7}$$

For the time interval q:

$$\lambda = 4, t = \frac{1}{6}, x = 1 \tag{7.3.8}$$

$$\Pr\left(X_q = 1\right) = f_X\left(1; \frac{1}{6}\right) \tag{7.3.9}$$

$$= \frac{2}{3}e^{\frac{-2}{3}} \tag{7.3.10}$$

For the time interval r:

$$\lambda = 4, t = \frac{2}{3}, x = 1 \tag{7.3.11}$$

$$\Pr\left(X_r = 1\right) = f_X\left(1; \frac{2}{3}\right) \tag{7.3.12}$$

$$= \frac{8}{3}e^{\frac{-8}{3}} \tag{7.3.13}$$

Substituting (7.3.6) (7.3.10) (7.3.13) in (7.3.3):

$$\Pr\left(X_p = 0 | Y = 2\right) = \frac{1}{4} \tag{7.3.14}$$

7.4. Men arrive in a queue according to a Poisson process with rate λ_1 and women arrive in the same queue according to another Poisson process with rate λ_2. The arrivals of men and women are independent. The probability that the first person to arrive in the queue is a man is:

a) $\dfrac{\lambda_1}{\lambda_1 + \lambda_2}$

b) $\dfrac{\lambda_2}{\lambda_1 + \lambda_2}$

c) $\dfrac{\lambda_1}{\lambda_2}$

d) $\dfrac{\lambda_2}{\lambda_1}$

Solution: Let X and Y be Poisson random variables, with the values X takes being the number of men joining the queue in an arbitrary time t, and the values Y takes being the number of women joining the queue in an arbitrary time t.

$$Pr\left(X = i\right) = \frac{\lambda_1^i \cdot e^{-\lambda_1}}{i!} \tag{7.4.1}$$

$$Pr\left(Y = i\right) = \frac{\lambda_2^i \cdot e^{-\lambda_2}}{i!} \tag{7.4.2}$$

For 2 independent Poisson distributions with means λ_1 and λ_2, the simultaneous distribution can be represented by:

$$Pr\left(X + Y = i\right) = \frac{(\lambda_1 + \lambda_2)^i \cdot e^{-(\lambda_1 + \lambda_2)}}{i!} \tag{7.4.3}$$

Now we take conditional probability that if only one person entered the queue within a certain time t, then the probability of them being a man and not a woman is given by:

$$Pr\left(X = 1 | (X + Y) = 1\right) = \frac{Pr\left((X = 1) + (Y = 0)\right)}{Pr\left(X + Y = 1\right)} \tag{7.4.4}$$

$$\tag{7.4.5}$$

Since X and Y are independent,

$$Pr\left(X = 1 | (X + Y) = 1\right) = \frac{Pr\left(X = 1\right) \cdot Pr\left(Y = 0\right)}{Pr\left(X + Y = 1\right)} \tag{7.4.6}$$

$$= \frac{\frac{\lambda_1^1 \cdot e^{-\lambda_1}}{1!} \cdot \frac{\lambda_2^0 \cdot e^{-\lambda_2}}{0!}}{\frac{(\lambda_1+\lambda_2)^1 \cdot e^{-(\lambda_1+\lambda_2)}}{1!}} \tag{7.4.7}$$

$$= \frac{\lambda_1}{\lambda_1 + \lambda_2} \tag{7.4.8}$$

The probability that the first person to arrive in the queue is a man is option A, i.e $\frac{\lambda_1}{\lambda_1+\lambda_2}$

8 Gaussian Distribution

8.1. Let $X_1, X_2, \dots$ be independent random variables with X_n being uniformly distributed between -n and 3n, n=1,2,.... Let $S_N = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \frac{X_n}{n}$ for N=1,2,...and let F_N be the distribution function of S_N. Also, let ϕ denote the distribution function of a standard normal variable. Which of the following is/are true?

A) $\lim_{N\to\infty} F_N(0) \leq \phi(0)$

B) $\lim_{N\to\infty} F_N(0) \geq \phi(0)$

C) $\lim_{N\to\infty} F_N(1) \leq \phi(1)$

D) $\lim_{N\to\infty} F_N(1) \geq \phi(1)$

Solution: Given, $X_1, X_2, \dots$ are independent random variables with $X_i \sim \mathcal{U}(-i, 3i)$. Let us define:

$$Y_i = \frac{X_i}{i} \ \forall \text{ i} \tag{8.1.1}$$

$$\implies Y_i \sim \mathcal{U}(-1, 3) \tag{8.1.2}$$

Now, $Y_1, Y_2, \dots$ are i.i.d (independent and identically distributed) uniform random variables.

Definition 5. *The probability density function of a uniformly distributed continuous random variable in the interval [a,b] is*

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{for } a \le x \le b \\ 0, & \text{otherwise} \end{cases} \tag{8.1.3}$$

Lemma 8.1. *For a uniformly distributed continuous random variable in the interval [a,b], the mean and variance are given by:*

$$\mu = \frac{a+b}{2} \tag{8.1.4}$$

$$\sigma^2 = \frac{(b-a)^2}{12} \tag{8.1.5}$$

Proof.

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x)\,dx \tag{8.1.6}$$

$$\mu = \int_{-\infty}^{a} 0x\,dx + \int_{a}^{b} \frac{1}{b-a} x\,dx + \int_{b}^{+\infty} 0x\,dx \tag{8.1.7}$$

$$= 0 + \frac{1}{b-a}\left(\frac{x^2}{2}\right)\Bigg|_a^b + 0 \tag{8.1.8}$$

$$= \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2} \tag{8.1.9}$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x)\,dx \tag{8.1.10}$$

$$E(X^2) = \int_{-\infty}^{a} 0x^2\,dx + \int_{a}^{b} \frac{1}{b-a} x^2\,dx + \int_{b}^{+\infty} 0x^2\,dx \tag{8.1.11}$$

$$= 0 + \frac{1}{b-a}\left(\frac{x^3}{3}\right)\Bigg|_a^b + 0 \tag{8.1.12}$$

$$= \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3} \tag{8.1.13}$$

$$\sigma^2 = E(X^2) - [E(X)]^2 \tag{8.1.14}$$

$$= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 \tag{8.1.15}$$

$$= \frac{(b-a)^2}{12} \tag{8.1.16}$$

□

Using (8.1.4) and (8.1.5), we get:

$$\mu = \frac{-1+3}{2} = 1 \tag{8.1.17}$$

$$\sigma^2 = \frac{[3-(-1)]^2}{12} = \frac{16}{12} = \frac{4}{3} \tag{8.1.18}$$

And, we have:

$$S_N = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \frac{X_n}{n} \tag{8.1.19}$$

$$= \frac{1}{\sqrt{N}} \sum_{n=1}^{N} Y_n \tag{8.1.20}$$

Lemma 8.2 (Central Limit theorem)**.** *Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with finite μ and σ^2. If Z_n is defined as*

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \tag{8.1.21}$$

Then,

$$\lim_{n\to\infty} Z_n \sim \mathcal{N}(\mu, \sigma^2) \tag{8.1.22}$$

By (8.1.22), we can conclude

$$\lim_{N\to\infty} S_N \sim \mathcal{N}(\mu, \sigma^2) \tag{8.1.23}$$

$$\implies \lim_{N\to\infty} F_N(x) = \phi\left(\frac{x-\mu}{\sigma}\right) \tag{8.1.24}$$

$$= \phi\left(\frac{\sqrt{3}(x-1)}{2}\right) \tag{8.1.25}$$

Substituting $x = 1$, we get

$$\lim_{N\to\infty} F_N(1) = \phi\left(\frac{\sqrt{3}(1-1)}{2}\right) \tag{8.1.26}$$

$$= \phi(0) \tag{8.1.27}$$

The distribution function F_N is non-decreasing. So,

$$\lim_{N\to\infty} F_N(0) \le \lim_{N\to\infty} F_N(1) = \phi(0) \tag{8.1.28}$$

$$\lim_{N\to\infty} F_N(0) \le \phi(0) \tag{8.1.29}$$

The distribution function ϕ is non-decreasing. So,

$$\lim_{N\to\infty} F_N(1) = \phi(0) \le \phi(1) \tag{8.1.30}$$

$$\lim_{N\to\infty} F_N(1) \le \phi(1) \tag{8.1.31}$$

Answer: Option (A) and (C)

8.2. Let U and V be two independent zero mean Gaussian random variables of variances $\frac{1}{4}$ and $\frac{1}{9}$ respectively. The probability $\Pr(3V \geq 2U)$ is ...

Solution: From the given information,

$$U = \mathcal{N}\left(0, \frac{1}{4}\right) V \quad = \mathcal{N}\left(0, \frac{1}{9}\right) \tag{8.2.1}$$

Let $Y = 3V - 2U$. Then,

$$\mathrm{E}(Y) = 3\mathrm{E}(V) - 2\mathrm{E}(U) = 0 \tag{8.2.2}$$

$$\mathrm{var}(Y) = 3^2\mathrm{var}(V) + 2^2\mathrm{var}(U) = 2 \tag{8.2.3}$$

$$\therefore Y = \mathcal{N}(0, 2) \tag{8.2.4}$$

Thus,

$$\Pr(3V \geq 2U) = \Pr(3V - 2U \geq 0) \tag{8.2.5}$$

$$= \Pr(Y \geq 0) = \frac{1}{2} \tag{8.2.6}$$

$\because$ Y is symmetric about the origin.

8.3. (X, Y) follows bivariate normal distribution N_2(0,0,1,1,ρ), -1 $< \rho <$ 1. Then,

a) $X + Y$ and $X - Y$ are uncorrelated only if $\rho = 0$
b) $X + Y$ and $X - Y$ are uncorrelated only if $\rho < 0$
c) $X + Y$ and $X - Y$ are uncorrelated only if $\rho > 0$
d) $X+Y$ and $X-Y$ are uncorrelated for all values of ρ

Solution: Given that

$$\mathbf{M} = \begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{8.3.1}$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{8.3.2}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{8.3.3}$$

Also,

$$X + Y = \mathbf{A}^\top \mathbf{M} \tag{8.3.4}$$

$$X - Y = \mathbf{B}^\top \mathbf{M} \tag{8.3.5}$$

where

$$\mathbf{A} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{B} \quad = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{8.3.6}$$

Thus,

$$Cov(X + Y, X - Y) = \mathbf{A}^\top \boldsymbol{\Sigma} \mathbf{B} \tag{8.3.7}$$

$$= 0 \tag{8.3.8}$$

$\therefore$ X+Y and X-Y are uncorrelated irrespective of value of ρ where $\rho \in (-1, 1)$.

8.4. Let U and V be two independent zero mean Gaussian random variables of variances $\frac{1}{4}$ and $\frac{1}{9}$ respectively. The probability $\Pr(3V \geq 2U)$ is ...

Solution: From the given information,

$$U = \mathcal{N}\left(0, \frac{1}{4}\right) V \quad = \mathcal{N}\left(0, \frac{1}{9}\right) \tag{8.4.1}$$

Let $Y = 3V - 2U$. Then,

$$\mathrm{E}(Y) = 3\mathrm{E}(V) - 2\mathrm{E}(U) = 0 \tag{8.4.2}$$

$$\mathrm{var}(Y) = 3^2\mathrm{var}(V) + 2^2\mathrm{var}(U) = 2 \tag{8.4.3}$$

$$\therefore Y = \mathcal{N}(0, 2) \tag{8.4.4}$$

Thus,

$$\Pr(3V \geq 2U) = \Pr(3V - 2U \geq 0) \tag{8.4.5}$$

$$= \Pr(Y \geq 0) = \frac{1}{2} \tag{8.4.6}$$

$\because$ Y is symmetric about the origin.

8.5. Let X_1, X_2, X_3, X_4, X_5 be a random sample of size 5 from a population having standard normal distribution. If

$$\overline{X} = \frac{1}{5} \sum_{i=1}^{5} X_i \tag{8.5.1}$$

$$T = \sum_{i=1}^{5} \left(X_i - \overline{X}\right)^2 \tag{8.5.2}$$

then $\mathrm{E}\left(T^2 \overline{X}^2\right)$ is equal to

a) 3
b) 3.6
c) 4.8
d) 5.2

Solution:

Lemma 8.3. *Let*

$$\mathbf{x}=\begin{pmatrix}X_1\\X_2\\X_3\\X_4\\X_5\end{pmatrix},\mathbf{y}=\begin{pmatrix}Y_1\\Y_2\\Y_3\\Y_4\\Y_5\end{pmatrix},\sim\mathcal{N}(\mathbf{0},\mathbf{I}) \tag{8.5.3}$$

$$\mathbf{u}=\begin{pmatrix}1\\1\\1\\1\\1\end{pmatrix},\mathbf{v}=\begin{pmatrix}1\\1\\1\\1\\0\end{pmatrix}. \tag{8.5.4}$$

Then

$$\overline{X}=\frac{1}{5}\mathbf{u}^\top\mathbf{x} \tag{8.5.5}$$

$$T=\left\|\mathbf{v}^\top\mathbf{y}\right\|^2 \tag{8.5.6}$$

where

$$\mathbf{y}=\mathbf{P}\mathbf{x}, \tag{8.5.7}$$

$$\mathbf{M}=\mathbf{I}-\frac{1}{5}\mathbf{1} \tag{8.5.8}$$

$$=\mathbf{P}^\top\mathbf{D}\mathbf{P}, \tag{8.5.9}$$

$$\mathbf{D}=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&0\end{pmatrix} \tag{8.5.10}$$

and **1** *is the all ones matrix.*

Proof. From (8.5.1) and (8.5.2), it is easy to verify that

$$T=\mathbf{x}^\top\mathbf{M}\mathbf{x},\mathbf{M}^2=\mathbf{M} \tag{8.5.11}$$

$$\implies T=\mathbf{x}^\top\mathbf{P}\mathbf{D}\mathbf{P}^\top\mathbf{x} \tag{8.5.12}$$

$$=\mathbf{y}^\top\mathbf{D}\mathbf{y} \tag{8.5.13}$$

yielding (8.5.6). We have used spectral decomposition above. □

Lemma 8.4. $\mathbf{y}=\mathbf{P}\mathbf{x}\sim\mathcal{N}(\mathbf{0},\mathbf{I})$ *if* $\mathbf{P}^\top\mathbf{P}=\mathbf{I}$.

Proof. The moment generating function

$$M_{\mathbf{x}}(\mathbf{t})=exp\left(\mathbf{t}^\top\mu+\frac{1}{2}\mathbf{t}^\top\mathbf{V}\mathbf{t}\right) \tag{8.5.14}$$

$\because \mu=\mathbf{0}$ and $\mathbf{V}=\mathbf{I}$,

$$M_{\mathbf{x}}(\mathbf{t})=\exp\left(\frac{1}{2}\mathbf{t}^\top\mathbf{t}\right) \tag{8.5.15}$$

Therefore the joint moment generating function of **y** is

$$M_{\mathbf{y}}(\mathbf{t})=M_{\mathbf{x}}(\mathbf{P}\mathbf{t}) \tag{8.5.16}$$

$$=exp\left(\frac{1}{2}\mathbf{t}^\top\mathbf{P}^\top\mathbf{P}\mathbf{t}\right) \tag{8.5.17}$$

$$=M_{\mathbf{x}}(\mathbf{t}) \tag{8.5.18}$$

□

Corollary 8.1. $\overline{X}\sim\mathcal{N}\left(0,\frac{1}{5}\right)$

Definition 6. *a)* Random Variable: *A random variable X is a real-valued function defined on the "sample space" Ω (the set of outcomes being studied via probability).*

b) Borel Set: *A random variable X is studied by means of the probabilities that its value lies within various intervals of real numbers (or, more generally, sets constructed in simple ways out of intervals: these are the Borel measurable sets of real numbers).*

c) sigma-algebra: Corresponding to any Borel measurable set I is the event $X^*(I)$ consisting of all outcomes ω for which $X(\omega)$ lies in I. The sigma-algebra generated by X is determined by the collection of all such events.

d) Independent random variables: *The naive definition says two random variables X and Y are independent "when their probabilities multiply." That is, when I is one Borel measurable set and J is another, then*

$$\Pr(X(\omega)\in I,Y(\omega)\in J)$$
$$=\Pr(X(\omega)\in I)\Pr(Y(\omega)\in J). \tag{8.5.19}$$

But in the language of events (and sigma algebras) that's the same as

$$\Pr(\omega\in X^*(I),\omega\in Y^*(J))$$
$$=\Pr(\omega\in X^*(I))\Pr(\omega\in Y^*J)). \tag{8.5.20}$$

e) Function of a random variable:

$$(f\circ X)(\omega)=f(X(\omega)) \tag{8.5.21}$$

$$(f\circ X)^*(I)=X^*(f^*(I)) \tag{8.5.22}$$

Lemma 8.5. *Functions of independent random variables are themselves independent.*

Proof. Consider now two functions $f,g:\mathbb{R}\to\mathbb{R}$ and suppose that $f\circ X$ and $g\circ Y$ are random variables. In other words, every event generated by $f\circ X$ (which is on the left) in (8.5.22) is automatically an event generated by X (as exhibited by the form of the right hand side). Therefore (8.5.22) automatically holds for $f \circ X$ and $g \circ Y$. This covers the case of vector-valued random variables as well. $\square$

Corollary 8.2. *Let* $\mathbf{y}$ *and* $\mathbf{z}$ *be two independent normal random vectors. Then* $\mathbf{y}$ *and* $\|\mathbf{z}\|$ *are also independent.*

Lemma 8.6. $\mathbf{Ax}, \mathbf{Bx}$ *are independent if and only if* $\mathbf{AB}^\top = 0$

Proof. The given vectors are independent $\iff$

$$\mathrm{E}\left((\mathbf{Ax} - \mathrm{E}(\mathbf{Ax}))(\mathbf{Bx} - \mathrm{E}(\mathbf{Bx}))^\top\right) = 0$$
$$\implies \mathbf{A}\mathrm{E}\left((\mathbf{x} - \mathrm{E}(\mathbf{x}))(\mathbf{x} - \mathrm{E}(\mathbf{x}))^\top\right)\mathbf{B}^\top = 0$$
$$\implies \mathbf{A}\mathrm{var}(\mathbf{x})\,\mathbf{B}^\top = 0$$
$$\text{or, } \mathbf{AB}^\top = 0, \quad \because \mathrm{var}(\mathbf{x}) = \mathbf{I} \tag{8.5.23}$$

$\square$

Theorem 8.3. $\mathbf{u}^\top\mathbf{x}$ *and* $\mathbf{v}^\top\mathbf{y}$ *are independent.*

Proof. From Lemma 8.3,

$$\mathbf{y} = \mathbf{Px}. \tag{8.5.24}$$

The given statement can be proved using Lemma 8.6 and noting that

$$\mathbf{u}^\top\mathbf{P}^\top\mathbf{v} = 0, \tag{8.5.25}$$

$\square$

Corollary 8.4. $\overline{X} andT$ *are independent.*

Definition 7. chi-square distribution: $\|\mathbf{x}\|^2$ *is said to be chi-square distributed. If the length of* $\mathbf{x}$ *is k, The mean and variance are given by*

$$E(Y) = k \tag{8.5.26}$$
$$var(Y) = 2k \tag{8.5.27}$$

From Corollary 8.1 and Definition 7,

$$\mathrm{E}\left(\overline{X}^2\right) = \frac{1}{5}, \tag{8.5.28}$$
$$\mathrm{E}\left(T^2\right) = \mathrm{var}(T) + (\mathrm{E}(T))^2 \tag{8.5.29}$$
$$= 24 \tag{8.5.30}$$
$$\implies \mathrm{E}\left(T^2\overline{X}^2\right) = 4.8 \tag{8.5.31}$$

8.6. Suppose (X_1, X_2) follows a bivariate normal distribution with

$$E(X_1) = E(X_2) = 0 \tag{8.6.1}$$
$$V(X_1) = V(X_2) = 2 \tag{8.6.2}$$
$$Cov(X_1, X_2) = -1 \tag{8.6.3}$$

If $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{y^2}{2}}\,dy$, then $\Pr(X_1 - X_2 > 6)$ = ?

a) $\Phi(-1)$
b) $\Phi(-3)$
c) $\Phi(\sqrt{6})$
d) $\Phi(-\sqrt{6})$

Solution:

Definition 8. $\mathbf{x} \sim \mathcal{N}(\mu_\mathbf{x}, \Sigma_\mathbf{x})$ *is a bivariate random vector given by*

$$\mathbf{x} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \tag{8.6.4}$$
$$\mu_\mathbf{x} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \tag{8.6.5}$$
$$\mathbf{\Sigma_x} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \tag{8.6.6}$$

where $\mathbf{\Sigma_x}$ *is covariance matrix of* $\mathbf{x}$ *and* ρ *is correlation of* X_1 *and* X_2 *which is given by*

$$\rho = \frac{cov(X_1,, X_2)}{\sigma_1\sigma_2} \tag{8.6.7}$$

Lemma 8.7. *Let* $\mathbf{x}$ *be a* $K \times 1$ *multivariate normal random vector with mean* μ_x *and covariance matrix* $\mathbf{\Sigma_x}$. *Let* $\mathbf{a}$ *be an* $L \times 1$ *real vector and* $\mathbf{B}$ *an* $L \times K$ *full rank real matrix with* $K > L$ *.Then the* $L \times 1$ *random vector* $\mathbf{y}$ *defined by*

$$\mathbf{y} = \mathbf{a} + \mathbf{Bx} \tag{8.6.8}$$

has a multivariate normal distribution with mean

$$\mu_\mathbf{y} = \mathbf{a} + \mathbf{B}\mu_\mathbf{x} \tag{8.6.9}$$

and covariance matrix

$$\mathbf{\Sigma_y} = \mathbf{B\Sigma_x B}^\top \tag{8.6.10}$$

Proof. The joint moment generating function of $\mathbf{x}$ is

$$M_\mathbf{x}(\mathbf{t}) = \mathbb{E}[\exp(\mathbf{t}^\top\mathbf{x})] \tag{8.6.11}$$
$$\implies M_\mathbf{x}(\mathbf{t}) = \exp\left(\mathbf{t}^\top\mu_\mathbf{x} + \frac{1}{2}\mathbf{t}^\top\mathbf{\Sigma_x t}\right) \tag{8.6.12}$$

$\therefore$, the joint moment generating function of $\mathbf{y}$ is

$$M_{\mathbf{y}}(\mathbf{t}) = \mathbb{E}[\exp(\mathbf{t}^\top \mathbf{y})] \quad (8.6.13)$$
$$= \mathbb{E}[\exp(\mathbf{t}^\top(\mathbf{a}+\mathbf{B}\mathbf{x}))] \quad (8.6.14)$$
$$= \exp(\mathbf{t}^\top \mathbf{a})\mathbb{E}[(\mathbf{t}^\top \mathbf{B}\mathbf{x})] \quad (8.6.15)$$
$$= \exp(\mathbf{t}^\top \mathbf{a})\mathbb{E}[(\mathbf{B}^\top \mathbf{t})^\top \mathbf{x}] \quad (8.6.16)$$
$$= \exp(\mathbf{t}^\top \mathbf{a})M_{\mathbf{x}}(\mathbf{B}^\top \mathbf{t}) \quad (8.6.17)$$

which can be expressed as

$$M_{\mathbf{y}}(\mathbf{t})$$
$$= \exp(\mathbf{t}^\top \mathbf{a})\exp\left((\mathbf{B}^\top \mathbf{t})^\top \mu_{\mathbf{x}} + \frac{1}{2}(\mathbf{B}^\top \mathbf{t})^\top \mathbf{\Sigma_x}\mathbf{B}^\top \mathbf{t}\right)$$
$$= \exp\left(\mathbf{t}^\top(\mathbf{a}+\mathbf{B}\mu_{\mathbf{x}}) + \frac{1}{2}\mathbf{t}\mathbf{B}\mathbf{\Sigma_x}\mathbf{B}^\top \mathbf{t}^\top\right) \quad (8.6.18)$$

which is the moment generating function of a multivariate normal distribution with mean $\mathbf{a}+\mathbf{B}\mu$ and covariance matrix $\mathbf{B}\mathbf{\Sigma_x}\mathbf{B}^\top$.

$\therefore$ $\mu_{\mathbf{y}} = \mathbf{a}+\mathbf{B}\mu_{\mathbf{x}}$ and $\mathbf{\Sigma_y} = \mathbf{B}\mathbf{\Sigma_x}\mathbf{B}^\top$ $\square$

Theorem 8.5. *If* (X_1, X_2) *follows a bivariate distribution then*

$$\Pr(X_1 - X_2 > \alpha) = \Phi\left(\frac{-(\alpha + \mathbf{u}^\top \mu_{\mathbf{x}})}{\sqrt{\mathbf{u}^\top \mathbf{\Sigma_x}\mathbf{u}}}\right) \quad (8.6.19)$$

Proof. Let $\mathbf{u} = \begin{pmatrix}-1\\1\end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix}X_1\\X_2\end{pmatrix}$. Then

$$Y = X_2 - X_1 = \mathbf{u}^\top \mathbf{x} \quad (8.6.20)$$

with

$$\mu_y = \mathbf{u}^\top \mu_{\mathbf{x}} \quad (8.6.21)$$
$$\sigma_y^2 = \mathbf{u}^\top \mathbf{\Sigma_x}(\mathbf{u}^\top)^\top \quad (8.6.22)$$
$$\implies \sigma_y^2 = \mathbf{u}^\top \mathbf{\Sigma_x}\mathbf{u} \quad (8.6.23)$$
$$\text{or, } Y \sim \mathcal{N}(\mathbf{u}^\top \mu_{\mathbf{x}}, \mathbf{u}^\top \mathbf{\Sigma_x}\mathbf{u}) \quad (8.6.24)$$

from Lemma 8.7. Also,

$$Z = \frac{Y-\mu_y}{\sigma_y} \sim \mathcal{N}(0,1) \quad (8.6.25)$$

Thus,

$$\Pr(Y < -\alpha) = \Pr\left(\frac{Y-\mu_y}{\sigma_y} < \frac{-\alpha-\mu_y}{\sigma_y}\right) \quad (8.6.26)$$
$$= \Pr\left(Z < \frac{-(\alpha + \mathbf{u}^\top \mu_{\mathbf{x}})}{\sqrt{\mathbf{u}^\top \mathbf{\Sigma_x}\mathbf{u}}}\right) \quad (8.6.27)$$

$$\implies \Pr(X_1 - X_2 > \alpha) = \Phi\left(\frac{-(\alpha + \mathbf{u}^\top \mu_{\mathbf{x}})}{\sqrt{\mathbf{u}^\top \mathbf{\Sigma_x}\mathbf{u}}}\right) \quad (8.6.28)$$

$\square$

From the given information,

$$\mu_1 = \mu_2 = 0 \quad (8.6.29)$$
$$\sigma_1^2 = \sigma_2^2 = 2 \quad (8.6.30)$$
$$Cov(X_1, X_2) = -1 \quad (8.6.31)$$
$$\rho = \frac{Cov(X_1, X_2)}{\sigma_1\sigma_2} = \frac{-1}{2} \quad (8.6.32)$$
$$\mu_{\mathbf{x}} = \begin{pmatrix}\mu_1\\\mu_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \quad (8.6.33)$$
$$\mathbf{\Sigma_x} = \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix} = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix} \quad (8.6.34)$$

which, upon substituting in (8.6.28) yields

$$\implies \Pr(X_1 - X_2 > 6) = \Phi\left(\frac{-6}{\sqrt{6}}\right) \quad (8.6.35)$$
$$= \Phi(-\sqrt{6}) \quad (8.6.36)$$

8.7. Let C be a circle of unit area with centre at origin and let S be a square of unit area with $\left(\frac{1}{2}, \frac{1}{2}\right)$, $\left(\frac{-1}{2}, \frac{1}{2}\right)$, $\left(\frac{-1}{2}, \frac{-1}{2}\right)$ and $\left(\frac{1}{2}, \frac{-1}{2}\right)$ as the four vertices. If X and Y be two independent standard variates, show that

$$\iint_C \phi(x)\,\phi(y)\,dx\,dy \geq \iint_S \phi(x)\,\phi(y)\,dx\,dy$$

where $\phi(.)$ is the pdf of $N(0,1)$ distribution.
Solution:

Definition 9. *PDF of normal distribution is given as*

$$\phi_Z(z) = N\left(\mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(z-\mu)^2}{2\sigma^2}} \quad (8.7.1)$$

Corollary 8.6.

$$\phi_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \tag{8.7.2}$$

$$\phi_Y(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} \tag{8.7.3}$$

Proof. Since $\phi(.)$ is the pdf of $N(0,1)$ distribution (given in question), $\mu = 0$ and $\sigma^2 = 1$.

$$\implies \phi_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \text{ and } \phi_Y(y) = \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}}$$

□

Lemma 8.8. *For circle C,*

$$\iint_C \phi(x)\, \phi(y)\, dx\, dy = 1 - e^{\frac{-1}{2\pi}} \tag{8.7.4}$$

Proof. C has unit area with centre at origin.

$$\implies \pi \times r^2 = 1 \tag{8.7.5}$$

$$\implies |r| = \frac{1}{\sqrt{\pi}} \tag{8.7.6}$$

For the area inside circle C we get,

$$x^2 + y^2 \le \frac{1}{\pi} \implies |y| \le \sqrt{\frac{1}{\pi} - x^2} \tag{8.7.7}$$

From (8.7.7) we get,

$$\iint_C \phi(x)\, \phi(y)\, dx\, dy =$$

$$\int_{x=\frac{-1}{\sqrt{\pi}}}^{x=\frac{1}{\sqrt{\pi}}} \int_{y=-\sqrt{\frac{1}{\pi}-x^2}}^{y=\sqrt{\frac{1}{\pi}-x^2}} \phi(x)\, \phi(y)\, dy\, dx \tag{8.7.8}$$

Using (8.7.2) and (8.7.3) in (8.7.8) we get,

$$\int_{x=\frac{-1}{\sqrt{\pi}}}^{x=\frac{1}{\sqrt{\pi}}} \int_{y=-\sqrt{\frac{1}{\pi}-x^2}}^{y=\sqrt{\frac{1}{\pi}-x^2}} \frac{1}{2\pi} e^{\frac{-\left(x^2+y^2\right)}{2}}\, dy\, dx \tag{8.7.9}$$

Converting it to polar coordinates we get,

$$\int_{r=0}^{r=\frac{1}{\sqrt{\pi}}} \int_{\theta=0}^{\theta=2\pi} \frac{1}{2\pi} e^{\frac{-r^2}{2}}\, r\, d\theta\, dr \tag{8.7.10}$$

$$= \int_0^{\frac{1}{\sqrt{\pi}}} e^{\frac{-r^2}{2}}\, r\, dr \tag{8.7.11}$$

$$= 1 - e^{\frac{-1}{2\pi}} \tag{8.7.12}$$

□

Definition 10. *The Q function is defined as*

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-u^2}{2}}\, du \tag{8.7.13}$$

Lemma 8.9.

$$\int_{-a}^{a} e^{\frac{-x^2}{2}}\, dx = \sqrt{2\pi}\,(1 - 2Q(a)) \tag{8.7.14}$$

Proof. Since $e^{\frac{-x^2}{2}}$ is an even function,

$$\int_{-a}^{a} e^{\frac{-x^2}{2}}\, dx = 2\int_0^a e^{\frac{-x^2}{2}}\, dx \tag{8.7.15}$$

$$\implies 2\int_0^a e^{\frac{-x^2}{2}}\, dx = 2\left(\int_0^\infty e^{\frac{-x^2}{2}}\, dx - \int_a^\infty e^{\frac{-x^2}{2}}\, dx\right)$$

$$= 2\sqrt{2\pi} \times \frac{1}{\sqrt{2\pi}} \left(\int_0^\infty e^{\frac{-x^2}{2}}\, dx - \int_a^\infty e^{\frac{-x^2}{2}}\, dx\right) \tag{8.7.16}$$

Comparing (8.7.16) with (8.7.13) we get,

$$\int_{-a}^{a} e^{\frac{-x^2}{2}}\, dx = 2\sqrt{2\pi}\,(Q(0) - Q(a)) \tag{8.7.17}$$

$$= 2\sqrt{2\pi}\left(\frac{1}{2} - Q(a)\right) \tag{8.7.18}$$

$$= \sqrt{2\pi}\,(1 - 2Q(a)) \tag{8.7.19}$$

□

Lemma 8.10. *For square S,*

$$\iint_S \phi(x)\, \phi(y)\, dx\, dy = \left(1 - 2Q\left(\frac{1}{2}\right)\right)^2 \tag{8.7.20}$$

Proof. For square S (given in question),

$$\frac{-1}{2} \le x \le \frac{1}{2} \tag{8.7.21}$$

$$\frac{-1}{2} \le y \le \frac{1}{2} \tag{8.7.22}$$

From this we get,

$$\iint_S \phi(x)\, \phi(y)\, dx\, dy =$$

$$\int_{x=\frac{-1}{2}}^{x=\frac{1}{2}} \int_{y=\frac{-1}{2}}^{y=\frac{1}{2}} \phi(x)\, \phi(y)\, dy\, dx \tag{8.7.23}$$

Using (8.7.2) and (8.7.3) in (8.7.23) we get,

$$\int_{x=\frac{-1}{2}}^{x=\frac{1}{2}} \int_{y=\frac{-1}{2}}^{y=\frac{1}{2}} \frac{1}{2\pi} e^{\frac{-(x^2+y^2)}{2}}\, dy\, dx \tag{8.7.24}$$

Using (8.7.14) twice we get,

$$\int_{x=\frac{-1}{2}}^{x=\frac{1}{2}} \int_{y=\frac{-1}{2}}^{y=\frac{1}{2}} \frac{1}{2\pi} e^{\frac{-(x^2+y^2)}{2}}\, dy\, dx = \left(1 - 2Q\left(\frac{1}{2}\right)\right)^2 \tag{8.7.25}$$

□

Calculating the values of (8.7.4) and (8.7.20) we get,

$$1 - e^{\frac{-1}{2\pi}} = 0.147136 \tag{8.7.26}$$

$$\left(1 - 2Q\left(\frac{1}{2}\right)\right)^2 = 0.146631 \tag{8.7.27}$$

This proves that

$$1 - e^{\frac{-1}{2\pi}} \geq \left(1 - 2Q\left(\frac{1}{2}\right)\right)^2 \tag{8.7.28}$$

$$\implies \iint_C \phi(x)\, \phi(y)\, dx\, dy \geq \iint_S \phi(x)\, \phi(y)\, dx\, dy \tag{8.7.29}$$

8.8. Suppose $\begin{pmatrix} X \\ Y \end{pmatrix}$ is a random vector such that the marginal distribution of X and the marginal distribution of Y are the same and each is normally distributed with mean 0 and variance 1. Then, which of the following conditions imply independence of X and Y?

a) $\text{Cov}(X, Y) = 0$
b) $aX + bY$ is normally distributed with mean 0 and variance $a^2 + b^2$ for all real a and b
c) $\Pr(X \leq 0, Y \leq 0) = \frac{1}{4}$
d) $E\left[e^{itX+isY}\right] = E\left[e^{itX}\right] E\left[e^{isY}\right]$ for all real s and t

Solution:
An important property of dirac delta function that will be used at multiple ocassions in this solution is

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a) \tag{8.8.1}$$

Given $X \sim N(0,1), Y \sim N(0,1)$

a)

$$Cov(X, Y) = 0 \tag{8.8.2}$$

$$E[XY] - E[X]E[Y] = 0 \tag{8.8.3}$$

$$E[XY] = 0 \tag{8.8.4}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy = 0 \tag{8.8.5}$$

This doesn't imply independence. Counter example given below
Lets consider a case where X and Y are dependent based on the following relation, Y being independent of K

$$X = KY \tag{8.8.6}$$

PMF for K is given as

$$p_K(k) = \begin{cases} \frac{1}{2} & k = 1 \\ \frac{1}{2} & k = -1 \\ 0 & \text{otherwise} \end{cases} \tag{8.8.7}$$

A simulation is given below, Y is gaussian, then X also follows gaussian Theoretically

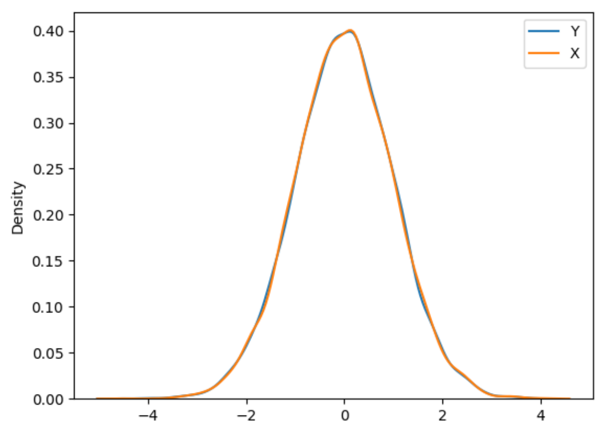


Fig. 8.8.1: X and Y, if Y is normal

it can be proved in the following manner, Since K and Y are independent

$$f_X(x) = \Pr(K = 1) f_Y(x) + \Pr(K = -1) f_Y(-x) \tag{8.8.8}$$

$$= \frac{1}{2}\left(f_Y(x) + f_Y(-x)\right) \tag{8.8.9}$$

$$= f_Y(x) \tag{8.8.10}$$

Therefore, X follows identical but not independent distribution as Y, An alternative proof is given below as a proof for marginal probability

Now consider that X is normally distributed, we will establish Y is also normally distributed. The joint probability distribution is therefore

$$f_{XY}(x,y) = f_{X|Y}(x|y)f_X(x) = f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y)) \tag{8.8.11}$$

The marginal probability distribution function for X is given as

$$\int_{-\infty}^{\infty} f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dy \tag{8.8.12}$$

Using (8.8.1), we get

$$\int_{-\infty}^{\infty} f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dy = f_X(x) \tag{8.8.13}$$

We know that $X \sim N(0,1)$, $f_X(x)$ represents gaussian probability distribution function.

Futher, using symmetry of (8.8.6), we can establish that marginal distribution of Y is gaussian. Here is a proof anyways

$$f_Y(y) = \int_{-\infty}^{\infty} f_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dx \tag{8.8.14}$$

Using (8.8.1), we get

$$f_Y(y) = \frac{1}{2}\left(f_X(y)+f_X(-y)\right) = f_X(y) \tag{8.8.15}$$

Since Y has identical probability distribution function, $Y \sim N(0,1)$

The covariance is given as

$$Cov(X,Y) = E[XY] - E[X]E[Y] = E[XY] \tag{8.8.16}$$

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_{XY}(x,y)dydx \tag{8.8.17}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)\frac{1}{2}(\delta(x+y)+\delta(x-y))dydx \tag{8.8.18}$$

$$= \int_{-\infty}^{\infty} xf_X(x)\int_{-\infty}^{\infty} y\frac{1}{2}(\delta(x+y)+\delta(x-y))dydx \tag{8.8.19}$$

Using (8.8.1)

$$E[XY] = \int_{-\infty}^{\infty} xf_X(x)\frac{1}{2}(x-x)dx = 0 \tag{8.8.20}$$

b) Defining the following matrices/vectors Given

vector/matrix	expression
$\mathbf{Z}$	$\begin{pmatrix} X & Y \end{pmatrix}^\top$
$\mathbf{C}$	$\begin{pmatrix} a & b \end{pmatrix}^\top$
$\boldsymbol{\mu}$	$\begin{pmatrix} 0 & 0 \end{pmatrix}^\top$
$\boldsymbol{\Sigma}$	$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$

TABLE 8.8.1: vectors/matrices and their expressions

$$\mathbf{C}^\top\mathbf{Z} \sim N\left(0, a^2+b^2\right) \tag{8.8.21}$$

Since this is true for all a and b, it is equivalent to X and Y being jointly gaussian

$$\mathbf{Z} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{8.8.22}$$

For correlated random variables X and Y in bivariate normal distribution, we have

$$\sigma_Z^2 = \sum_{i,j}\Sigma_{ij} \tag{8.8.23}$$

$$a^2+b^2 = a^2+b^2+2\rho ab \tag{8.8.24}$$

$$\therefore \rho = 0 \tag{8.8.25}$$

The joint distribution is given as

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^2\,|\boldsymbol{\Sigma}|}} \tag{8.8.26}$$

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\begin{pmatrix}x & y\end{pmatrix} I_2 \begin{pmatrix}x & y\end{pmatrix}^\top\right)}{\sqrt{(2\pi)^2}} \tag{8.8.27}$$

Where I_2 is the identity matrix of order 2

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\begin{pmatrix}x & y\end{pmatrix}\begin{pmatrix}x & y\end{pmatrix}^\top\right)}{\sqrt{(2\pi)^2}} \tag{8.8.28}$$

$$f_{\mathbf{Z}}(x,y) = \frac{\exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right)}{\sqrt{(2\pi)^2}} = f_X(x) f_Y(y) \tag{8.8.29}$$

$\therefore$ **Option(2) is correct**, A simulation for bivariate gaussian is given below

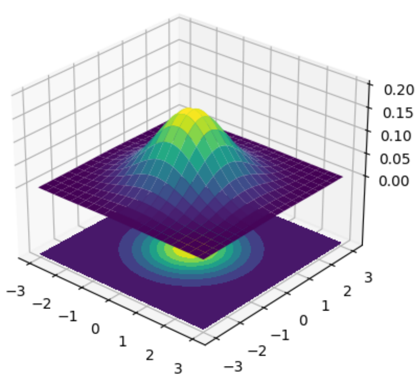


Fig. 8.8.2: bivariate gaussian while 0 mean vector and identity covariance matrix

c)

$$\Pr(X \le 0, Y \le 0) = \frac{1}{4} \tag{8.8.30}$$

This doesn't imply independence, it can be true even for dependent X and Y, the counter example is (8.8.11), the joint probability function is symmetric across all 4 quadrants

$$\therefore \Pr(X \le 0, Y \le 0) = \frac{1}{4} \tag{8.8.31}$$

Alternatively, here is proof

$$\Pr(X \le 0) = F_X(0) = \frac{1}{2} \tag{8.8.32}$$

Using (8.8.6)

$$\Pr(Y \le 0 | X \le 0) = \frac{1}{2} \tag{8.8.33}$$

Using (8.8.32) and (8.8.33)

$$\Pr(X \le 0, Y \le 0) = \frac{1}{4} \tag{8.8.34}$$

d)

$$E\left[e^{itX+isY}\right] = E\left[e^{itX}\right] E\left[e^{isY}\right] \tag{8.8.35}$$

$$E\left[e^{itX+isY}\right] = \varphi_X(t)\varphi_Y(s) \tag{8.8.36}$$

The inverse is given as

$$f_{XY}(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-itX-isY} E\left[e^{itX+isY}\right] dsdt \tag{8.8.37}$$

Using (8.8.36)

$$f_{XY}(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-itX-isY} \varphi_X(t)\varphi_Y(s) dsdt \tag{8.8.38}$$

$$f_{XY}(x,y) = f_X(x) f_Y(y) \tag{8.8.39}$$

$\therefore$ **Option(4) is correct**

9 Geometric Distribution

9.1. Suppose X follows an exponential distribution with parameter $\lambda > 0$. Fix $a > 0$. Define the random variable Y by
$Y = k, \quad$ if $ka \le X < (k+1)a$,
$k = 0, 1, 2 \ldots$
Which of the following statements are correct?

a) $\Pr(4 < Y < 5) = 0$
b) Y follows an Exponential distribution
c) Y follows a Geometric distribution
d) Y follows a Poisson distribution

Solution:

9.2. Suppose X has density

$$f(x|\theta) = \frac{1}{\theta} e^{-x/\theta}, x > 0 \tag{9.2.1}$$

Define

$$Y = k, \quad k \le X < k+1, \quad k = 0,1,2\dots \tag{9.2.2}$$

Then the distribution of Y is

a) Normal
b) Binomial
c) Poisson
d) Geometric

Solution:

$$\Pr(Y = k) = \Pr(k \le X < k+1) \tag{9.2.3}$$
$$= \int_k^{k+1} f(x|\theta)\,dx \tag{9.2.4}$$
$$= \int_k^{k+1} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \tag{9.2.5}$$
$$= \left[-e^{-\frac{x}{\theta}}\right]_k^{k+1} \tag{9.2.6}$$
$$= e^{-\frac{k}{\theta}}\left(1 - e^{-\frac{1}{\theta}}\right) \tag{9.2.7}$$
$$\implies \Pr(Y = k) = (1-p)^k\, p\, k = 0,1,2\dots \tag{9.2.8}$$

where

$$p = 1 - e^{-\frac{1}{\theta}} \tag{9.2.9}$$

Therefore, the distribution of Y is 4) Geometric.

10 Two Dimensions

10.1. Let $c \in \mathbb{R}$ be a constant. Let X, Y be random variables with joint probability density function

$$f(x,y) = \begin{cases} cxy & 0 < x < y < 1, \\ 0 & \text{otherwise} \end{cases} \tag{10.1.1}$$

Which of the following statements are correct ?
a) $c = \frac{1}{8}$
b) $c = 8$
c) X and Y are independent
d) $\Pr(X = Y) = 0$

Solution:

a) False

b) By definition,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \tag{10.1.2}$$
$$= \int_0^y cxy\,dx \tag{10.1.3}$$
$$= cy\left(\frac{x^2}{2}\right)\Bigg|_0^y \tag{10.1.4}$$
$$= \frac{cy^3}{2} \tag{10.1.5}$$

$$\implies f_Y(y) = \begin{cases} \frac{cy^3}{2}, & 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{10.1.6}$$

$\because$ the area under the pdf is 1, from (10.1.6),

$$\implies \int_{-\infty}^{\infty} f_Y(y)\,dy = 1 \tag{10.1.7}$$
$$\implies \int_0^1 c\frac{y^3}{2} = 1 \tag{10.1.8}$$
$$\implies \frac{c}{8} = 1 \tag{10.1.9}$$
$$\text{or, } c = 8 \tag{10.1.10}$$

Also,

$$f_Y(y) = \begin{cases} 4y^3 & \text{, if } 0 < y < 1 \\ 0 & \text{, otherwise} \end{cases} \tag{10.1.11}$$

c)

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \tag{10.1.12}$$
$$= \int_x^1 cxy\,dy \tag{10.1.13}$$
$$= cx\left(\frac{y^2}{2}\right)\Bigg|_x^1 \tag{10.1.14}$$
$$= cx\left(\frac{1-x^2}{2}\right) \tag{10.1.15}$$

$$\implies f_X(x) = \begin{cases} 4x\left(1-x^2\right), & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{10.1.16}$$

From (10.1.16) and (10.1.11)

$$f_X(x) \times f_Y(y) = \begin{cases} 16xy^3\left(1-x^2\right) & \text{, if } 0 < x, y < 1 \\ 0 & \text{, otherwise} \end{cases} \tag{10.1.17}$$

$$\neq f(x,y) \tag{10.1.18}$$

Hence, X and Y are not independent.

d)

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx \tag{10.1.19}$$

$$= \int_{0}^{x} 4x(1-x^2)\,dx \tag{10.1.20}$$

$$= \int_{0}^{x} 4x - 4x^3\,dx \tag{10.1.21}$$

$$= 2x^2 - 4x^4 \text{ for } 0 < x < 1 \tag{10.1.22}$$

yielding

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ 2x^2 - 4x^4 & 0 < x < 1 \\ 1 & x \ge 1 \end{cases} \tag{10.1.23}$$

From (10.1.23),

$$\Pr(Y - \epsilon < X < Y + \epsilon)$$

$$= F_X(Y+\epsilon) - F_X(Y-\epsilon)$$

$$= 8\epsilon Y\left(1 - Y^2 - \epsilon^2\right) \tag{10.1.24}$$

upon simplification. Letting

$$g(Y) = 8\epsilon Y(1 - Y^2 - \epsilon^2), \tag{10.1.25}$$

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y) f_Y(y)\,dy \tag{10.1.26}$$

$$= \int_{0}^{1} (4y^3)(8\epsilon y)(1 - y^2 - \epsilon^2)\,dy \tag{10.1.27}$$

$$\implies \Pr(Y - \epsilon < X < Y + \epsilon)$$

$$= 32\epsilon\left(\frac{2 - 7\epsilon^2}{35}\right) \tag{10.1.28}$$

Now substituting $\epsilon = 0$ in the above,

$$\Pr(X = Y) = 0 \tag{10.1.29}$$

10.2. Let X and Y be random variables having the joining probability density function

$$f_{XY}(x,y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{\frac{-1}{2y}(x-y)^2} & x \in (-\infty, \infty), \\ & y \in (0,1) \\ 0 & \text{otherwise} \end{cases} \tag{10.2.1}$$

The covariance between the random variables X and Y is

Solution:

10.3. Let a random variable X follow exponential distribution with mean 2. Define $Y = [X-2|X > 2]$. The value of $\Pr(Y \ge t)$ is . . .

Solution: From the given information,

$$\Pr(Y \ge t) = \frac{\Pr(X - 2 \ge t, X > 2)}{\Pr(X > 2)} \tag{10.3.1}$$

$$= \frac{\Pr(X \ge t + 2, X > 2)}{\Pr(X > 2)} \tag{10.3.2}$$

$\because$ X has an exponential distribution with parameter $\lambda = \frac{1}{2}$,

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } 0 < x < \infty \\ 0, & \text{otherwise} \end{cases} \tag{10.3.3}$$

and

$$\Pr(X > 2) = 1 - F_X(2) = e^{-2\lambda} \tag{10.3.4}$$

Also,

$$\Pr(X \ge t + 2, X > 2) = \begin{cases} \Pr(X \ge t + 2) & t \ge 0 \\ \Pr(X > 2) & t < 0 \end{cases} \tag{10.3.5}$$

Substituting (10.3.5) in (10.3.2), using (10.3.4) and simplifying,

$$\Pr(Y \ge t) = \begin{cases} e^{-\frac{t}{2}} & t \ge 0 \\ 1 & t < 0 \end{cases} \tag{10.3.6}$$

10.4. Let X and Y be two random variables with joint probability density function

$$f(x.y) = \begin{cases} \frac{1}{\pi} & 0 \le x^2 + y^2 \le 1 \\ 0 & otherwise \end{cases}$$

Which of the following statements are correct?

a) X and Y are independent.

b) $\Pr(X > 0) = \frac{1}{2}$

c) E(Y)=0

d) Cov(X,Y)=0

Solution:

a) The marginal PDF of X is given by

$$f_X(x) = \int_{y=-\infty}^{y=\infty} f_{XY}(x,y)dy \tag{10.4.1}$$

$$= \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} \frac{1}{\pi}dy \tag{10.4.2}$$

$$= \frac{2\sqrt{1-x^2}}{\pi} \tag{10.4.3}$$

The marginal PDF of Y is given by

$$f_Y(x) = \int_{x=-\infty}^{x=\infty} f_{XY}(x,y)dx \tag{10.4.4}$$

$$= \int_{x=-\sqrt{1-y^2}}^{x=\sqrt{1-y^2}} \frac{1}{\pi}dx \tag{10.4.5}$$

$$= \frac{2\sqrt{1-y^2}}{\pi} \tag{10.4.6}$$

Now,

$$f_X(x) \times f_Y(x) = \frac{2\sqrt{1-x^2}}{\pi} \times \frac{2\sqrt{1-y^2}}{\pi} \tag{10.4.7}$$

$$= \frac{4(1-x^2)(1-y^2)}{\pi^2} \tag{10.4.8}$$

$$\neq \frac{1}{\pi} \tag{10.4.9}$$

$$\neq f_{XY}(x,y) \tag{10.4.10}$$

Therefore, X and Y are not independent.

b) Now,

$$\Pr(X > 0) = \int_{x=0}^{x=\infty} f_X(x)dx \tag{10.4.11}$$

$$= \int_{x=0}^{x=1} \frac{2\sqrt{1-x^2}}{\pi}dx \tag{10.4.12}$$

$$= \left(\frac{\arcsin(x) + x\sqrt{1-x^2}}{\pi}\right)_0^1 \tag{10.4.13}$$

$$= \frac{1}{2} \tag{10.4.14}$$

Therefore, option(2) is correct.

c) Now,

$$E[Y] = \int_{y=-\infty}^{y=\infty} y f_Y(y)dy \tag{10.4.15}$$

$$= \int_{y=-1}^{y=1} \frac{2y\sqrt{1-y^2}}{\pi}dy \tag{10.4.16}$$

$$= \left(\frac{-2(1-y^2)^{\frac{3}{2}}}{3\pi}\right)_{-1}^1 \tag{10.4.17}$$

$$= 0 \tag{10.4.18}$$

Therefore, option(3) is also correct.

d) Now,

$$E[XY] = \int_x \int_y xy f_{XY}(x,y)dydx \tag{10.4.19}$$

$$= \int_{x=-1}^{x=1} \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} \frac{xy}{\pi}dydx \tag{10.4.20}$$

$$= \frac{x}{\pi} \int_{x=-1}^{x=1} \left(\frac{y^2}{2}\right)_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dx \tag{10.4.21}$$

$$= 0 \tag{10.4.22}$$

Now,

$$Cov(X,Y) = E[XY] - E[X]E[Y] \quad (10.4.23)$$
$$= 0 - E[X] \times 0 \quad (10.4.24)$$
$$= 0 \quad (10.4.25)$$

Therefore, option(4) is also correct.

10.5. Let X and Y be two random variables satisfying $X \geq 0, Y \geq 0, E(X) = 3, Var(X) = 9, E(Y) = 2$ and $Var(Y) = 4$. Which of the following statements are correct?

A) $0 \leq Cov(X,Y) \leq 4$
B) $E(XY) \leq 3$
C) $Var(X+Y) \leq 25$
D) $E(X+Y)^2 \geq 25$

Solution:

$$E(X^2) = Var(X) + (E(X))^2 = 18 \quad (10.5.1)$$

Similarly,

$$E(Y^2) = Var(Y) + (E(Y))^2 = 8 \quad (10.5.2)$$

We can use the Covariance inequality for this question,

$$(Cov(X,Y))^2 \leq Var(X)Var(Y) \quad (10.5.3)$$

The proof of this inequality is as shown,

$$Var\left(\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}\right) = Var\left(\frac{X}{\sigma_X}\right) + Var\left(\frac{\pm Y}{\sigma_Y}\right)$$
$$+ 2Cov\left(\frac{X}{\sigma_X}, \frac{\pm Y}{\sigma_Y}\right) \quad (10.5.4)$$
$$= \frac{1}{{\sigma_X}^2}Var(X) + \frac{1}{{\sigma_Y}^2}Var(Y)$$
$$+ 2Cov\left(\frac{X}{\sigma_X}, \frac{\pm Y}{\sigma_Y}\right) \quad (10.5.5)$$
$$= 2 \pm 2\frac{Cov(X,Y)}{\sigma_X\sigma_Y} \quad (10.5.6)$$

Since Variance is always positive,

$$Var\left(\frac{X}{\sigma_X} \pm \frac{Y}{\sigma_Y}\right) \geq 0 \quad (10.5.7)$$
$$2 \pm 2\frac{Cov(X,Y)}{\sigma_X\sigma_Y} \geq 0 \quad (10.5.8)$$
$$1 \pm 1\frac{Cov(X,Y)}{\sigma_X\sigma_Y} \geq 0 \quad (10.5.9)$$
$$|(Cov(X,Y))| \leq (\sigma_X)(\sigma_Y) \quad (10.5.10)$$
$$(Cov(X,Y))^2 \leq Var(X)Var(Y) \quad (10.5.11)$$

a) Substituting values of variance we get,

$$-6 \leq Cov(X,Y) \leq 6 \quad (10.5.12)$$

Therefore, option A is incorrect.

b) From equation (10.5.12),

$$Cov(X,Y) = E(XY) - E(X)E(Y) \quad (10.5.13)$$
$$-6 \leq E(XY) - E(X)E(Y) \leq 6 \quad (10.5.14)$$
$$0 \leq E(XY) \leq 12 \quad (10.5.15)$$

Also, if X and Y are independent,

$$E(XY) = E(X)E(Y) = 6 \quad (10.5.16)$$

Therefore, Option B is incorrect.

c) Now,

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y) \quad (10.5.17)$$
$$= 13 + 2Cov(X,Y) \quad (10.5.18)$$

From equation (10.5.12),

$$1 \leq Var(X+Y) \leq 25 \quad (10.5.19)$$

Therefore, Option C is correct.

d) Now,

$$E(X+Y)^2 = E(X^2) + E(Y^2) + 2E(XY) \quad (10.5.20)$$
$$E(X+Y)^2 = 26 + 2E(XY) \quad (10.5.21)$$

From equation (10.5.15),

$$26 \leq E(X+Y)^2 \leq 50 \quad (10.5.22)$$

Therefore, Option D is correct.

10.6. The joint probability density function of (X,Y) is

$$f(x,y) = \begin{cases} 6(1-x) & if \quad 0 < y < x, 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{10.6.1}$$

Which among the following are correct?

a) X and Y are not independent

b) $f_Y(y) = \begin{cases} 3\,(y-1)^2 & if \quad 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$

c) X and Y are independent

d) $f_Y(y) = \begin{cases} 3\left(y - \frac{1}{2}y^2\right) & if \quad 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$

Solution: Given joint probability density function of X and Y, marginal probability density functions are as follows:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy \tag{10.6.2}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx \tag{10.6.3}$$

Calculating $f_X(x)$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy \tag{10.6.4}$$

$$= \int_0^x 6(1-x)dy \tag{10.6.5}$$

$$f_X(x) = \begin{cases} 6x(1-x) & 0 < x < 1 \\ 0 & otherwise \end{cases} \tag{10.6.6}$$

Calculating $f_Y(y)$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx \tag{10.6.7}$$

$$= \int_y^1 6(1-x)dx \tag{10.6.8}$$

$$= 6x - 3x^2\Big|_y^1 \tag{10.6.9}$$

$$= 3 - 6y + 3y^2 \tag{10.6.10}$$

$$= 3(y-1)^2 \tag{10.6.11}$$

$$f_Y(y) = \begin{cases} 3(y-1)^2 & 0 < y < 1 \\ 0 & otherwise \end{cases} \tag{10.6.12}$$

To check whether X and Y are independent, we calculate $f_X(x) \times f_Y(y)$. From (10.6.6) and (10.6.12)

$$f_X(x) \times f_Y(y) = \begin{cases} 18x(1-x)(y-1)^2 \\ \qquad 0 < x < 1, 0 < y < 1 \\ 0 \qquad \text{otherwise} \end{cases} \tag{10.6.13}$$

$$\neq f(x,y) \tag{10.6.14}$$

Since f(x,y) and $f_X(x) \times f_Y(y)$ are different, random variables X and Y are not independent.

Options 1 and 2 are correct

10.7. Suppose that (X, Y) has a joint probability distribution with the marginal distribution of X being N(0,1) and $E(Y|X = x) = x^3$ for all $x \in R$. Then, which of the following statements are true?

a) $\text{Corr}(X, Y) = 0$

b) $\text{Corr}(X, Y) > 0$

c) $\text{Corr}(X, Y) < 0$

d) X and Y are independent

Solution: The following result shall be useful later. For $n \in N$

$$\int_{-\infty}^{\infty} \frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = \begin{cases} 0 & n \text{ is odd} \\ (n-1) \times ... \times 3 \times 1 & n \text{ is even} \end{cases} \tag{10.7.1}$$

The proof for the above can be found at the end of the solution.

$$Corr(X,Y) = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y} \tag{10.7.2}$$

We know $X \sim N(0,1)$. Thus,

$$f_X(x) = \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} \tag{10.7.3}$$

$$E(X) = 0 \tag{10.7.4}$$

$$\sigma_X^2 = 1 \tag{10.7.5}$$

$$\sigma_Y^2 = E(Y^2) - E(Y)^2 \tag{10.7.6}$$

$$E(Y) = \int_{-\infty}^{\infty} E(Y|X = x) f_X(x)dx \tag{10.7.7}$$

$$= \int_{-\infty}^{\infty} \frac{x^3 e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \tag{10.7.8}$$

$$= 0 \tag{10.7.9}$$

$$E(Y^2) = \int_{-\infty}^{\infty} E(Y^2|X = x) f_X(x) dx \quad (10.7.10)$$

$$= \int_{-\infty}^{\infty} \frac{x^6 e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \quad (10.7.11)$$

$$= 15 \quad (10.7.12)$$

Substituting in (10.7.6)

$$\sigma_Y^2 = 15 \quad (10.7.13)$$

$$\sigma_{XY}^2 = E(XY) - E(X)E(Y) \quad (10.7.14)$$

$$E(XY) = \int_{-\infty}^{\infty} E(XY|X = x) f_X(x) dx \quad (10.7.15)$$

$$= \int_{-\infty}^{\infty} E(xY|X = x) f_X(x) dx \quad (10.7.16)$$

$$= \int_{-\infty}^{\infty} xE(Y|X = x) f_X(x) dx \quad (10.7.17)$$

$$= \int_{-\infty}^{\infty} \frac{x^4 e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \quad (10.7.18)$$

$$= 3 \quad (10.7.19)$$

Substituting in (10.7.14)

$$\sigma_{XY}^2 = 3 \quad (10.7.20)$$

Substituting in (10.7.2)

$$Corr(X, Y) = \frac{3}{\sqrt{15}} > 0 \quad (10.7.21)$$

Since $Corr(X, Y) \neq 0$, X and Y are dependent. Thus option 2 is the only correct option. **Proof for the integral:** If n is odd, $\frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$ is an odd function, thus

$$\int_{-\infty}^{\infty} \frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = 0 \quad (10.7.22)$$

If n is even,

$$\int_{-\infty}^{\infty} \frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = \int_{-\infty}^{\infty} (x^{n-1}) (\frac{x e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}) dx \quad (10.7.23)$$

Using integration by parts,

$$\int_{-\infty}^{\infty} \frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = \left(x^{n-1} \int \frac{x e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \right) \Bigg|_{-\infty}^{\infty} - (n-1) \int_{-\infty}^{\infty} x^{n-2} \left(\int \frac{x e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \right) dx \quad (10.7.24)$$

$$= \left(x^{n-1} (-\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}) \right) \Bigg|_{-\infty}^{\infty} - (n-1) \int_{-\infty}^{\infty} x^{n-2} (-\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}) dx \quad (10.7.25)$$

$$= (n-1) \int_{-\infty}^{\infty} \frac{x^{n-2} e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \quad (10.7.26)$$

$$= (n-1)(n-3) \int_{-\infty}^{\infty} \frac{x^{n-4} e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \quad (10.7.27)$$

$$= (n-1) \times ... \times 3 \times 1 \int_{-\infty}^{\infty} \frac{x^0 e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx \quad (10.7.28)$$

$$= (n-1) \times ... \times 3 \times 1 \quad (10.7.29)$$

Alternative proof for the integral:

If n is odd, $\frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$ is an odd function, thus

$$\int_{-\infty}^{\infty} \frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = 0 \quad (10.7.30)$$

If n is even, let $n = 2k$. We differentiate the following identity k times w.r.t. α.

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\left(\frac{\pi}{\alpha}\right)} \quad (10.7.31)$$

On differentiating k times, we get

$$\int_{-\infty}^{\infty} x^{2k} e^{-\alpha x^2} = \frac{1 \times 3 \times ... \times (2k-1)}{2^k} \sqrt{\left(\frac{\pi}{\alpha^{2k+1}}\right)} \quad (10.7.32)$$

On substituting $\alpha = \frac{1}{2}$, we get

$$\int_{-\infty}^{\infty} x^n e^{-\frac{x^2}{2}} = 1 \times 3 \times ... \times (n-1) \sqrt{2\pi} \quad (10.7.33)$$

Thus

$$\int_{-\infty}^{\infty} \frac{x^n e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} dx = (n-1) \times \ldots \times 3 \times 1 \tag{10.7.34}$$

10.8. Consider the quadratic equation $x^2+2Ux+V=0$ where U and V are chosen independently and randomly from $\{1,2,3\}$ with equal probabilities. Then probability that the equation has both roots real

a) $\frac{2}{3}$
b) $\frac{1}{2}$
c) $\frac{7}{9}$
d) $\frac{1}{3}$

Solution: Let $U \in \{1.2,3\}$ and $V \in \{1,2,3\}$

TABLE 10.8.1: Probability of selecting values for U

k	1	2	3
$\Pr(U=k)$	1/3	1/3	1/3

TABLE 10.8.2: Probability of selecting values for V

k	1	2	3
$\Pr(V=k)$	1/3	1/3	1/3

For $x^2+2Ux+V=0$ to have real roots,

$$b^2 - 4ac \geq 0 \tag{10.8.1}$$
$$(2U)^2 - 4(1)(V) \geq 0 \tag{10.8.2}$$
$$U^2 \geq V \tag{10.8.3}$$

$$\Pr\left(U^2 \geq V\right) = 1 - \Pr\left(U^2 < V\right) \tag{10.8.4}$$

The possible pairs (U,V) for $\Pr\left(U^2 < V\right)$,

TABLE 10.8.3: Table for $\Pr\left(U^2 < V\right)$

(U,V) for $U^2 < V$	Probability
$(1,2)$	$\Pr(U=1)\Pr(V=2)$=1/9
$(1,3)$	$\Pr(U=1)\Pr(V=3) = \frac{1}{9}$
Total	$\Pr\left(U^2 < V\right) = \frac{2}{9}$

$$\Pr\left(U^2 \geq V\right) = 1 - \frac{2}{9} \tag{10.8.5}$$
$$\Pr\left(U^2 \geq V\right) = \frac{7}{9} \tag{10.8.6}$$

Hence, Option 3 is correct.

11 Integral Transforms

11.1. X and Y are independent random variables each having the density

$$f(t) = \frac{1}{\pi}\frac{1}{1+t^2} \quad -\infty < t < +\infty \tag{11.1.1}$$

Then the density function of $\frac{X+Y}{3}$ for $-\infty <$ t $< +\infty$ is

a) $\frac{6}{\pi}\frac{1}{4+9t^2}$

b) $\frac{6}{\pi}\frac{1}{9+4t^2}$

c) $\frac{3}{\pi}\frac{1}{1+9t^2}$

d) $\frac{3}{\pi}\frac{1}{9+t^2}$

Solution: Let us consider the random variables X and Y. The Characteristic function of the probability density $f(t)$ is

$$g(w) = \int_{-\infty}^{\infty} f(t)e^{iwt}dt \tag{11.1.2}$$
$$= \int_{-\infty}^{\infty} \frac{1}{\pi}\frac{1}{1+t^2}e^{iwt}dt \tag{11.1.3}$$
$$= e^{-|w|} \quad, -\infty < w < \infty \tag{11.1.4}$$

The product of the Characteristic function of probability density of X and Y is

$$h(w) = g_1(w) \times g_2(w) = e^{-2|w|} \tag{11.1.5}$$

To get the probability density of X+Y, we find the inverse characteristic function of h(w). But since there is a one to one correspondence between a function and its fourier transform and $h(w) = g(2w)$

$$F_{X+Y}(t) = \frac{1}{2}f\left(\frac{t}{2}\right) \tag{11.1.6}$$
$$= \frac{1}{2\pi}\frac{4}{4+t^2} \quad, -\infty < t < \infty \tag{11.1.7}$$

We know that if a random variable M has a probability density $f_M(x)$, then the probability density of random variable kM is

$$f_{kM}(x) = \frac{1}{|k|} f_M\left(\frac{x}{|k|}\right) \tag{11.1.8}$$

Probability density of $Z = \frac{X+Y}{3}$ given $F_{X+Y}(t)$ is

$$F_Z(t) = 3 \times f_{X+Y}(3t) \tag{11.1.9}$$
$$= \frac{6}{\pi}\frac{1}{4+9t^2} \tag{11.1.10}$$

11.2. Suppose X and Y are independent and identically distributed random variables and let $Z = X + Y$. Then the distribution of Z is in the same family as that of X and Y if X is
Solution:

1) Normal 2) Exponential
3) Uniform 4) Binomial

1) Let X and Y be independent and identically distributed normal random variables. Then the characteristic function of X and Y is given by

$$\Phi_X(\omega) = e^{j\eta\omega - \sigma^2\omega^2/2} \tag{11.2.1}$$

The characteristic function of Z is given by

$$\Phi_Z(\omega) = \Phi_X^2(\omega) \tag{11.2.2}$$
$$= e^{2j\eta\omega - \sigma^2\omega^2} \tag{11.2.3}$$

Thus Z is a normal random variable with parameters 2η and $2\sigma^2$. Thus option (1) is correct.
2) Let X and Y be independent and identically distributed exponential random variables. Then the characteristic function of X and Y is given by

$$\Phi_X(\omega) = \frac{\lambda}{1 - j\omega} \tag{11.2.4}$$

The characteristic function of Z is given by

$$\Phi_Z(\omega) = \Phi_X^2(\omega) \tag{11.2.5}$$
$$= \frac{\lambda^2}{(1 - j\omega)^2} \tag{11.2.6}$$

Thus Z is not an exponential random variable. Therefore option (2) is wrong.
3) Let X and Y be independent and identically distributed uniform random variables such that X, Y ~ U(a,b). Then the characteristic function of X and Y is given by

$$\Phi_X(\omega) = \frac{e^{jb\omega} - e^{ja\omega}}{j\omega(b-a)} \tag{11.2.7}$$

The characteristic function of Z is given by

$$\Phi_Z(\omega) = \Phi_X^2(\omega) \tag{11.2.8}$$
$$= -\frac{(e^{jb\omega} - e^{ja\omega})^2}{\omega^2(b-a)^2} \tag{11.2.9}$$

Thus Z is not a uniform random variable. Thus option (3) is wrong.
4) Let X and Y be independent and identically distributed binomial random variables. Then the characteristic function of X and Y is given by

$$\Phi_X(\omega) = (pe^{j\omega} + q)^n \tag{11.2.10}$$

The characteristic function of Z is given by

$$\Phi_Z(\omega) = \Phi_X^2(\omega) \tag{11.2.11}$$
$$= (pe^{j\omega} + q)^{2n} \tag{11.2.12}$$

Thus Z is a binomial random variable with parameter 2n. Thus option (4) is correct.

The following figures show the experimental distributions for Z in each case. The simulation length was kept one million.

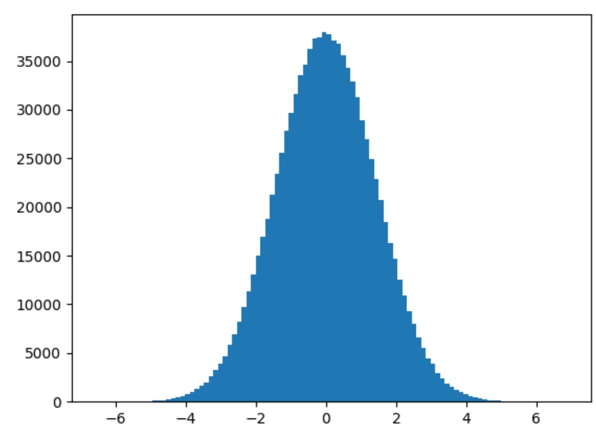


Fig. 11.2.1: Z when X is standard normal

11.3. $N, A_1, A_2 \cdots$ are independent real valued ran-

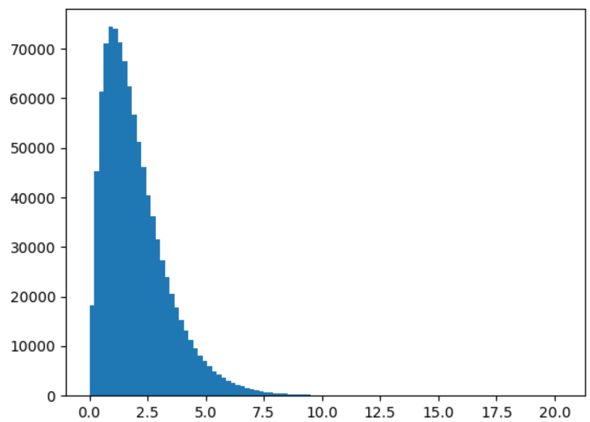


Fig. 11.2.2: Z when X is exponential with $\lambda = 1$

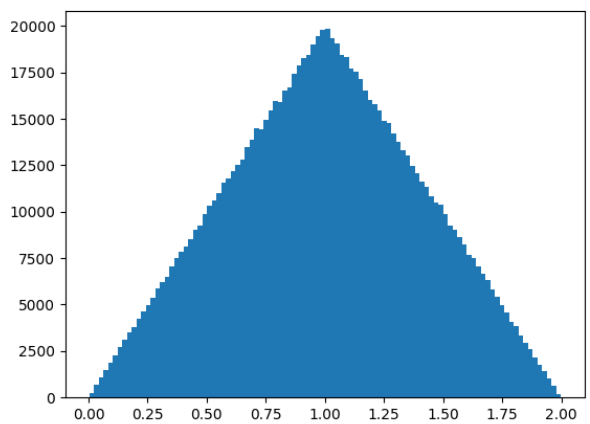


Fig. 11.2.3: Z when X ~ U(0,1)

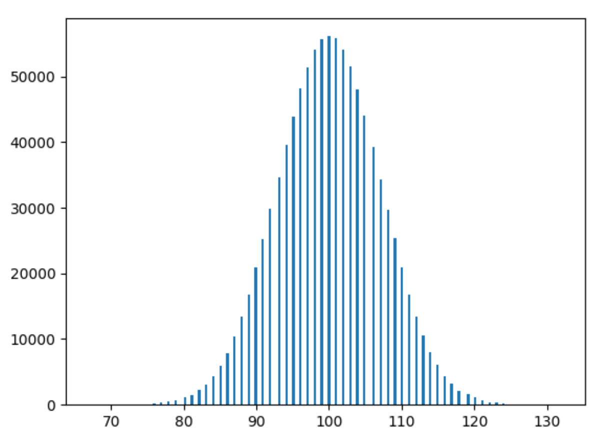


Fig. 11.2.4: Z when X ~ B(100,0.5)

dom variables such that

$$\Pr(N = k) = (1-p)p^k, k = 0, 1, 2, 3\cdots \tag{11.3.1}$$

where $0 < p < 1$ and $\{A_i : i = 1, 2, \cdots\}$ is a sequence of independent and identically distributed bounded random variables. Let

$$X(w) = \begin{cases} 0 & \text{if } N(w) = 0 \\ \sum_{j=1}^{k} A_j & \text{if } N(w) = k, k = 1, 2, 3\cdots \end{cases} \tag{11.3.2}$$

Which of the following are necessarily correct?

a) X is a bounded random variable.
b) Moment generating function m_X of X is

$$m_X(t) = \frac{1-p}{1-pm_A(t)}, t \in \mathbb{R}, \tag{11.3.3}$$

where m_A is moment generating function of A_1.
c) Characteristic function φ_X of X is

$$\varphi_X(t) = \frac{1-p}{1-p\varphi_A(t)}, t \in \mathbb{R}, \tag{11.3.4}$$

where φ_A is the characteristic function of A_1.
d) X is symmetric about 0.

12 Markov Chain

12.1. **Step 1.** Flip a coin twice.
Step 2. If the outcomes are (TAILS, HEADS) then output Y and stop.
Step 3. If the outcomes are either (HEADS, HEADS) or (HEADS, TAILS), then output N and stop.
Step 4. If the outcomes are (TAILS, TAILS), then go to Step 1.
The probability that the output of the experiment is Y is (upto two decimal places)...... **Solution:** The given problem can be represented using Table 12.1.1 and the Markov chain in Fig. 12.1.1. The state transition matrix for the

State	Description
1	$\{T, T\}$
2	$Y = \{T, H\}$
3	$N = \{\{H, H\}, \{H, T\}\}$

TABLE 12.1.1: States and their notations

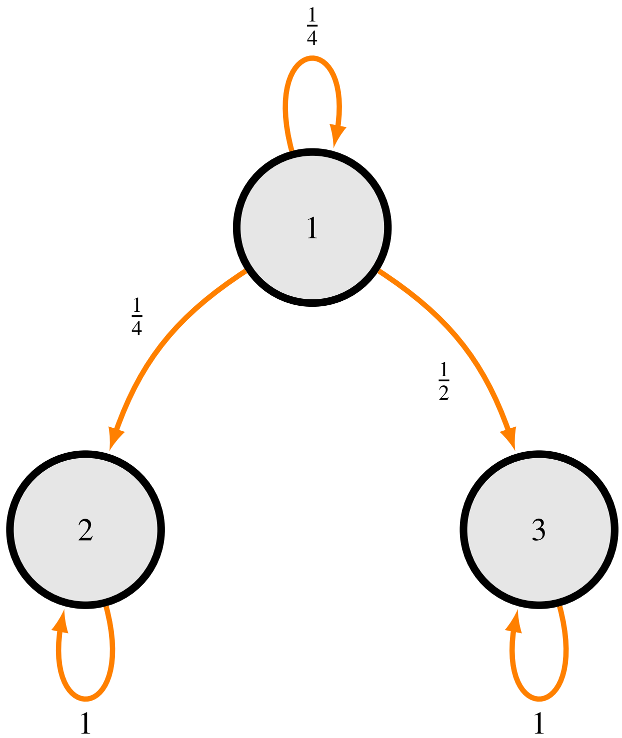


Fig. 12.1.1

Markov chain can be expressed as

$$P = \begin{array}{c c} & \begin{array}{ccc} 2 & 3 & 1 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 1 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} \end{array} \tag{12.1.1}$$

Clearly, the state 1 is transient, while $2, 3$ are absorbing. Comparing (12.1.1) with the standard form of the state transition matrix

$$P = \begin{array}{c c} & \begin{array}{cc} A & N \end{array} \\ \begin{array}{c} A \\ N \end{array} & \begin{bmatrix} I & O \\ R & Q \end{bmatrix} \end{array} \tag{12.1.2}$$

where, From (12.1.1) and (12.1.2),

TABLE 12.1.2: Notations and their meanings

Notation	Meaning
A	All absorbing states
N	All non-absorbing states
I	Identity matrix
O	Zero matrix
R, Q	Other submatices

$$R = \begin{pmatrix} 0.25 & 0.5 \end{pmatrix}, Q = \begin{pmatrix} 0.25 \end{pmatrix} \tag{12.1.3}$$

The limiting matrix for absorbing Markov chain is

$$\bar{P} = \begin{pmatrix} I & O \\ FR & O \end{pmatrix} \tag{12.1.4}$$

where

$$F = (I - Q)^{-1} = (1 - 0.25)^{-1} = \frac{4}{3} \tag{12.1.5}$$

is called the fundamental matrix of P. Upon substituting from (12.1.3) in (12.1.5),

$$F = (1 - 0.25)^{-1} = \frac{4}{3} \tag{12.1.6}$$

and

$$FR = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \end{pmatrix} \tag{12.1.7}$$

which, upon substituting in (12.1.4) yields

$$\bar{P} = \begin{array}{c c} & \begin{array}{ccc} 2 & 3 & 1 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 1 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{bmatrix} \end{array} \tag{12.1.8}$$

$$\therefore \bar{p}_{12} = \frac{1}{3} \tag{12.1.9}$$

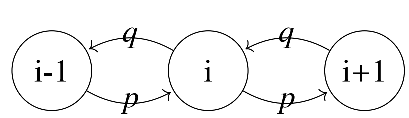
12.2. Consider a simple symmetric random walk on integers ,Where from every state i you to move to i-1 and i+1 with probability half each. Then which of the following are correct?

a) The random walk is aperiodic
b) The random walk is irreducible
c) The random walk is null recurrent
d) The random walk is positive recurrent

Solution: This is a Markov Chain ,Where the state space consists of the integers ($i = 0, \pm1, \pm2, \pm3, ...$) and transition probability is given as

$$P_{i,i+1} = p = \frac{1}{2} \tag{12.2.1}$$

$$P_{i,i-1} = q = \frac{1}{2} \tag{12.2.2}$$



Let $P^n_{i,j}$ denotes the probability of being in state j after nth transition starting from state i.

a) We know that for state j in Markov chain to be **aperiodic** ,Then their exist k such that $P^n_{j,j} > 0$ for all $n \geq k$. but for to return to same state j after n transitions ,Number of forward steps should be equal to Backward steps , i.e for odd n in (2m +1)form

$$P^{2m+1}_{j,j} = 0 \tag{12.2.3}$$

when n is even in 2m form

$$P^{2m}_{j,j} = \binom{2m}{m} p^m q^m \tag{12.2.4}$$

$$= \frac{(2m)!}{m!.m!} p^m q^m \tag{12.2.5}$$

,As for odd n $P^n_{j,j} = 0$,$P^n_{j,j} > 0$ for all $n \geq k$ is not possible .which implies all states are **Periodic**
Option (1) is **incorrect**.

b) In a Markov Chain for state j to be recurrent then it should satisfy following condition

$$\lim_{t\to\infty} \sum_{n=1}^{t} P^n_{j,j} = \infty \tag{12.2.6}$$

using Stirling approximation in equation (12.2.5)

$$P^{2m}_{j,j} = \frac{((2m)^{2m+\frac{1}{2}}).e^{-2m}.(2\pi)^{\frac{1}{2}}}{m^{m+\frac{1}{2}}.e^{-m}.m^{m+\frac{1}{2}}.e^{-m}.2\pi}.p^m q^m \tag{12.2.7}$$

$$= \frac{(4pq)^{2m}}{(m\pi)^{\frac{1}{2}}} \tag{12.2.8}$$

In this question $p = \frac{1}{2} = q$,then using (12.2.3) and (12.2.8)

$$\lim_{t\to\infty} \sum_{n=1}^{t} P^n_{j,j} = \sum_{n=2k,k=1}^{\infty} \frac{1}{(\frac{n}{2}\pi)^{\frac{1}{2}}} \tag{12.2.9}$$

Since $\frac{1}{n^{\frac{1}{2}}}$ is divergent,

$$\lim_{t\to\infty} \sum_{n=1}^{t} P^n_{j,j} = \infty \tag{12.2.10}$$

Therefore state j is recurrent ,as what we calculated is independent of j ,all states are **recurrent** . The first-passage-time probability, $f_{i,j}(n)$, of a Markov chain is the probability,given as

$$f_{i,j}(n) = \Pr(X_n = j, X_{n-1} \neq j, X_{n-2} \neq j, \ldots X_1 \neq j | X_0 = i) \tag{12.2.11}$$

The first-passage time $T_{j,j}$from a state j back to itself is of particular importance. It has the PMF $f_{j,j}(n)$ amd Distribution function $F_{j,j}(n)$

$$F_{j,j}(n) = \sum_{k=0}^{n} f_{j,j}(k) \tag{12.2.12}$$

We Know that all states are recurrent .Now i will find whether they are null recurrent or positive recurrent . For positive recurrent

$$\overline{T_{j,j}} < \infty \tag{12.2.13}$$

For null recurrent

$$\overline{T_{j,j}} = \infty \tag{12.2.14}$$

Where $\overline{T_{j,j}}$ is mean time to enter state j starting from j. Now calculating $\overline{T_{j,j}}$ using below formula,

$$\overline{T_{j,j}} = 1 + \sum_{k=0}^{n} (1 - F_{j,j}(k)) \tag{12.2.15}$$

Using (12.2.15) and (12.2.12),We get

$$\overline{T_{j,j}} = \infty \tag{12.2.16}$$

Therefore all states are null recurrent. Option(3) is **correct**

c) Since all states are recurrent,they communicate with each other ,therefore Markov chain is irreducible , option (2) is **correct**

d) As all states are null recurrent , option (4) is **incorrect**

Therefore correct options are **2,3**

12.3. Consider a Markov Chain with state space $\{0, 1, 2\}$ and transition matrix

$$P = \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} \frac{1}{4} & \frac{5}{8} & \frac{1}{8} \\ \frac{1}{4} & 0 & \frac{3}{4} \\ \frac{1}{2} & \frac{3}{8} & \frac{1}{8} \end{pmatrix} \end{array} \tag{12.3.1}$$

Then which of the following are true?

a) $\lim_{n\to\infty} p^{(n)}_{12} = 0$
b) $\lim_{n\to\infty} p^{(n)}_{12} = \lim_{n\to\infty} p^{(n)}_{21}$
c) $\lim_{n\to\infty} p^{(n)}_{22} = \frac{1}{8}$
d) $\lim_{n\to\infty} p^{(n)}_{21} = \frac{1}{3}$

12.4. Consider a Markov chain with state space 1,2,....,100. Suppose states 2i and 2j communicate with each other and states 2i-1 and 2j-1 communicate with each other for every i,j = 1,2,...,50. Further suppose that $p_{3,3}^{(2)}$ ¿ 0,$p_{4,4}^{(3)}$ ¿ 0 and $p_{2,5}^{(7)}$ ¿ 0. Then

a) The Markov chain is irreducible.
b) The Markov chain is aperiodic.
c) State 8 is recurrent.
d) State 9 is recurrent.

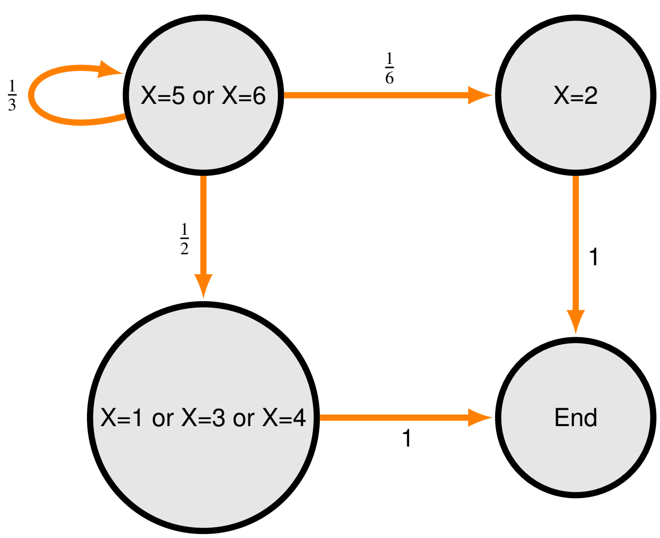
Solution:

12.5. A standard fair die is rolled until some face other than 5 or 6 turns up.Let X denote the face value of the last roll.Let A={X is even} and B={X is atmost 2} Then,

a) $\Pr(A \cap B) = 0$
b) $\Pr(A \cap B) = \frac{1}{6}$
c) $\Pr(A \cap B) = \frac{1}{4}$
d) $\Pr(A \cap B) = \frac{1}{3}$

Solution: Let us assume the following table.

Fig. 12.5.1: Markov chain



Let us represent the markov chain diagram in a

TABLE 12.5.1

state 1	state 2	state 3	state 4
$X = 5$ or $X = 6$	$X = 2$	$X = 1$ or $X = 3$ or $X = 4$	end

matrix.Let P_{ij} represent the element of a matrix which is in i^{th} row and j^{th} column.The value of P_{ij} is equal to probability of transition from state i to state j

$$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{12.5.1}$$

We need the probability that $X = 2$.Hence required probability is

$$P_{12} + (P_{12})^2 + \cdots + \infty \tag{12.5.2}$$

where P_{12}^n represents the 1st row ,2nd column element in the P^n

$$P^2 = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{12.5.3}$$

$$= \begin{bmatrix} \frac{1}{9} & \frac{1}{18} & \frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{12.5.4}$$

$$P^3 = (P^2)(P^1) \tag{12.5.5}$$

$$= \begin{bmatrix} \frac{1}{9} & \frac{1}{18} & \frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \frac{1}{3} & \frac{1}{6} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{12.5.6}$$

$$= \begin{bmatrix} \frac{1}{27} & \frac{1}{54} & \frac{1}{18} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{12.5.7}$$

From above we can notice that each time P_{12} reduces by $\frac{1}{3}$.Hence from (12.5.2),

$$\sum_{i=0}^{\infty} \left(\frac{1}{3}\right)^i \frac{1}{6} \tag{12.5.8}$$

From Geometric progression we can write ,required probability $=\frac{1}{4}$ $\therefore$ **option C is correct**

12.6. Consider a Markov chain with five states $\{1, 2, 3, 4, 5\}$ and transition matrix

$$P = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{7} & 0 & 0 & \frac{6}{7} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} & 0 \\ 0 & \frac{5}{8} & 0 & 0 & \frac{3}{8} \end{pmatrix} \tag{12.6.1}$$

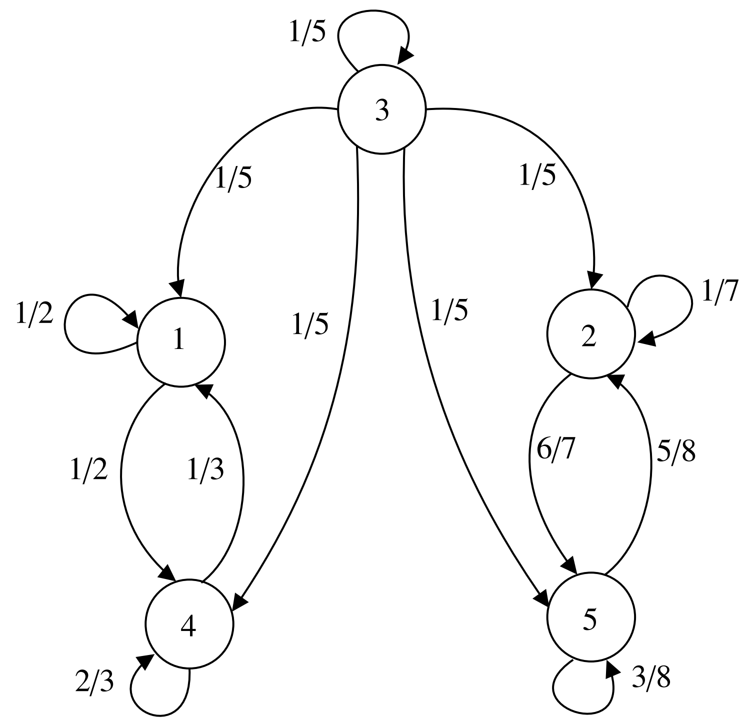
Which of the following are true?

Accessibility of states in Markov's chain	We say that state j is accessible from state i, written as $i \to j$, if $p_{ij}^{(n)} > 0$ for some n. Every state is accessible from itself since $p_{ii}^{(0)} = 1$
Communication between states	Two states i and j are said to communicate, written as $i \leftrightarrow j$, if they are accessible from each other. In other words, $$i \leftrightarrow j \text{ means } i \to j \text{ and } j \to i.$$
Communicating class	For each Markov chain, there exists a unique decomposition of the state space S into a sequence of disjoint subsets $C_1, C_2, ...,$ $$S = \bigcup_{i=1}^{\infty} C_i$$ in which each subset has the property that all states within it communicate. Each such subset is called a communication class of the Markov chain.

TABLE 12.6.1: Definition and Result used

a) 3 and 1 are in the same communicating class
b) 1 and 4 are in the same communicating class
c) 4 and 2 are in the same communicating class
d) 2 and 5 are in the same communicating class

Solution: See Tables 12.6.1 and 12.6.2

<table>
<tr>
<td>Drawing Transition diagram</td>
<td>


</td>
</tr>
<tr>
<td>Checking whether the states 3 and 1 are in the same communicating class</td>
<td>Here,
State 1 is accessible from the state 3.
But, State 3 is not accessible from the state 1
i.e. $3 \rightarrow 1,\ 1 \nrightarrow 3$
$\implies \boxed{3 \nleftrightarrow 1}$

Therefore, 3 and 1 are not in the same communicating class.</td>
</tr>
<tr>
<td>Checking whether the states 1 and 4 are in the same communicating class</td>
<td>Here,
State 1 is accessible from the state 4.
Also, State 4 is accessible from the state 1
i.e. $3 \rightarrow 1,\ 1 \rightarrow 3$
$\implies \boxed{3 \leftrightarrow 1}$

Therefore, 1 and 4 are in the same communicating class.</td>
</tr>
<tr>
<td>Checking whether the states 4 and 2 are in the same communicating class</td>
<td>Here,
State 2 is not accessible from the state 4.
Also, State 4 is not accessible from the state 2
i.e. $4 \nrightarrow 2,\ 2 \nrightarrow 4$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</table>

	$\implies \boxed{4 \nleftrightarrow 2}$ Therefore, 4 and 2 are not in the same communicating class.
Checking whether the states 2 and 5 are in the same communicating class	Here, State 2 is accessible from the state 5. Also, State 5 is accessible from the state 2 i.e. $5 \to 2,\ 2 \to 5$ $\implies \boxed{2 \leftrightarrow 5}$ Therefore, 2 and 5 are in the same communicating class.
Conclusion	Communication classes are: $\boxed{S = \{1,4\} \cup \{3\} \cup \{2,5\}}$ Option 2) and 4) are true.

TABLE 12.6.2: Solution

12.7. A and B play a game of tossing a fair coin. A starts the game by tossing the coin once and B then tosses the coin twice, followed by A tossing the coin once and B tossing the coin twice and this continues until a head turns up. Whoever gets the first head wins the game. Then,

a) $P(B \text{ Wins}) > P(A \text{ Wins})$
b) $P(B \text{ Wins}) = 2P(A \text{ Wins})$
c) $P(A \text{ Wins}) > P(B \text{ Wins})$
d) $P(A \text{ Wins}) = 1 - P(B \text{ Wins})$

Solution: Given, a fair coin is tossed till heads turns up.

$$p = \frac{1}{2}, q = \frac{1}{2} \tag{104.1}$$

Let's define a Markov chain $\{X_n, n = 0, 1, 2, \dots\}$, where $X_n \in S = \{1, 2, 3, 4, 5\}$, such that The state transition matrix for the Markov

TABLE 12.7.1: States and their notations

Notation	State
$S = 1$	A's turn
$S = 2$	B's first turn
$S = 3$	B's second turn
$S = 4$	A wins
$S = 4$	B wins

chain is

$$P = \begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \tag{104.2}$$

Clearly, the states $1, 2, 3$ are transient, while $4, 5$ are absorbing. The standard form of a state transition matrix is

$$P = \begin{array}{c c} & \begin{array}{cc} A & N \end{array} \\ \begin{array}{c} A \\ N \end{array} & \begin{bmatrix} I & O \\ R & Q \end{bmatrix} \end{array} \tag{104.3}$$

where, Converting (104.2) to standard form, we

TABLE 12.7.2: Notations and their meanings

Notation	Meaning
A	All absorbing states
N	All non-absorbing states
I	Identity matrix
O	Zero matrix
R, Q	Other submatices

get

$$P = \begin{array}{c c} & \begin{array}{ccccc} 4 & 5 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 4 \\ 5 \\ 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 & 0 \end{bmatrix} \end{array} \tag{104.4}$$

From (104.4),

$$R = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix}, Q = \begin{bmatrix} 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \\ 0.5 & 0 & 0 \end{bmatrix} \tag{104.5}$$

The limiting matrix for absorbing Markov chain is

$$\bar{P} = \begin{bmatrix} I & O \\ FR & O \end{bmatrix} \tag{104.6}$$

where,

$$F = (I - Q)^{-1} \tag{104.7}$$

is called the fundamental matrix of P.
On solving, we get

$$\bar{P} = \begin{array}{c} \\ 4 \\ 5 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccccc} 4 & 5 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.5714 & 0.4285 & 0 & 0 & 0 \\ 0.1428 & 0.8571 & 0 & 0 & 0 \\ 0.2857 & 0.7142 & 0 & 0 & 0 \end{bmatrix} \end{array} \tag{104.8}$$

A element $\bar{p}_{ij}$ of $\bar{P}$ denotes the absorption probability in state j, starting from state i. Then,

a) $Pr(A$ wins$)=\bar{p}_{14} \approx 0.5714$
b) $Pr(B$ wins$)=\bar{p}_{15} \approx 0.4285$

$$\therefore \bar{p}_{14} > \bar{p}_{15} \tag{104.9}$$

Also, in $\bar{P}$, all the terms in every row should sum to 1.

$$\Rightarrow \bar{p}_{14} + \bar{p}_{15} + 0 + 0 + 0 = 1 \tag{104.10}$$
$$\therefore \bar{p}_{14} = 1 - \bar{p}_{15} \tag{104.11}$$

Therefore, options 3), 4) are correct.

12.8. Consider a Markov chain with state space 1,2,....,100. Suppose states 2i and 2j communicate with each other and states 2i-1 and 2j-1 communicate with each other for every i,j = 1,2,...,50. Further suppose that $p_{3,3}^{(2)}$ ¿ 0,$p_{4,4}^{(3)}$ ¿ 0 and $p_{2,5}^{(7)}$ ¿ 0. Then

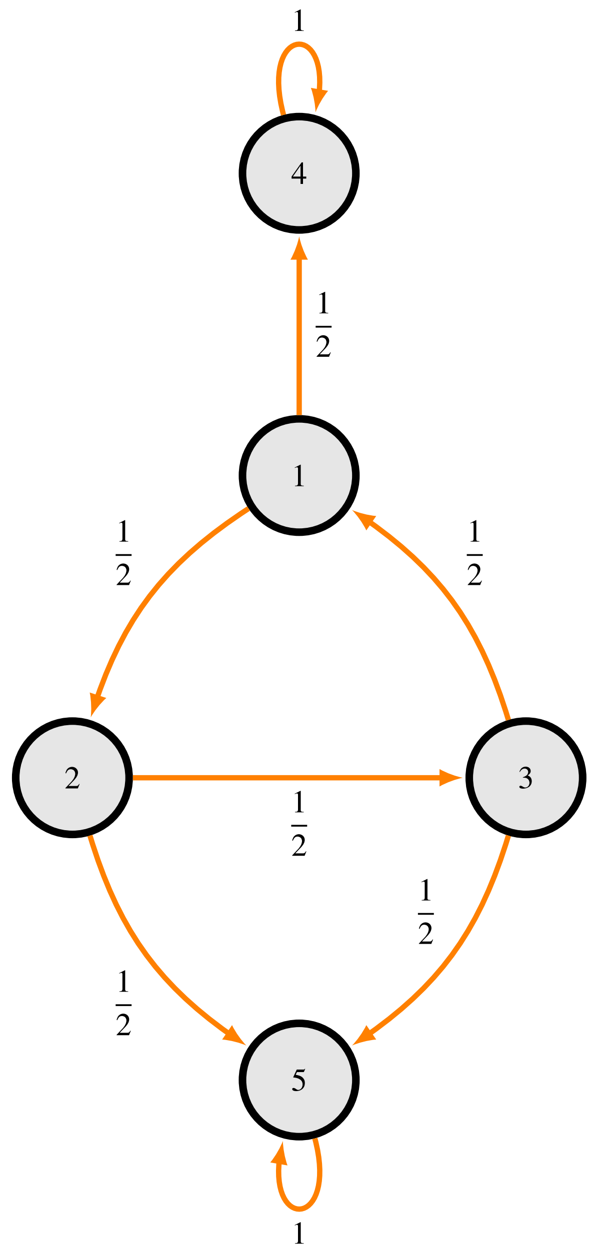
a) The Markov chain is irreducible.
b) The Markov chain is aperiodic.
c) State 8 is recurrent.
d) State 9 is recurrent.

Solution:

13 Inequalities

13.1. Let $X_i's$ be independent random variables such that $X_i's$ are symmetric about 0

Markov chain diagram



and $var(X_i) = 2i - 1$,for $i \geq 1$.then, $\lim_{n\to\infty} \Pr\left(X_1 + X_2 + \cdots + X_n > n \log n\right)$

a) does not exist.
b) equals $\frac{1}{2}$.
c) equals 1.
d) equals 0.

Solution: Let $X = X_1 + X_2 + \cdots + X_n$, as $X_i's$ are symmetric about 0. The mean of X is given by,

$$E[X] = 0 \tag{13.1.1}$$

the variance of X is given by,

$$var[X] = \sum_{i=1}^{n}(2i-1) \tag{13.1.2}$$
$$= \frac{2n(n+1)}{2} - n \tag{13.1.3}$$
$$= n^2 \tag{13.1.4}$$

the standard deviation,

$$\sigma_X = n \tag{13.1.5}$$

Applying Chebyshev's Inequality for the random variable X, for any $k > 0$

$$\Pr(|X - E[X]| > k\sigma_X) \le \frac{1}{k^2} \tag{13.1.6}$$

let $k = \log n$, using (13.1.1) and (13.1.5) in (13.1.6),

$$\Pr(|X| > n\log n) \le \frac{1}{(\log n)^2} \tag{13.1.7}$$

$$\Pr(X > n\log n) + \Pr(X < -n\log n) \le \frac{1}{(\log n)^2} \tag{13.1.8}$$

As, X is symmetric about 0,

$$\Pr(X > n\log n) = \Pr(X < -n\log n) \tag{13.1.9}$$

using (13.1.9) in (13.1.8),

$$2\Pr(X > n\log n) \le \frac{1}{(\log n)^2} \tag{13.1.10}$$
$$\Pr(X > n\log n) \le \frac{1}{2(\log n)^2} \tag{13.1.11}$$

as any probability is greater than 0,

$$0 < \Pr(X > n\log n) \le \frac{1}{2(\log n)^2} \tag{13.1.12}$$

applying sandwich principle to (13.1.12),

$$\lim_{n\to\infty} 0 < \lim_{n\to\infty} \Pr(X > n\log n) \le \lim_{n\to\infty} \frac{1}{2(\log n)^2} \tag{13.1.13}$$

$$\lim_{n\to\infty} \Pr(X_1 + X_2 + \cdots + X_n > n\log n) = 0 \tag{13.1.14}$$

Hence the option.4 is correct.

13.2. Let X_1,X_2,... be independent random variables each following exponential distribution with mean 1. Then which of the following statements are correct?

a) $\mathrm{P}(X_n > \log n$ for infinitely many $n \ge 1) = 1$
b) $\mathrm{P}(X_n > 2$ for infinitely many $n \ge 1) = 1$
c) $\mathrm{P}(X_n > \frac{1}{2}$ for infinitely many $n \ge 1) = 0$
d) $\mathrm{P}(X_n > \log n, X_{n+1} > \log(n+1)$ for infinitely many $n \ge 1) = 0$

Solution: PDF of X_i is

$$f_{X_i}(x) = \begin{cases} \lambda_i e^{-\lambda_i x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

Mean of X_i is expressed as

$$\mathrm{E}(X_i) = \int_{-\infty}^{\infty} x f_{X_i}(x) dx$$
$$= \int_{-\infty}^{0} 0 dx + \int_{0}^{\infty} x\lambda_i e^{-\lambda_i x}$$
$$= \frac{1}{\lambda_i} \tag{13.2.1}$$

From (13.2.1)and $\mathrm{E}(X_i) = 1$, we have $\lambda_i = 1 \forall i \ge 1$ Now, for some constant $c \ge 0$

$$\Pr(X_n > c) = \int_{c}^{\infty} f_{X_n}(x) dx$$
$$= \int_{c}^{\infty} e^{-x} dx$$
$$= e^{-c} \tag{13.2.2}$$

Borel-Cantelli Lemma:
Let E_1,E_2,... be a sequence of events in some probability space. The Borel–Cantelli lemma states that, if the sum of the probabilities of the events E_n is finite

$$\sum_{n=1}^{\infty} \Pr(E_n) < \infty \tag{13.2.3}$$

then the probability that infinitely many of them occur is 0

$$\Pr\left(\lim_{n\to\infty} \sup E_n\right) = 0 \tag{13.2.4}$$

Second Borel-Cantelli Lemma:
If the events E_n are independent and the sum of the probabilities of the E_n diverges to infinity, then the probability that infinitely many of them occur is 1. If for independent events

$E_1, E_2, ...$

$$\sum_{n=1}^{\infty} \Pr(E_n) = \infty \tag{13.2.5}$$

Then

$$\Pr\left(\limsup_{n\to\infty} E_n\right) = 1 \tag{13.2.6}$$

a) **Option 1:** We can say the events $X_n > \log n$ are independent $\forall n \geq 1$ as X_n are independent random variable.
From (13.2.2)

$$\sum_{n=1}^{\infty} \Pr(X_n > \log n) = \sum_{n=1}^{\infty} e^{-\log n}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n}$$
$$= \infty \text{ (Cauchy's Criterion)}$$

Now, from second Borel-Cantelli lemma

$$\Pr(X_n > \log n \text{ for infinitely many } n \geq 1)$$
$$= \Pr\left(\limsup_{n\to\infty} X_n > \log n\right)$$
$$= 1$$

$\therefore$ Option 1 is correct.

b) **Option 2:** We can say the events $X_n > 2$ are independent $\forall n \geq 1$ as X_n are independent random variable.
From (13.2.2)

$$\sum_{n=1}^{\infty} \Pr(X_n > 2) = \sum_{n=1}^{\infty} e^{-2}$$
$$= \infty$$

Now, from second Borel-Cantelli lemma

$$\Pr(X_n > 2 \text{ for infinitely many } n \geq 1)$$
$$= \Pr\left(\limsup_{n\to\infty} X_n > 2\right)$$
$$= 1$$

$\therefore$ Option 2 is correct.

c) **Option 3:** We can say the events $X_n > \frac{1}{2}$ are independent $\forall n \geq 1$ as X_n are independent random variable.
From (13.2.2)

$$\sum_{n=1}^{\infty} \Pr\left(X_n > \frac{1}{2}\right) = \sum_{n=1}^{\infty} e^{-\frac{1}{2}}$$
$$= \infty$$

Now, from second Borel-Cantelli lemma

$$\Pr\left(X_n > \frac{1}{2} \text{ for infinitely many } n \geq 1\right)$$
$$= \Pr\left(\limsup_{n\to\infty} X_n > \frac{1}{2}\right)$$
$$= 1$$

$\therefore$ Option 3 is incorrect.

d) **Option 4:** We can say the events $X_n > \log n$ are independent $\forall n \geq 1$ as X_n are independent random variable.
Let the event $X_n > \log n, X_{n+1} > \log(n+1)$ be represented by E_n'
From (13.2.2)

$$\sum_{n=1}^{\infty} \Pr(E_n)$$
$$= \sum_{n=1}^{\infty} \Pr(X_n > \log n) \Pr(X_{n+1} > \log(n+1))$$
$$= \sum_{n=1}^{\infty} e^{-\log n} e^{-\log(n+1)}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n(n+1)}$$
$$= \sum_{n=1}^{\infty} \frac{1}{n} - \frac{1}{n+1}$$
$$= 1 \tag{13.2.7}$$

Now, from Borel-Cantelli lemma

$$\Pr(E_n \text{ for infinitely many } n \geq 1)$$
$$= \Pr\left(\limsup_{n\to\infty} (X_n > \log n, X_{n+1} > \log(n+1))\right)$$
$$= 0$$

$\therefore$ Option 4 is correct.

$\boxed{\textbf{Solution: } \text{Options 1, 2, 4}}$

14 Convergence

14.1. Let $X_1, X_2, \ldots$ be independent and identically distributed random variables each following a uniform distribution on (0,1). Denote

$$T_n = \max\{X_1, X_2, \dots, X_n\}. \quad (14.1.1)$$

Then, which of the following statements are true?

a) T_n converges to 1 in probability.
b) $n(1 - T_n)$ converges in distribution.
c) $n^2(1 - T_n)$ converges in distribution.
d) $\sqrt{n}(1 - T_n)$ converges to 0 in probability.

Solution:

Definition 11. Random Sampling : *A collection of random variables $X_1, X_2, \dots, X_n$ is said to be a random sample of size n if they are independent and identically distributed, i.e,*

a) $X_1, X_2, \dots, X_n$ are independent random variables
b) They have the same distribution (Let us denote it by $F_X(x)$), i.e,

$$F_X(x) = F_{X_i}(x), i = 1, 2, \dots, n \forall x \in \mathbb{R} \quad (14.1.2)$$

Definition 12. Order Statistics : *Given a random sample $X_1, X_2, \dots, X_n$, the sequence $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ is called the order statistics of it. Here,*

$$X_{(1)} = \min(X_1, X_2, \dots, X_n) \quad (14.1.3)$$

$$X_{(2)} = \textit{the } 2^{nd} \textit{ smallest of } X_1, X_2, \dots, X_n \quad (14.1.4)$$

$$\vdots \quad (14.1.5)$$

$$X_{(n)} = \max(X_1, X_2, \dots, X_n) \quad (14.1.6)$$

Lemma 14.1. Distribution of the maximum :

$$f_{T_n}(x) = \begin{cases} nx^{n-1}, & 0 < x < 1 \\ 0, & otherwise \end{cases} \quad (14.1.7)$$

$$F_{T_n}(x) = \begin{cases} x^n, & 0 < x < 1 \\ 1, & x \geq 1 \\ 0, & otherwise \end{cases} \quad (14.1.8)$$

Proof:

$$F_{X_{(n)}}(x) = \Pr(X_{(n)} \leq x) \quad (14.1.9)$$

$$= \Pr(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \quad (14.1.10)$$

$$= \Pr(X_1 \leq x)\Pr(X_2 \leq x)\dots\Pr(X_n \leq x) \quad (14.1.11)$$

$$= [\Pr(X_1 \leq x)]^n \text{ (i.i.d)} \quad (14.1.12)$$

$$= [F_X(x)]^n \quad (14.1.13)$$

and

$$f_{X_{(n)}}(x) = \frac{d}{dx}\left(F_{X_{(n)}}(x)\right) = \frac{d}{dx}\left([F_X(x)]^n\right) \quad (14.1.14)$$

$$= n\left([F_X(x)]^{n-1}\right)\frac{d}{dx}(F_X(x)) \quad (14.1.15)$$

$$= n[F_X(x)]^{n-1} f_X(x) \quad (14.1.16)$$

$\because$

$$f_{X_i}(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & otherwise \end{cases}, \quad (14.1.17)$$

$$F_{X_i}(x) = \begin{cases} x, & 0 < x < 1 \\ 1, & x \geq 1 \\ 0, & otherwise, \end{cases} \quad (14.1.18)$$

$\forall i \in \mathbb{N}$. Substituting the above in (14.1.16) and (14.1.13) yields (14.1.7) and (14.1.8) respectively. Then, as $T_n = max\{X_1, X_2, \dots, X_n\} = X_{(n)}$,

Lemma 14.2. *If $Y = aX + b$ and $a < 0$, then*

$$F_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right) \quad (14.1.19)$$

Definition 13. Convergence in Probability : *A sequence of random variables $X_1, X_2, X_3, \dots$ converges in probability to a random variable X, shown by $X_n \xrightarrow{p} X$, if*

$$\lim_{n\to\infty} \Pr(|X_n - X| \geq \epsilon) = 0, \forall \epsilon > 0 \quad (14.1.20)$$

Definition 14. Convergence in Distribution : *A sequence of random variables $X_1, X_2, X_3, \dots$ converges in distribution to a random variable X, shown by $X_n \xrightarrow{d} X$, if*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad (14.1.21)$$

for all x at which $F_X(x)$ is continuous.

a)

$$\lim_{n\to\infty} \Pr(|T_n - 1| \geq \epsilon) = \lim_{n\to\infty} \Pr(1 - T_n \geq \epsilon)$$
$$= \lim_{n\to\infty} \Pr(T_n \leq 1 - \epsilon) = \lim_{n\to\infty} F_{T_n}(1-\epsilon) \tag{14.1.22}$$

$$\because F_{T_n}(1-\epsilon) = \begin{cases} (1-\epsilon)^n, & 0 < \epsilon < 1 \\ 0, & \epsilon \geq 1 \end{cases} \tag{14.1.23}$$

and

$$\because \lim_{n\to\infty}(1-\epsilon)^n = 0 \text{ for } 0 < \epsilon < 1 \tag{14.1.24}$$
$$\tag{14.1.25}$$

from (14.1.24), (14.1.23) and (14.1.22),

$$\lim_{n\to\infty} \Pr(|T_n - 1| \geq \epsilon) = 0, \forall \epsilon > 0 \tag{14.1.26}$$

$\therefore T_n$ converges to 1 in probability.

b) Substituting $a = -n, b = n$ in (14.1.19),

$$F_{n(1-T_n)}(x) = 1 - F_{T_n}\left(1 - \frac{x}{n}\right) \tag{14.1.27}$$
$$\tag{14.1.28}$$

where

$$F_{T_n}\left(1 - \frac{x}{n}\right) = \begin{cases} \left(1 - \frac{x}{n}\right)^n, & 0 < x < n \\ 1, & x \leq 0 \\ 0, & x \geq n \end{cases} \tag{14.1.29}$$
$$from (14.1.8) \tag{14.1.30}$$

$$\because \lim_{n\to\infty}\left(1 - \frac{y}{n}\right)^n = e^{-y}, \tag{14.1.31}$$
$$\tag{14.1.32}$$

$$\therefore \lim_{n\to\infty} F_{T_n}\left(1 - \frac{x}{n}\right) = \begin{cases} e^{-x}, & x > 0 \\ 1, & x \leq 0 \end{cases} \tag{14.1.33}$$

$$\implies \lim_{n\to\infty} F_{n(1-T_n)}(x) = 1 - \lim_{n\to\infty} F_{T_n}\left(1 - \frac{x}{n}\right) \tag{14.1.34}$$

which can be expressed as

$$\therefore \lim_{n\to\infty} F_{n(1-T_n)}(x) = \begin{cases} 1 - e^{-x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{14.1.35}$$

$\therefore n(1 - T_n)$ converges in distribution to the random variable $X \sim Exponential(1)$.

c) Substituting $a = -n^2, b = n^2$ in (14.1.19),

$$F_{n^2(1-T_n)}(x) = 1 - F_{T_n}\left(1 - \frac{x}{n^2}\right) \tag{14.1.36}$$

$$F_{T_n}\left(1 - \frac{x}{n^2}\right) = \begin{cases} \left(1 - \frac{x}{n^2}\right)^n, & 0 < x < n^2 \\ 1, & x \leq 0 \\ 0, & x \geq n^2 \end{cases} \tag{14.1.37}$$

$$= \begin{cases} 1, & x > 0 \\ 1, & x \leq 0 \end{cases} \tag{14.1.38}$$

$$\because \lim_{n\to\infty}\left(1 - \frac{x}{n^2}\right)^n = 1 \tag{14.1.39}$$

yielding

$$\lim_{n\to\infty} F_{n^2(1-T_n)}(x) = \begin{cases} 0, & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{14.1.40}$$

which is not a valid CDF. Hence, $n^2(1 - T_n)$ does not converge in distribution.

d)

$$\lim_{n\to\infty} \Pr\left(|\sqrt{n}(1 - T_n) - 0| \geq \epsilon\right)$$
$$= \lim_{n\to\infty} \Pr\left(1 - T_n \geq \frac{\epsilon}{\sqrt{n}}\right)$$
$$= \lim_{n\to\infty} \Pr\left(T_n \leq 1 - \frac{\epsilon}{\sqrt{n}}\right)$$
$$= \lim_{n\to\infty} F_{T_n}\left(1 - \frac{\epsilon}{\sqrt{n}}\right)$$
$$= \begin{cases} \left(1 - \frac{\epsilon}{\sqrt{n}}\right)^n, & 0 < \epsilon < \sqrt{n} \\ 0, & \epsilon \geq \sqrt{n} \end{cases} \tag{14.1.41}$$

$$\because \lim_{n\to\infty}\left(1-\frac{\epsilon}{\sqrt{n}}\right)^n = 0 \text{ for } 0<\epsilon<\sqrt{n},$$

$$\lim_{n\to\infty}\Pr\left(|\sqrt{n}(1-T_n)-0|\geq\epsilon\right)=0, \forall\epsilon>0 \quad (14.1.42)$$

$\therefore\ \sqrt{n}(1-T_n)$ converges to 0 in probability.

Hence, options 1), 2), 4) are correct.

14.2. Let $\{X_i\}_{i\geq 1}$ be a sequence of i.i.d. random variables with $\mathrm{E}(X_i) = 0$ and $\mathrm{var}(X_i) = 1$. Which of the following are true?

a) $\frac{1}{n}\sum_{i=1}^n X_i^2 \to 0$ in probability

b) $\frac{1}{n^{3/4}}\sum_{i=1}^n X_i \to 0$ in probability

c) $\frac{1}{n^{1/2}}\sum_{i=1}^n X_i \to 0$ in probability

d) $\frac{1}{n}\sum_{i=1}^n X_i^2 \to 1$ in probability

Solution:

Definition 15. *(Convergence in distribution) A sequence of random variables Y, Y_1, $Y_2\dots$ converges in distribution to a random variable Y, if*

$$\lim_{n\to\infty} F_{X_n}(a) = F_X(a)\ \forall a\in\mathbb{R}. \quad (14.2.1)$$

Definition 16. *(Convergence in probability) A sequence of random variables Y, Y_1, $Y_2\dots$ is said to converge in probability to Y, if*

$$\lim_{n\to\infty}\Pr(|Y_n-Y|>\epsilon)=0\ \forall\epsilon>0. \quad (14.2.2)$$

Lemma 14.3. *If $Y_n\to Y$ in probability, $Y_n\to Y$ in distribution.*

Lemma 14.4. *(Strong Law of Large Numbers) Let X_1, X_2, $\dots X_n$ be i.i.d. random variables with expected value $E(X_i)=\mu<\infty$, then,*

$$X_i \xrightarrow{p} \mu \quad (14.2.3)$$

Or,

$$\frac{1}{n}\sum_{i=1}^n X_i \xrightarrow{p} \mu \quad (14.2.4)$$

Lemma 14.5. *If X_i is a sequence of i.i.d. random variables, with*

$$F_{X_i}(x) = F_X(x), \quad (14.2.5)$$

then,

$$F_{X_i^2}(x) = F_{X^2}(x) \quad (14.2.6)$$

$\forall x\in\mathbb{R}$, where $F_X(x)$ is the c.d.f. of X_i.

Proof.

$$F_{X_i{}^2}(y) = \Pr\left(X_i^2\leq y\right) \quad (14.2.7)$$

$$\implies F_{Y_i}(y) = \Pr(-\sqrt{y}\leq X_i\leq\sqrt{y}) \quad (14.2.8)$$

$$\implies F_{X_i{}^2}(y) = F_{X_i}(\sqrt{y}) - F_{X_i}(-\sqrt{y}) \quad (14.2.9)$$

$$= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \quad (14.2.10)$$

$$= F_{X^2}(y) \quad (14.2.11)$$

Using (14.2.5), □

Corollary 14.1. *If X_i are i.i.d, X_i^2 are i.i.d.*

Proof. Let $Y_i = X_i^2$.

$$\begin{aligned} &F_{Y_1,Y_2,\dots,Y_n}(y_1,y_2,\dots,y_n) \\ &\quad = \Pr(Y_1\leq y_1, Y_2\leq y_2,\dots,Y_n\leq y_n) \\ &\quad = \Pr\left(X_1^2\leq y_1, X_2^2\leq y_2,\dots,X_n^2\leq y_n\right) \\ &\quad = \prod_{i=1}^n [\Pr(-\sqrt{y_i}\leq X_i\leq\sqrt{y_i})] \end{aligned}$$

$$= \prod_{i=1}^n F_Y(y_i) \quad (14.2.12)$$

$\implies X_i^2$ are i.i.d. □

Lemma 14.6. *If $X_1, X_2,\dots X_n$ are independent random variables,*

$$var\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n var(X_i) \quad (14.2.13)$$

Lemma 14.7. *(Chebyshev's Inequality) Let the random variable X have a finite mean μ and a finite variance σ^2. For every $\epsilon>0$,*

$$\Pr(|X-\mu|\geq\epsilon)\leq\frac{\sigma^2}{\epsilon^2} \quad (14.2.14)$$

Lemma 14.8. *(Central Limit Theorem) Let X_1, X_2, $\dots X_n$ be i.i.d. random variables with expected value $E(X_i)=\mu<\infty$ and $0<V(X_i)=\sigma^2<\infty$. Then the random variable*

$$Z_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} \mathcal{N}(0,1) \quad (14.2.15)$$

a) From Lemma 14.1, $\{X_i^2\}$ is a sequence of i.i.d. random variables. Hence,

$$E(X_i^2) = \mathrm{var}(X_i) + [\mathrm{E}(X_i)]^2 = 1 \quad (14.2.16)$$

From Lemma 14.4,

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} 1 \neq 0 \tag{14.2.17}$$

Therefore, option 1 is incorrect.

b)

Lemma 14.9. *Let*

$$Y_n = \frac{1}{n^{3/4}}\sum_{i=1}^{n} X_i, \tag{14.2.18}$$

where X_i are i.i.d. with

$$E(X_i) = 0, var(X_i) = 1 \tag{14.2.19}$$

Then

$$E(Y_n) = 0, var(Y_n) = \frac{1}{n^{1/2}} \tag{14.2.20}$$

Proof.

$$\mathrm{E}(Y_n) = \frac{1}{n^{3/4}}\mathrm{E}\left(\left(\sum_{i=1}^{n} X_i\right)\right) = 0 \tag{14.2.21}$$

Since $E(X_i) = 0$. Also, from Lemma 14.6,

$$\mathrm{var}(Y_n) = \frac{1}{n^{3/2}}\left(\sum_{i=1}^{n} \mathrm{E}(X_i)\right) \tag{14.2.22}$$

$$= \frac{1}{n^{3/2}} \times n = \frac{1}{n^{1/2}} \tag{14.2.23}$$

$\because \mathrm{var}(X_i) = 1$. □

Now, from Lemma 14.7 and Lemma 14.9

$$\lim_{n\to\infty} \Pr(|Y_n - 0| \geq \epsilon) \leq \lim_{n\to\infty} \frac{1}{n^{1/2}\epsilon^2} \tag{14.2.24}$$

$$= 0 \tag{14.2.25}$$

From Definition 16,

$$\frac{1}{n^{3/4}}\sum_{i=1}^{n} X_i \xrightarrow{p} 0 \tag{14.2.26}$$

Thus, option 2 is correct.

c) Substituting $\mu = 0$ and $\sigma = 1$, in Lemma 14.8,

$$Z_n = \frac{1}{n^{1/2}}\sum_{i=1}^{n} X_i \xrightarrow{d} \mathcal{N}(0,1) \neq 0 \tag{14.2.27}$$

Therefore, option 3 is incorrect.

d) From (14.2.17), option 4 is correct.

Therefore, options 2 and 4 are correct.

14.3. Let $\{X_n\}$ be a sequence of independent random variables where the distribution of X_n is normal with mean μ and variance n for n = 1,2,... Define

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{14.3.1}$$

$$S_n = \frac{\sum_{i=1}^{n} \frac{1}{i} X_i}{\sum_{i=1}^{n} \frac{1}{i}} \tag{14.3.2}$$

Which of the following are true?

a) $\mathrm{E}(\bar{X}_n) = \mathrm{E}(S_n)$ for sufficiently large n
b) $\mathrm{Var}(S_n) < \mathrm{Var}(\bar{X}_n)$ for sufficiently large n
c) $\bar{X}_n$ is consistent for μ
d) $\bar{X}_n$ is sufficient for μ

Solution: As X_i for $i = 1, 2, \ldots, n$ are independent random variables we can use this property to state

$$Var\left(\sum_{i=1}^{n} g(X_i)\right) = \sum_{i=1}^{n} Var(g(X_i)) \tag{14.3.3}$$

Definition 17. *Random Sample: The random variables $X_1, X_2, X_3, \ldots, X_n$ are said to be random sample if*

a) the X_i's are independent
b) $F_{X_1}(x) = F_{X_2}(x) = \cdots = F_{X_n}(x) = F_X(x)$
c) $EX_i = EX = \mu < \infty$
d) $0 < \mathrm{Var}(X_i) = \mathrm{Var}(X) = \sigma^2 < \infty$

Let $n = 2$ and hence X_1 and X_2 are sequence of independent random variables and

$$Var(X_1) = 1 \tag{14.3.4}$$

$$Var(X_2) = 2 \tag{14.3.5}$$

$$Var(X_1) \neq Var(X_2) \tag{14.3.6}$$

The equation (14.3.6) doesn't follow point(14.3d) in definition(17) and hence the random variables are not a random sample.

a) Expectation of $\bar{X}_n$ and S_n

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \tag{14.3.7}$$
$$= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) \tag{14.3.8}$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mu \tag{14.3.9}$$
$$= \mu \tag{14.3.10}$$
$$E(S_n) = E\left(\frac{\sum_{i=1}^{n}\frac{1}{i}X_i}{\sum_{i=1}^{n}\frac{1}{i}}\right) \tag{14.3.11}$$
$$= \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{i}\right)}\sum_{i=1}^{n} E\left(\frac{1}{i}X_i\right) \tag{14.3.12}$$
$$= \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{i}\right)}\sum_{i=1}^{n}\frac{\mu}{i} \tag{14.3.13}$$
$$= \mu \tag{14.3.14}$$

From (14.3.10) and (14.3.14) we get option(14.3a) is correct.

b) Variance of $\bar{X}_n$ and S_n using (14.3.3)

$$Var(\bar{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \tag{14.3.15}$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var(X_i)\right) \tag{14.3.16}$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} i\right) \tag{14.3.17}$$
$$= \frac{1}{2} + \frac{1}{2n} \tag{14.3.18}$$
$$Var(S_n) = Var\left(\frac{\sum_{i=1}^{n}\frac{1}{i}X_i}{\sum_{i=1}^{n}\frac{1}{i}}\right) \tag{14.3.19}$$
$$= \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{i}\right)^2}\sum_{i=1}^{n}\frac{1}{i^2}Var\left(X_i\right) \tag{14.3.20}$$
$$= \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{i}\right)^2}\sum_{i=1}^{n}\frac{1}{i^2}i \tag{14.3.21}$$
$$= \frac{1}{\sum_{i=1}^{n}\frac{1}{i}} \tag{14.3.22}$$

As n is sufficiently large

$$Var(\bar{X}_n) = \frac{1}{2} \tag{14.3.23}$$
$$Var(S_n) = 0 \tag{14.3.24}$$
$$Var(S_n) < Var(\bar{X}_n) \tag{14.3.25}$$

from (14.3.25) we get option(14.3b) as correct.

c)

Definition 18. *Point Estimator : Let θ be an unknown fixed(non-random) parameter be estimated. To estimate θ we define a point estimator $\hat{\Theta}$ that is a function of the random sample $X_1, X_2, X_3, \ldots, X_n$ i.e.,*

$$\hat{\Theta} = h(X_1, X_2, \cdots, X_n) \tag{14.3.26}$$

Definition 19. *Consistent Estimator : Let $\hat{\Theta}_1$,$\hat{\Theta}_2$ $\ldots$,$\hat{\Theta}_n$,$\ldots$ be a sequence of point estimators of θ. We say that $\hat{\Theta}_n$ is a consistent estimator of θ , if*

$$\lim_{n\to\infty} P(|\hat{\Theta}_n - \theta| \geq \epsilon) = 0, \text{ for all } \epsilon > 0. \tag{14.3.27}$$

From (14.3.6) as given data is not a random sample we don't define point estimator and hence option(14.3c) is incorrect.

d)

Definition 20. *Statistic : A statistic is a function $T = r(X_1, X_2, \ldots, X_n)$ of the random sample $X_1, X_2, \ldots, X_n$.*

Definition 21. *Sufficient Statistics : A statistic t = T(X) is sufficient for θ if the conditional probability distribution of data X, given the statistic t = T(X), doesn't depend on the parameter θ.*

Equation (14.3.6) suggests that given data is not a random sample we don't define statistic and hence option(14.3d) is incorrect.

Hence option(14.3a) and option(14.3b) are correct.

14.4. Let $\{X_n : n > 0\}$ and X be random variables defined on a common probability space. Further assume that $\{X_n \geq 0 \ \forall n > 0\}$ and

$$\Pr\left(X = x\right) = \begin{cases} p, & x = 0 \\ 1-p, & x = 1 \\ 0, & \text{otherwise} \end{cases} \tag{14.4.1}$$

where, $0 \leq p \leq 1$. Which of the following statements are necessarily true?

a) If $p = 0$ and $X_n \xrightarrow{\text{d}} X$, $X_n \xrightarrow{\text{p}} X$
b) If $p = 1$ and $X_n \xrightarrow{\text{d}} X$, then $X_n \xrightarrow{\text{p}} X$
c) If $0 < p < 1$ and $X_n \xrightarrow{\text{d}} X$, then $X_n \xrightarrow{\text{p}} X$
d) If $X_n \xrightarrow{\text{p}} X$, then $X_n \xrightarrow{\text{a.s}} X$

Solution:

Definition 22 (Convergence in Distribution or Weak convergence)**.** *For any given sequence of real-valued random variables $X_1, X_2, X_3, \ldots, X_n$ and a random variable X,*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{14.4.2}$$

where F_{X_n} and F_X are the cumulative probability distribution functions of X_n and X respectively.

Definition 23 (Convergence in Probability)**.**

$$\lim_{n\to\infty} \Pr\left(|X_n - X| > \epsilon\right) = 0 \forall \epsilon > 0 \tag{14.4.3}$$

This is stronger than the convergence in distribution but weaker than Almost sure convergence.

Definition 24 (Almost sure Convergence)**.**

$$\Pr\left(\lim_{n\to\infty} X_n = X\right) = 1 \tag{14.4.4}$$

This type of convergence is stronger than both convergence in distribution and probability.

- In general, stronger statements imply weaker statements but not vice versa, i.e. Convergence in probability implies convergence in distribution and Almost sure convergence implies convergence in probability.
- We shall use the following statement from Portmanteau's Lemma in the following proof:

Lemma 14.10 (Portmanteau's Lemma)**.** *The sequence $X_1, X_2, X_3, \ldots, X_n$ converges in distribution to X if and only if*

$$\limsup \Pr\left(X_n \in F\right) \leq \Pr\left(X \in F\right) \tag{14.4.5}$$

for every closed set F;

Lemma 14.11. *Convergence in distribution implies convergence in probability if X is a constant.*

Proof:

- Let $\epsilon > 0$. Let X = c and the sequence $X_1, X_2, X_3, \ldots, X_n$ converges to X in distribution.
- Let

$$S = \{X : |X - c| > \epsilon\} \tag{14.4.6}$$

$$\implies \Pr\left(|X_n - c| > \epsilon\right) = \Pr\left(X_n \in S\right) \tag{14.4.7}$$

- From Lemma 14.10,

$$\because \lim_{n\to\infty} X_n = c \tag{14.4.8}$$

$$\implies \limsup_{n\to\infty} \Pr\left(X_n \in S\right) \leq \Pr\left(c \in S\right) \tag{14.4.9}$$

$$\because \Pr\left(c \in S\right) = 0 \text{ (By defn)} \tag{14.4.10}$$

$$\implies \limsup_{n\to\infty} \Pr\left(X_n \in S\right) \leq 0 \tag{14.4.11}$$

$$\implies \lim_{n\to\infty} \Pr\left(X_n \in S\right) \leq 0 \tag{14.4.12}$$

$$\implies \lim_{n\to\infty} \Pr\left(X_n \in S\right) = 0 \text{ (Probability} \geq 0) \tag{14.4.13}$$

- Thus, by definition,

$$\begin{gathered} \Pr\left(|X_n - c| > \epsilon\right) = 0 \text{ for any } \epsilon > 0 \text{ given,} \\ \lim_{n\to\infty} F_{X_n}(x) = F_X(x) \text{ and X is constant} \end{gathered} \tag{14.4.14}$$

Let us look at each option one after another.

a) Given,

$$p = 0 \implies X = 1$$

Since X is a constant, from Lemma **??**, we can say that option 1 is true.

b) Given,

$$p = 1 \implies X = 0$$

Since X is a constant, from Lemma **??**, we can say that option 2 is true.

c) Given,

$$0 < p < 1 \implies X \neq 0, 1$$

Since X is not a constant, we can say that option 3 is false.

d) Since Convergence in probability is weaker than Almost sure convergence, we can say that option 4 is false as a weaker statement does not imply a stronger statement.

Therefore, the true statements from the options are options 1 and 2.

14.5. Let $X_1, X_2, \dots$ be i.i.d. $N(0,1)$ random variables. Let

$$S_n = X_1^2 + X_2^2 + \cdots + X_n^2 \forall n \geq 1. \quad (14.5.1)$$

Which of the following statements are correct?

a)

$$\frac{S_n - n}{\sqrt{2}} \sim N(0,1) \quad \forall n \geq 1 \quad (14.5.2)$$

b)

$$\forall \epsilon > 0, \Pr\left(\left|\frac{S_n}{n} - 2\right| > \epsilon\right) \to 0, n \to \infty \quad (14.5.3)$$

c) $\frac{S_n}{n} \to 1$ with probability 1

d)

$$\Pr\left(S_n \leq n + \sqrt{n}x\right) \to \Pr\left(Y \leq x\right) \quad \forall x \in \mathbb{R}, Y \sim N(0,2) \quad (14.5.4)$$

Solution:

Lemma 14.12. *For independent random variables $X_i, i = 1, 2, \dots, n$,*

$$Var\left(\sum_{i=1}^{n} g(X_i)\right) = \sum_{i=1}^{n} Var(g(X_i)) \quad (14.5.5)$$

Definition 25. *Random Sample: The set $\{X_i\}_{i=1}^{n}$ constitutes a random sample if*

a) X_i are independent
b) $F_{X_i}(x) = F_X(x)$
c) $E(X_i) = E(X) = \mu < \infty$
d) $0 < var(X_i) = var(X) = \sigma^2 < \infty$

a) Expectation of $\bar{X}_n$ and S_n

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \quad (14.5.6)$$
$$= \frac{1}{n}\sum_{i=1}^{n} E\left(X_i\right) \quad (14.5.7)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mu \quad (14.5.8)$$
$$= \mu \quad (14.5.9)$$
$$E(S_n) = E\left(\frac{\sum_{i=1}^{n} \frac{1}{i} X_i}{\sum_{i=1}^{n} \frac{1}{i}}\right) \quad (14.5.10)$$
$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{i}\right)} \sum_{i=1}^{n} E\left(\frac{1}{i} X_i\right) \quad (14.5.11)$$
$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{i}\right)} \sum_{i=1}^{n} \frac{\mu}{i} \quad (14.5.12)$$
$$= \mu \quad (14.5.13)$$

From (14.5.9) and (14.5.13) we get option(**??**) is correct.

b) Variance of $\bar{X}_n$ and S_n using (14.5.5)

$$Var(\bar{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \quad (14.5.14)$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var(X_i)\right) \quad (14.5.15)$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} i\right) \quad (14.5.16)$$
$$= \frac{1}{2} + \frac{1}{2n} \quad (14.5.17)$$
$$Var(S_n) = Var\left(\frac{\sum_{i=1}^{n} \frac{1}{i} X_i}{\sum_{i=1}^{n} \frac{1}{i}}\right) \quad (14.5.18)$$
$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{i}\right)^2} \sum_{i=1}^{n} \frac{1}{i^2} Var\left(X_i\right) \quad (14.5.19)$$
$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{i}\right)^2} \sum_{i=1}^{n} \frac{1}{i^2} i \quad (14.5.20)$$
$$= \frac{1}{\sum_{i=1}^{n} \frac{1}{i}} \quad (14.5.21)$$

As n is sufficiently large

$$Var(\bar{X}_n) = \frac{1}{2} \quad (14.5.22)$$

$$Var(S_n) = 0 \quad (14.5.23)$$

$$Var(S_n) < Var(\bar{X}_n) \quad (14.5.24)$$

from (14.5.24) we get option(**??**) as correct.

c)

Definition 26. *Point Estimator : Let θ be an unknown fixed(non-random) parameter be estimated. To estimate θ we define a point estimator $\hat{\Theta}$ that is a function of the random sample $X_1, X_2, X_3, \ldots, X_n$ i.e.,*

$$\hat{\Theta} = h(X_1, X_2, \cdots, X_n) \quad (14.5.25)$$

Definition 27. *Consistent Estimator : Let $\hat{\Theta}_1, \hat{\Theta}_2 \ldots, \hat{\Theta}_n, \ldots$ be a sequence of point estimators of θ. We say that $\hat{\Theta}_n$ is a consistent estimator of θ , if*

$$\lim_{n\to\infty} P(|\hat{\Theta}_n - \theta| \geq \epsilon) = 0, \text{ for all } \epsilon > 0. \quad (14.5.26)$$

From the given sample $Var(X_i) = i$ for i = 1,2,... Hence the given random variables doesn't follow point(14.5d) in definition(25) and hence the random variables are not a random sample.
As given random variables are not random sample we don't define point estimator and hence option(**??**) is incorrect.

d)

Definition 28. *Statistic : A statistic is a function $T = r(X_1, X_2, \ldots, X_n)$ of the random sample $X_1, X_2, \ldots, X_n$.*

Definition 29. *Sufficient Statistics : A statistic $t = T(X)$ is sufficient for θ if the conditional probability distribution of data X, given the statistic $t = T(X)$, doesn't depend on the parameter θ.*

As given random variables are not random sample we don't define statistic and hence option(**??**) is incorrect.

Hence option(**??**) and option(**??**) are correct.

14.6. Let $X_1, X_2, ..., X_n$ be independent and identically distributed, each having a uniform distribution on $(0, 1)$. Let $S_n = \sum_{i=1}^{n} X_i$ for $n \geq 1$. Then, which of the following statements are true?

A) $\frac{S_n}{n \log n} \to 0$ as $n \to \infty$ with probability 1.
B) $\Pr\left(\left(S_n > \frac{2n}{3}\right) \text{occurs for infinitely many n}\right) = 1$
C) $\frac{S_n}{\log n} \to 0$ as $n \to \infty$ with probability 1.
D) $\Pr\left(\left(S_n > \frac{n}{3}\right) \text{occurs for infinitely many n}\right) = 1$

Solution:

Symbol	expression/definition
S_n	$\sum_{i=1}^{n} X_i$
μ_n	$\frac{1}{n}\sum_{i=1}^{n} X_i$
X	Independent continuous random variable identical to $X_1, X_2, ..., X_n$

TABLE 14.6.1: Variables and their definitions

a) Given

$$S_n = \sum_{i=1}^{n} X_i, n \geq 1 \quad (14.6.1)$$

Dividing by n on both sides

$$\frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \mu_n \quad (14.6.2)$$

It can be said that $X_1, X_2, ..., X_n$ are the trials of X. By definition

$$E[X] = \lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{n} = \lim_{n\to\infty} \frac{S_n}{n} \quad (14.6.3)$$

$$\lim_{n\to\infty} \frac{S_n}{n} = E[X] = \frac{1}{2} \quad (14.6.4)$$

$$\therefore \lim_{n\to\infty} \frac{S_n}{n \log n} = 0 \quad (14.6.5)$$

b) Using weak law, (14.6.4), and table (14.6.1)

$$\lim_{n\to\infty} \Pr\left(|\mu_n - E[X]| > \epsilon\right) = 0, \forall \epsilon > 0 \quad (14.6.6)$$

$$\lim_{n\to\infty} \Pr\left(S_n = \frac{n}{2}\right) = 1 \quad (14.6.7)$$

It can be easily implied from (14.6.7) that option B is false.

c) It is easy to observe from (14.6.4) that option C is false.

d) Using (14.6.7), we get

$$\Pr\left(\left(S_n > \frac{n}{3}\right)\text{occurs for infinitely many n}\right) = 1 \tag{14.6.8}$$

14.7. Let $U_1, U_2 \ldots, U_n$ be independent and identically distributed random variables each having a uniform distribution on (0,1). Then,$\lim_{n\to+\infty} \Pr\left(U_1 + U_2 \ldots, U_n \le \frac{3}{4}n\right)$

a) does not exist
b) exists and equals 0
c) exists and equals 1
d) exists and equals $\frac{3}{4}$

Solution: We use Weak law for large numbers to solve this problem. Let the collection of identically distributed random variables $U_1, U_2 \ldots, U_n$ have a finite mean μ and finite variance σ^2.

$$\mu = \mathrm{E}\left(U_i\right) \quad \text{for i} \in \text{(1,2,3...,n)} \tag{14.7.1}$$

Since the distribution is uniform on (0,1), $\mu = 0.5$. Let M_n be the sample mean

$$M_n = \frac{U_1 + U_2 + U_3 \cdots + U_n}{n} \tag{14.7.2}$$

Expected value of M_n (using (14.7.2) and (14.7.1))is

$$\mathrm{E}\left(M_n\right) = \frac{\mathrm{E}\left(U_1 + U_2 + U_3 + \cdots + U_n\right)}{\mathrm{E}\left(n\right)} \tag{14.7.3}$$

$$= \frac{\mathrm{E}\left(U_1\right) + \mathrm{E}\left(U_2\right) + \cdots + \mathrm{E}\left(U_n\right)}{n} \tag{14.7.4}$$

$$= \frac{n \times \mu}{n} \tag{14.7.5}$$

$$= \mu \tag{14.7.6}$$

Variance of M

$$Var(M_n) = \frac{Var(U_1 + U_2 + U_3 \cdots + U_n)}{n^2} \tag{14.7.7}$$

$$= \frac{Var(U_1) + Var(U_2) \cdots + Var(U_n)}{n^2} \tag{14.7.8}$$

$$= \frac{n \times \sigma^2}{n^2} \tag{14.7.9}$$

$$= \frac{\sigma^2}{n} \tag{14.7.10}$$

From Chebyshev inequality, for any $\epsilon > 0$

$$\Pr\left(|M_n - \mu| \ge \epsilon\right) \le \frac{Var(M_n)}{\epsilon^2} \tag{14.7.11}$$

From (14.7.1) and (14.7.10)

$$\Pr\left(\left|\frac{U_1 + U_2 \cdots + U_n}{n} - \mu\right| \ge \epsilon\right) \le \frac{\sigma^2}{n \times \epsilon^2}$$

$$\lim_{n\to\infty} \Pr\left(\left|\frac{U_1 + U_2 \cdots + U_n}{n} - \mu\right| \ge \epsilon\right) \le \lim_{n\to\infty} \frac{\sigma^2}{n \times \epsilon^2} \le 0 \text{ for fixed } \epsilon > 0 \tag{14.7.12}$$

But since Probabilities are always non-negative,

$$\lim_{n\to\infty} \Pr\left(\left|\frac{U_1 + U_2 \cdots + U_n}{n} - \mu\right| \ge \epsilon\right) \to 0 \tag{14.7.13}$$

This is known as the weak law of large numbers
The inverse of (14.7.13) is also true

$$\lim_{n\to\infty} \Pr\left(\left|\frac{U_1 + U_2 \cdots + U_n}{n} - \mu\right| \le \epsilon\right) \to 1 \tag{14.7.14}$$

$$\left|\frac{U_1 + U_2 \cdots + U_n}{n} - \mu\right| \le \epsilon \text{ as n} \to \infty \tag{14.7.15}$$

From ϵ, n definition of limits, it is clear that

$$\frac{U_1 + U_2 \cdots + U_n}{n} \to \mu \tag{14.7.16}$$

$$U_1 + U_2 \ldots U_n \to n \times \mu \text{ as n} \to \infty \tag{14.7.17}$$

Since $\mu = \frac{1}{2}$,

$$\lim_{n\to+\infty} U_1 + U_2 \dots U_n = \frac{1}{2}n < \frac{3}{4}n \quad (14.7.18)$$

So

$$\lim_{n\to+\infty} \Pr\left(U_1 + U_2 \dots, U_n \le \frac{3}{4}n\right) = 1 \quad (14.7.19)$$

14.8. Let $X_1, X_2, \dots$ be independent and identically distributed random variables each following a uniform distribution on (0,1). Denote $T_n = max\{X_1, X_2, \dots, X_n\}$. Then, which of the following statements are true?

a) T_n converges to 1 in probability.
b) $n(1 - T_n)$ converges in distribution.
c) $n^2(1 - T_n)$ converges in distribution.
d) $\sqrt{n}(1 - T_n)$ converges to 0 in probability.

Solution: The PDF, CDF of each $X_1, X_2, X_3, \dots$ is

$$f_{X_i}(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & otherwise \end{cases} \quad (72.1)$$

$$F_{X_i}(x) = \begin{cases} x, & 0 < x < 1 \\ 1, & x \ge 1 \\ 0, & otherwise \end{cases} \quad (72.2)$$

$\forall i \in \mathbb{N}$. Then, as $T_n = max\{X_1, X_2, \dots, X_n\}$,

$$f_{T_n}(x) = \begin{cases} nx^{n-1}, & 0 < x < 1 \\ 0, & otherwise \end{cases} \quad (72.3)$$

$$F_{T_n}(x) = \begin{cases} x^n, & 0 < x < 1 \\ 1, & x \ge 1 \\ 0, & otherwise \end{cases} \quad (72.4)$$

NOTE : If $Y = aX + b$ and $a < 0$, then

$$F_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right) \quad (72.5)$$

a) OPTION-1:
Convergence in Probability :
A sequence of random variables $X_1, X_2, X_3, \dots$ converges in probability to a random variable X, shown by $X_n \xrightarrow{p} X$, if

$$\lim_{n\to\infty} \Pr(|X_n - X| \ge \epsilon) = 0, \forall \epsilon > 0 \quad (72.6)$$

To evaluate : $\lim_{n\to\infty} \Pr(|T_n - 1| \ge \epsilon), \forall \epsilon > 0$

$$\lim_{n\to\infty} \Pr(|T_n - 1| \ge \epsilon) = \lim_{n\to\infty} \Pr(1 - T_n \ge \epsilon) \quad (72.7)$$

$$= \lim_{n\to\infty} \Pr(T_n \le 1 - \epsilon) = \lim_{n\to\infty} F_{T_n}(1 - \epsilon) \quad (72.8)$$

$$F_{T_n}(1 - \epsilon) = \begin{cases} (1-\epsilon)^n, & 0 < \epsilon < 1 \\ 0, & \epsilon \ge 1 \end{cases} \quad (72.9)$$

$$\because \lim_{n\to\infty} (1-\epsilon)^n = 0 \text{ for } 0 < \epsilon < 1 \quad (72.10)$$

$$\therefore \lim_{n\to\infty} \Pr(|T_n - 1| \ge \epsilon) = 0, \forall \epsilon > 0 \quad (72.11)$$

$\therefore T_n$ converges to 1 in probability.

b) OPTION-2:
Convergence in Distribution :
A sequence of random variables $X_1, X_2, X_3, \dots$ converges in distribution to a random variable X, shown by $X_n \xrightarrow{d} X$, if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad (72.12)$$

for all x at which $F_X(x)$ is continuous.
To evaluate : $\lim_{n\to\infty} F_{n(1-T_n)}(x)$
Substituting $a = -n, b = n$ in (72.5),

$$F_{n(1-T_n)}(x) = 1 - F_{T_n}\left(1 - \frac{x}{n}\right) \quad (72.13)$$

$$F_{T_n}\left(1 - \frac{x}{n}\right) = \begin{cases} \left(1 - \frac{x}{n}\right)^n, & 0 < x < n \\ 1, & x \le 0 \\ 0, & x \ge n \end{cases} \quad (72.14)$$

$$\because \lim_{n\to\infty} \left(1 - \frac{y}{n}\right)^n = e^{-y} \quad (72.15)$$

$$\therefore \lim_{n\to\infty} F_{T_n}\left(1 - \frac{x}{n}\right) = \begin{cases} e^{-x}, & 0 < x < n \\ 1, & x \le 0 \\ 0, & x \ge n \end{cases} \quad (72.16)$$

$$\therefore F_{n(1-T_n)}(x) = \begin{cases} 1 - e^{-x}, & 0 < x < n \\ 0, & x \le 0 \\ 1, & x \ge n \end{cases} \tag{72.17}$$

$\therefore$ $n(1 - T_n)$ converges in distribution to a random variable with CDF in (72.17).

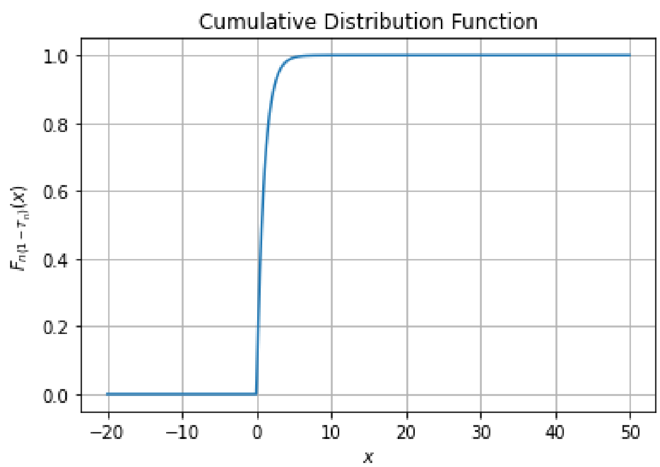


Fig. 14.8.1: CDF

c) OPTION-3:
Convergence in Distribution :
A sequence of random variables $X_1, X_2, X_3, \ldots$ converges in distribution to a random variable X, shown by $X_n \xrightarrow{d} X$, if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{72.18}$$

for all x at which $F_X(x)$ is continuous.
To evaluate : $\lim_{n\to\infty} F_{n^2(1-T_n)}(x)$
Substituting $a = -n^2, b = n^2$ in (72.5),

$$F_{n^2(1-T_n)}(x) = 1 - F_{T_n}\left(1 - \frac{x}{n^2}\right) \tag{72.19}$$

$$F_{T_n}\left(1 - \frac{x}{n^2}\right) = \begin{cases} \left(1 - \frac{x}{n^2}\right)^n, & 0 < x < n^2 \\ 1, & x \le 0 \\ 0, & x \ge n^2 \end{cases} \tag{72.20}$$

$$\because \lim_{n\to\infty} \left(1 - \frac{y}{n^2}\right)^n \text{ is not defined} \tag{72.21}$$

$\therefore$ $n^2(1 - T_n)$ does not converge in distribution.

d) OPTION-4:
Convergence in Probability :
A sequence of random variables $X_1, X_2, X_3, \ldots$ converges in probability to a random variable X, shown by $X_n \xrightarrow{p} X$, if

$$\lim_{n\to\infty} \Pr\left(|X_n - X| \ge \epsilon\right) = 0, \forall \epsilon > 0 \tag{72.22}$$

To evaluate :
$\lim_{n\to\infty} \Pr\left(|\sqrt{n}(1 - T_n) - 0| \ge \epsilon\right), \forall \epsilon > 0$

$$= \lim_{n\to\infty} \Pr\left(1 - T_n \ge \frac{\epsilon}{\sqrt{n}}\right) \tag{72.23}$$

$$= \lim_{n\to\infty} \Pr\left(T_n \le 1 - \frac{\epsilon}{\sqrt{n}}\right) \tag{72.24}$$

$$= \lim_{n\to\infty} F_{T_n}\left(1 - \frac{\epsilon}{\sqrt{n}}\right) \tag{72.25}$$

$$F_{T_n}\left(1 - \frac{\epsilon}{\sqrt{n}}\right) = \begin{cases} \left(1 - \frac{\epsilon}{\sqrt{n}}\right)^n, & 0 < \epsilon < \sqrt{n} \\ 0, & \epsilon \ge \sqrt{n} \end{cases} \tag{72.26}$$

$$\because \lim_{n\to\infty} \left(1 - \frac{\epsilon}{\sqrt{n}}\right)^n = 0 \text{ for } 0 < \epsilon < \sqrt{n} \tag{72.27}$$

$$\therefore \lim_{n\to\infty} \Pr\left(|\sqrt{n}(1 - T_n) - 0| \ge \epsilon\right) = 0, \forall \epsilon > 0 \tag{72.28}$$

$\therefore$ $\sqrt{n}(1 - T_n)$ converges to 0 in probability.

Hence, options 1), 2), 4) are correct.

14.9. Let $X_1, X_2, ...$ be i.i.d N(1,1) random variables.Let $S_n = X_1^2 + X_2^2 + ... + X_n^2$ for $n \ge 1$.Then

$$\lim_{n\to\infty} \frac{Var(S_n)}{n} =$$

(A) 4
(B) 6
(C) 1
(D) 0

15 Statistics

15.1. Let X_1 and X_2 be a random sample of size two from a distribution with probability density function

$$f_\theta(x) = \theta\left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}x^2} + (1-\theta)\left(\frac{1}{2}\right)e^{-|x|},$$

$-\infty < x < \infty$,

where $\theta \in \left\{0, \frac{1}{2}, 1\right\}$. If the observed values of X_1 and X_2 are 0 and 2, respectively, then the maximum likelihood estimate of θ is

a) 0
b) $\frac{1}{2}$
c) 1
d) not unique

Solution:

15.2. Let Y_1 denote the first order statistic in a random sample of size n from a distribution that has the pdf

$$f(x) = \begin{cases} e^{-(x-\theta)} & \text{when } \theta < x < \infty \\ 0 & \text{otherwise} \end{cases} \tag{15.2.1}$$

Obtain the distribution of $Z_n = n(Y_1 - \theta)$.

Solution: From the given information

$$Y_1 = \min\{X_1, X_2, ...X_n\} \tag{15.2.2}$$

and

$$F_{Z_n}(z) = \Pr\left(n(Y_1 - \theta) \le z\right) \tag{15.2.3}$$

$$= \Pr\left(Y_1 \le \frac{z}{n} + \theta\right) \tag{15.2.4}$$

$$= 1 - \Pr\left(Y_1 > \frac{z}{n} + \theta\right) \tag{15.2.5}$$

Let

$$\left(\frac{z}{n} + \theta\right) = z' \tag{15.2.6}$$

Then

$$F_{Z_n}(z) = 1 - \prod_{i=1}^{n} \Pr\left(X_i > z'\right) \tag{15.2.7}$$

$$= 1 - (1 - F(z'))^n \tag{15.2.8}$$

$$\implies F_{Z_n}(z) = 1 - \left(1 - F\left(\frac{z}{n} + \theta\right)\right)^n \tag{15.2.9}$$

where

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \tag{15.2.10}$$

$$= \begin{cases} 1 - e^{-(x-\theta)} & \text{when } \theta < x < \infty \\ 0 & \text{otherwise} \end{cases} \tag{15.2.11}$$

Substituting from (15.2.11) in (15.2.9),

$$F_{Z_n}(z) = \begin{cases} 1 - e^{-n\left(\frac{z}{n}+\theta-\theta\right)} & \theta < \frac{z}{n} + \theta < \infty \\ 0 & \text{otherwise} \end{cases} \tag{15.2.12}$$

$$= \begin{cases} 1 - e^{-z} & \text{when } 0 < z < \infty \\ 0 & \text{otherwise} \end{cases} \tag{15.2.13}$$

and

$$f_{Z_n}(z) = \frac{d}{dz}F_{Z_n}(z) \tag{15.2.14}$$

$$= \begin{cases} e^{-z} & 0 < z < \infty \\ 0 & \text{otherwise} \end{cases} \tag{15.2.15}$$

The plots for the cdf in (15.2.13) and the pdf in (15.2.15) are shown in Fig. 15.2.1 and Fig. 15.2.2 respectively:

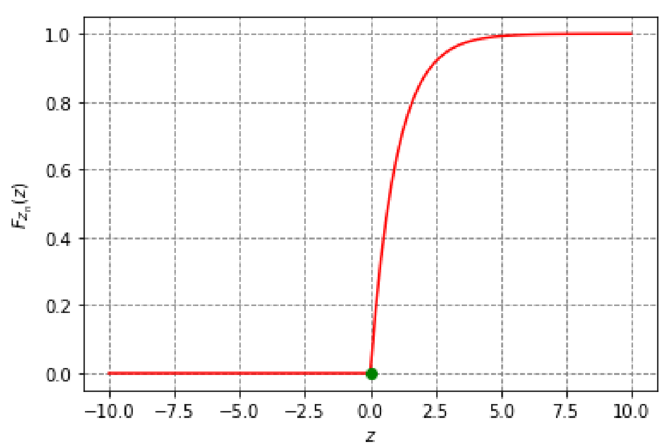


Fig. 15.2.1: cdf of Z_n

15.3. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n (≥ 2) from a distribution having the probability density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) & x > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{15.3.1}$$

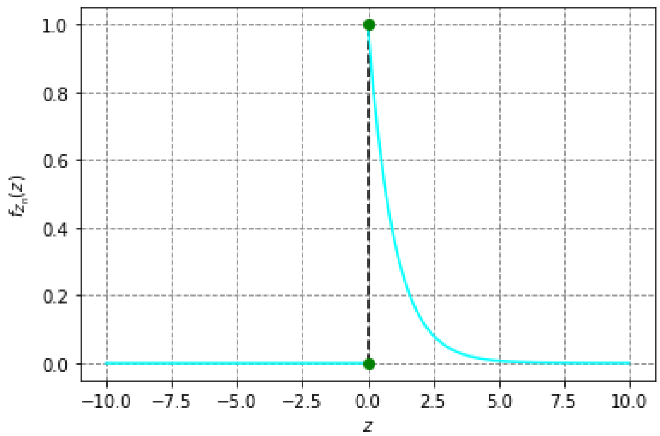


Fig. 15.2.2: pdf of Z_n

where $\theta \in (0, \infty)$. Let $X_{(1)} = \min\{X_1, X_2, \dots, X_n\}$ and $T = \sum_{i=1}^{n} X_i$. Then $E(X_{(1)}|T)$ equals

(A) $\frac{T}{n^2}$

(B) $\frac{T}{n}$

(C) $\frac{(n+1)T}{2n}$

(D) $\frac{(n+1)^2 T}{4n^2}$

Solution:

Lemma 15.1. Lehmann–Scheffé theorem : *If T is a complete sufficient statistic for θ and*

$$E(g(T)) = \tau(\theta) \tag{15.3.2}$$

then $g(T)$ is the uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(\theta)$.

We know that

$$T = \sum_{i=1}^{n} X_i \tag{15.3.3}$$

is a complete and sufficient statistic. By the law of total expectation,

$$E\left(E(X_{(1)}|T)\right) = E(X_{(1)}) \tag{15.3.4}$$

By Lehmann–Scheffé theorem, with

$$\theta = X_{(1)}, \tag{15.3.5}$$

$$\tau(x) = E(x), \tag{15.3.6}$$

$$g(T) = E(X_{(1)}|T). \tag{15.3.7}$$

it follows from (15.3.4) that $E(X_{(1)}|T)$ is the UMVUE of $E(X_{(1)})$.

$$\Pr(X_{(1)} > x) = \Pr(X_1 > x) \dots \Pr(X_n > x) \tag{15.3.8}$$

$$= (1 - F_{X_1}(x)) \dots (1 - F_{X_n}(x)) \tag{15.3.9}$$

$$= (1 - F_{X_1}(x))^n \tag{15.3.10}$$

$$= \exp\left(-\frac{nx}{\theta}\right) \tag{15.3.11}$$

$$F_{X_{(1)}}(x) = 1 - \exp\left(-\frac{nx}{\theta}\right) \tag{15.3.12}$$

$$f_{X_{(1)}}(x) = \frac{n}{\theta} \exp\left(-\frac{nx}{\theta}\right) \tag{15.3.13}$$

Therefore, $X_{(1)}$ follows an exponential distribution with mean $\frac{\theta}{n}$.

$$E(X_{(1)}) = \frac{\theta}{n} \tag{15.3.14}$$

Note that,

$$E\left(\frac{T}{n^2}\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n^2}\right) \tag{15.3.15}$$

$$= \frac{E(\sum_{i=1}^{n} X_i)}{n^2} \tag{15.3.16}$$

$$= \sum_{i=1}^{n} \frac{E(X_i)}{n^2} \tag{15.3.17}$$

$$= \sum_{i=1}^{n} \frac{\theta}{n^2} \tag{15.3.18}$$

$$= \frac{\theta}{n} \tag{15.3.19}$$

$$= E(X_{(1)}) \tag{15.3.20}$$

Therefore, by Lehmann–Scheffé theorem, with

$$\theta = X_{(1)}, \tag{15.3.21}$$

$$\tau(x) = E(x), \tag{15.3.22}$$

$$g(T) = \frac{T}{n^2}, \tag{15.3.23}$$

it follows that $\frac{T}{n^2}$ is UMVUE of $E(X_{(1)})$.

Since there exists a unique UMVUE for $E(X_{(1)})$, it follows that

$$E(X_{(1)}|T) = \frac{T}{n^2} \tag{15.3.24}$$

Hence, option A is correct.

15.4. Let X, Y, have the joint discrete distribution

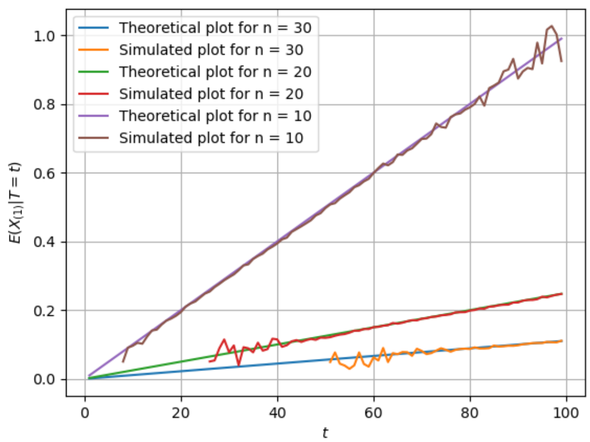


Fig. 15.3.1: Theory vs Simulated plot of $E(X_{(1)}|T)$

such that $X|Y = y \sim$ Binomial$(y, 0.5)$ and $Y \sim$ Poisson(λ), $\lambda > 0$, where λ is an unknown parameter. Let $T = T(X, Y)$ be any unbiased estimator of λ. Then

a) $Var(T) \leq Var(Y)$ for all λ
b) $Var(T) \geq Var(Y)$ for all λ
c) $Var(T) \geq \lambda$ for all λ
d) $Var(T) = Var(Y)$ for all λ

Solution:

Definition 30. Unbiased Estimator: *Suppose we have an estimator $T(X, Y)$; which estimates a parameter θ. If*

$$E\left(T(X, Y)\right) = \theta \tag{15.4.1}$$

the estimator is unbiased.

Lemma 15.2. *Since Y has a Poisson distribution, we know that:*

$$Var(Y) = \lambda \tag{15.4.2}$$

From Bayes Theorem,

$$\Pr\left(X = x, Y = y\right) = \Pr\left(Y = y\right)\Pr\left(X = x|Y = y\right) \tag{15.4.3}$$

$$= {}^{y}C_{x}\frac{1}{2^{y}}\frac{\lambda^{y}}{y!}e^{-\lambda} \tag{15.4.4}$$

Let us represent $\Pr\left(X = x, Y = y\right)$ as a function of x, y, and λ: $f(x, y; \lambda)$.

Definition 31. *If $T(X, Y)$ is an unbiased estimator of a parameter λ, the Cramer-Rao bound states that:*

$$Var\left(T\left(X, Y\right)\right) \geq -\frac{1}{E\left(\frac{\partial^2 \ln(f(x,y;\lambda))}{\partial\lambda^2}\right)} \tag{15.4.5}$$

From using (15.4.5) on $T(X, Y)$,

$$Var(T(X, Y)) \geq -\frac{1}{E\left(\frac{\partial^2 \ln(f(x,y;\lambda))}{\partial\lambda^2}\right)} \tag{15.4.6}$$

$$\geq -\frac{1}{E\left(-\frac{y}{\lambda^2}\right)} \tag{15.4.7}$$

$$\geq \frac{\lambda^2}{E\left(y\right)} \tag{15.4.8}$$

$$\geq \lambda \tag{15.4.9}$$

because the expectation value of a Poisson distribution with parameter λ is λ. The correct options are options (2) and (3), since by (15.4.9), we see that

$$Var(T) \geq \lambda = Var(Y)$$

(from (15.4.2)). (1) and (4) do not hold for all λ.

15.5. Let $X_1, \cdots, X_n$ be independent and identically distributed random variables with probability density function

$$f(x) = \frac{1}{2}\lambda^3 x^2 e^{-\lambda x}; x > 0; \lambda > 0 \tag{15.5.1}$$

Then which of the following statements are true?

a) $\frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i}$ is an unbiased estimator of λ
b) $\frac{3n}{\sum_{i=1}^{n}X_i}$ is an unbiased estimator of λ
c) $\frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i}$ is a consistent estimator of λ
d) $\frac{3n}{\sum_{i=1}^{n}X_i}$ is a consistent estimator of λ

Solution:

Definition 32. *An* ***estimator*** *is a statistic that estimates some fact about the population. The quantity that is being estimated is called the* ***estimand.***

Definition 33. *Let $\Theta = h(X_1, X_2, \cdots, X_n)$ be a point estimator for θ. The* ***bias*** *of the estimator Θ is defined by*

$$B(\Theta) = E\left(\Theta\right) - \theta \tag{15.5.2}$$

where $E\left(\Theta\right)$ is the expectation value of the *estimator Θ and θ is the estimand.*

Definition 34. *Let $\Theta = h(X_1, X_2, \cdots, X_n)$ be a point estimator for a parameter θ. We say that Θ is an* ***unbiased estimator*** *of θ if*

$$B(\Theta) = 0, \text{ for all possible values of } \theta. \tag{15.5.3}$$

Definition 35. *Let $\Theta_1, \Theta_2, \cdots, \Theta_n, \cdots$, be a sequence of point estimators of θ. We say that Θ_n is a* ***consistent*** *estimator of θ, if*

$$\lim_{n\to\infty} \Pr\left(|\Theta_n - \theta| \geq \epsilon\right) = 0 \text{ ,for all } \epsilon > 0. \tag{15.5.4}$$

Definition 36. *The* ***mean squared error (MSE)*** *of a point estimator Θ, shown by $MSE(\Theta)$, is defined as*

$$MSE(\Theta) = E\left((\Theta - \theta)^2\right) \tag{15.5.5}$$
$$= Var(\Theta) + B(\Theta)^2 \tag{15.5.6}$$

where $B(\Theta)$ is the bias of Θ.

Theorem 15.1. *Let $\Theta_1, \Theta_2, \cdots$ be a sequence of point estimators of θ. If*

$$\lim_{n\to\infty} MSE(\Theta_n) = 0, \tag{15.5.7}$$

then Θ_n is a consistent estimator of θ.

Definition 37. *The* ***moment generating function (MGF)*** *of a random variable X is a function $M_X(s)$ defined as*

$$M_X(s) = E[e^{sX}]. \tag{15.5.8}$$

Lemma 15.3. *The Gamma function is defined through the integral*

$$\int_0^\infty x^a e^{-bx}\,dx = \frac{\Gamma(a+1)}{b^{a+1}} \tag{15.5.9}$$

where $\Gamma(a) = (a-1)!$ for $a \in Z$

Theorem 15.2. *The Moment generating function of the gamma distribution with PDF*

$$f_X(x) = \begin{cases} \dfrac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & x > 0 \\ 0 & otherwise \end{cases} \tag{15.5.10}$$

where $\Gamma(\alpha) = \dfrac{1}{(\alpha-1)!}$ is given by,

$$M_X(s) = \left(\frac{\lambda}{\lambda - s}\right)^\alpha \tag{15.5.11}$$

Proof.

$$M_X(s) = E[e^{sX}] \tag{15.5.12}$$
$$= \int_{-\infty}^\infty e^{sx} f(x)\,dx \tag{15.5.13}$$
$$= \int_0^\infty e^{sx} \frac{1}{\Gamma(\alpha)} \lambda^\alpha x^{\alpha-1} e^{-\lambda x}\,dx \tag{15.5.14}$$
$$= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\lambda - s)x}\,dx \tag{15.5.15}$$

Substituting from Lemma 15.3 in the above,

$$M_X(s) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \times \frac{\Gamma(\alpha)}{(\lambda - s)^\alpha} \tag{15.5.16}$$
$$= \left(\frac{\lambda}{\lambda - s}\right)^\alpha \tag{15.5.17}$$

□

Theorem 15.3. *The MGF (if it exists) uniquely determines the distribution. That is, if two random variables have the same MGF, then they must have the same distribution.*

Lemma 15.4. *If X_i for $i = 1, 2, \ldots, n$ are i.i.d,*

$$var\left(a \sum_{i=1}^n g(X_i)\right) = a^2 \sum_{i=1}^n var\left(g(X_i)\right) \tag{15.5.18}$$

Lemma 15.5. *For X in* (15.5.1),

$$E\left[\frac{1}{\bar{X}}\right] = \frac{n\lambda}{3n-1} \tag{15.5.19}$$

Proof. Let

$$T = \sum_{i=1}^n X_i \sim \Gamma(3n, \lambda) \tag{15.5.20}$$

with pdf,

$$f_T(t) = \frac{\lambda^{3n} t^{3n-1} e^{-\lambda t}}{\Gamma(3n)}, t > 0 \tag{15.5.21}$$

$\because \dfrac{1}{\bar{X}} = \dfrac{n}{T}$,

$$E\left[\frac{1}{\bar{X}}\right] = \int_0^\infty \frac{n}{t} \frac{1}{\Gamma(3n)} \lambda^{3n} t^{3n-1} e^{-\lambda t}\,dt$$
$$= \frac{n\lambda}{(3n-1)} \int_0^\infty \frac{1}{\Gamma(3n-1)} \lambda^{3n-1} t^{3n-2} e^{-\lambda t}\,dt$$
$$\implies E\left[\frac{1}{\bar{X}}\right] = \frac{n\lambda}{3n-1} \tag{15.5.22}$$

□

Corollary 15.4.

$$E\left[\frac{1}{X_i}\right] = \frac{1\times\lambda}{3-1} \tag{15.5.23}$$
$$= \frac{\lambda}{2} \tag{15.5.24}$$

Lemma 15.6.

$$var\left(\frac{1}{\bar{X}}\right) = \frac{n^2\lambda^2}{(3n-1)^2(3n-2)} \tag{15.5.25}$$

Proof.

$$\text{var}\left(\frac{1}{\bar{X}}\right) = E\left[\frac{1}{\bar{X}^2}\right] - E\left[\frac{1}{\bar{X}}\right]^2 \tag{15.5.26}$$

$$E\left[\frac{1}{\bar{X}^2}\right] = \int_0^\infty \frac{n^2}{t^2}\frac{1}{\Gamma(3n)}\lambda^{3n}t^{3n-1}e^{-\lambda t}\,dt$$
$$= \frac{n^2\lambda^2}{(3n-1)(3n-2)} \tag{15.5.27}$$

$$\therefore \text{var}\left(\frac{1}{\bar{X}}\right) = \left(\frac{n^2\lambda^2}{(3n-1)(3n-2)} - \frac{n^2\lambda^2}{(3n-1)^2}\right) \tag{15.5.28}$$
$$= \frac{n^2\lambda^2}{3n-1}\left(\frac{1}{3n-2} - \frac{1}{3n-1}\right) \tag{15.5.29}$$
$$= \frac{n^2\lambda^2}{(3n-1)^2(3n-2)} \tag{15.5.30}$$

□

Corollary 15.5. *Substituting $n = 1$ in Lemma 15.6 as,*

$$var\left(\frac{1}{X_i}\right) = \frac{1^2\lambda^2}{(3-1)^2(3-2)} \tag{15.5.31}$$
$$= \frac{\lambda}{4} \tag{15.5.32}$$

a) In this case,

$$\Theta = \frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i} \text{ and } \theta = \lambda \tag{15.5.33}$$

Then

$$E[\Theta] = E\left[\frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right] \tag{15.5.34}$$
$$= \frac{2}{n}\sum_{i=1}^{n}E\left[\frac{1}{X_i}\right] \tag{15.5.35}$$

Using corollary 15.4,

$$E[\Theta] = \frac{2n}{n}\times\frac{\lambda}{2} \tag{15.5.36}$$
$$= \lambda \tag{15.5.37}$$

So the bias of estimator is given by,

$$B(\Theta) = E[\Theta] - \theta \tag{15.5.38}$$
$$= \lambda - \lambda = 0 \tag{15.5.39}$$

Therefore $\frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i}$ is an unbiased estimator of λ. Option 1 is correct.

b)

$$\Theta = \frac{3n}{\sum_{i=1}^{n}X_i} = \frac{3}{\bar{X}} \text{ and } \theta = \lambda \tag{15.5.40}$$

From Lemma 15.5 and equation (**??**),

$$E[\Theta] = E\left[\frac{3}{\bar{X}}\right] \tag{15.5.41}$$
$$= \frac{3n\lambda}{3n-1} \tag{15.5.42}$$

and

$$B(\Theta) = E[\Theta] - \lambda \tag{15.5.43}$$
$$= \frac{3n\lambda}{3n-1} - \lambda \tag{15.5.44}$$
$$= \frac{\lambda}{3n-1} \neq 0 \tag{15.5.45}$$

Therefore $\frac{3n}{\sum_{i=1}^{n}X_i}$ is not an unbiased estimator of λ. Option 2 is not correct.

c)

$$\Theta = \frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i} \text{ and } \theta = \lambda \tag{15.5.46}$$

Using Lemma **??**,

$$\text{var}(\Theta) = \text{var}\left(\frac{2}{n}\sum_{i=1}^{n}\frac{1}{X_i}\right) \tag{15.5.47}$$
$$= \frac{4}{n^2}\sum_{i=1}^{n}\text{var}\left(\frac{1}{X_i}\right) \tag{15.5.48}$$

Using Corollary 15.5,

$$\text{var}(\Theta) = \frac{4n}{n^2}\times\frac{\lambda^2}{4} \tag{15.5.49}$$
$$= \frac{\lambda^2}{n} \tag{15.5.50}$$

From option 1,

$$B(\Theta) = 0 \implies MSE(\Theta_n) = Var(\Theta) + B(\Theta)^2 \quad (15.5.51)$$

$$= \frac{\lambda^2}{n} \text{or, } \lim_{n\to\infty} MSE(\Theta_n) = \lim_{n\to\infty} \frac{\lambda^2}{n} \quad (15.5.52)$$

$$= 0 \quad (15.5.53)$$

$\therefore \frac{2}{n} \sum_{i=1}^{n} \frac{1}{X_i}$ is a consistent estimator of λ. Option 3 is correct.

d)

$$\Theta = \frac{3n}{\sum_{i=1}^{n} X_i} \text{ and } \theta = \lambda \quad (15.5.54)$$

The variance of estimator similar to option 2 and using Lemma 15.6,

$$\text{var}(\Theta) = \text{var}\left(\frac{3}{\bar{X}}\right) \quad (15.5.55)$$

$$= 9\text{var}\left(\frac{1}{\bar{X}}\right) \quad (15.5.56)$$

$$= \frac{9n^2\lambda^2}{(3n-1)^2(3n-2)} \quad (15.5.57)$$

From option 2,

$$B(\Theta) = \frac{\lambda}{3n-1} \quad (15.5.58)$$

$$\implies MSE(\Theta) = Var(\Theta) + B(\Theta)^2 \quad (15.5.59)$$

$$= \frac{9n^2\lambda^2}{(3n-1)^2(3n-2)} + \frac{\lambda^2}{(3n-1)^2} \quad (15.5.60)$$

Thus,

$$\lim_{n\to\infty} MSE(\Theta_n) \quad (15.5.61)$$

$$= \lim_{n\to\infty} \frac{9n^2\lambda^2}{(3n-1)^2(3n-2)} + \frac{\lambda^2}{(3n-1)^2} = 0 \quad (15.5.62)$$

$\therefore \frac{3n}{\sum_{i=1}^{n} X_i}$ is a consistent estimator of λ. Option 4 is correct.

Therefore option 1, option 3 and option 4 are correct.

15.6. If the marginal probability density function of the k^{th} order statistic of a random sample of size 8 from a uniform distribution on $[0, 2]$ is

$$f(x) = \begin{cases} \frac{7}{32} x^6 (2-x), & 0 < x < 2, \\ 0, & \text{otherwise,} \end{cases} \quad (15.6.1)$$

then k equals ______________

Solution:

Definition 38. *For a given statistical sample $\{X_1, X_2, \cdots X_n\}$, the order statistics is obtained by sorting the sample in ascending order. It denoted as $\{X_{(1)}, X_{(2)}, \cdots X_{(n)}\}$. The k^{th} smallest value $X_{(k)}$ is called k^{th} order statistic*

Theorem 15.6. *Let $\{X_1, X_2, \cdots X_n\}$ be n i.i.d random variables with common CDF = $F(x)$ and common PDF = $f(x)$, then the marginal probability distribution of k^{th} order statistic (CDF) is denoted by $F_{(k,n)}(x)$ and is given by*

$$F_{(k,n)}(x) = \sum_{j=k}^{n} {}^{n}C_j \,(F(x))^j \,(1 - F(x))^{n-j} \quad (15.6.2)$$

Proof. Let $P \sim B(n, F(x))$. Then

$$F_{(k,n)}(x) = \Pr(P \geq k) \quad (15.6.3)$$

$$= \sum_{j=k}^{n} {}^{n}C_j \,(F(x))^j \,(1 - F(x))^{n-j} \quad (15.6.4)$$

□

Corollary 15.7. *The marginal probability density of k^{th} order statistic (PDF) is denoted by $f_{(k,n)}(x)$ and given by*

$$f_{(k,n)}(x) = n \,{}^{n-1}C_{k-1} \,f(x)\,(F(x))^{k-1}\,(1 - F(x))^{n-k} \quad (15.6.5)$$

Definition 39 (Beta function)**.**

$$B(r, s) = \int_0^1 x^{r-1}\,(1-x)^{s-1}\,dx = \frac{\Gamma(r)\,\Gamma(s)}{\Gamma(r+s)} \quad (15.6.6)$$

Definition 40 (Beta Distribution)**.** *The Beta distribution is a continuous distribution defined on the range $(0, 1)$ whose PDF given by*

$$f(x) = \frac{1}{B(r, s)}\,x^{r-1}\,(1-x)^{s-1} \quad (15.6.7)$$

where $\int_0^1 f(x) = 1$ *as per definition* (39) *CDF, Mean value and Variance of Beta distribution*

$$F(x) = \frac{\int_0^x x^{r-1}\,(1-x)^{s-1}}{B(r,s)} = \frac{B_x(r,s)}{B(r,s)} \tag{15.6.8}$$

$$E(x) = \frac{r}{r+s} \tag{15.6.9}$$

$$var\,(()\,x) = \frac{rs}{(r+s)^2\,(r+s+1)} \tag{15.6.10}$$

Definition 41 (Uniform Order Statistics)**.** *Let* $\{X_1, \cdots X_n\}$ *be i.i.d form a uniform distribution on* $[0,1]$ *such that* $f_X(x) = 1$ *and* $F_X(x) = x$*. From Theorem* (15.7)*, equation* (15.6.5)

$$f_{(k,n)}(x) = n\,{}^{n-1}C_{k-1}\,x^{k-1}\,(1-x)^{n-k} \tag{15.6.11}$$

Lemma 15.7. *Uniform order statistics on [0,1] the PDF of k^{th} order statistic follows Beta distribution with $r = k$, $s = n-k+1$ and PDF is given by*

$$f_{(k,n)}(x) = \frac{1}{B(k,n-k+1)}\,x^{k-1}\,(1-x)^{(n-k+1)-1} \tag{15.6.12}$$

Proof. Since (15.6.11) is the PDF,

$$\int_0^1 n\,{}^{n-1}C_{k-1}\,x^{k-1}\,(1-x)^{n-k}\,dx = 1 \tag{15.6.13}$$

$$\int_0^1 x^{k-1}\,(1-x)^{n-k}\,dx = \frac{(k-1)!\,(n-k)!}{n!} \tag{15.6.14}$$

$$\int_0^1 x^{k-1}\,(1-x)^{n-k}\,dx = \frac{\Gamma(k)\,\Gamma(n-k+1)}{\Gamma\,(k+(n-k+1)))} \tag{15.6.15}$$

$$\int_0^1 x^{k-1}\,(1-x)^{n-k}\,dx = B(k,n-k+1) \tag{15.6.16}$$

$$\int_0^1 \frac{x^{k-1}\,(1-x)^{(n-k+1)-1}}{B(k,n-k+1)}\,dx = 1 \tag{15.6.17}$$

from definition (40) with $r = k$ and $s = n-k+1$ equation (15.6.11) follows beta distribution □

From lemma (15.7), PDF of k^{th} order statistic of a uniform distribution on $[0,1]$ follows beta distribution

$$\int_0^2 f_{(k,8)}(x)\,dx = \int_0^2 \frac{7}{32}\,x^6\,(2-x)\,dx \tag{15.6.18}$$

$$\int_0^2 f_{(k,8)}(x)\,dx = \int_0^2 56\left(\frac{x}{2}\right)^6\left(1-\frac{x}{2}\right)d\left(\frac{x}{2}\right) \tag{15.6.19}$$

Let new random variable be t such that $t = x/2$, New sample be $\{T_1, \cdots T_8\}$ such that $T_i = X_i/2$.

$$f_{(k,8)}(t) = 56\,t^6\,(1-t) \tag{15.6.20}$$

$$\int_0^2 f_{(k,8)}(x)\,dx = \int_0^1 f_{(k,8)}(t)\,dt = 1 \tag{15.6.21}$$

The Uniform distribution of new random sample is on $[0,1]$ such that PDF = 1 and CDF = t

$f(k,8)(x)$

$$f_{(k,8)}(t) = \begin{cases} 56\,t^6\ (1-t)\,, & 0 < t < 1, \\ 0, & \text{otherwise}, \end{cases} \tag{15.6.22}$$

Since equation (15.6.22) is a Beta distribution with $r = k,\ s = n - k + 1$

$$r - 1 = k - 1 = 6 \tag{15.6.23}$$

$$\therefore k = 7 \tag{15.6.24}$$

Hence the **value of k is 7**

15.7. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n (≥ 2) from a distribution having the probability density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) & x > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{15.7.1}$$

where $\theta \in (0,\infty)$. Let $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$ and $T = \sum_{i=1}^{n} X_i$. Then $E(X_{(1)}|T)$ equals ....

(A) $\dfrac{T}{n^2}$

(B) $\dfrac{T}{n}$

(C) $\dfrac{(n+1)T}{2n}$

(D) $\dfrac{(n+1)^2T}{4n^2}$

Solution:

Lemma 15.8 (Sum of Gamma random variables)**.** *Suppose that $X_i \sim \Gamma(k,\theta), i = 1,\ldots,n$. Then $T = \sum_{i=1}^{n} X_i \sim \Gamma(nk,\theta)$.*

Definition 42 (Laplace transform)**.** *Laplace transform is an integral transform that converts a real function of a real variable t to a function of a complex variable s. The laplace transform of a function $f(t)$ evaluated at s is defined by*

$$F(s) = \int_0^\infty f(t)e^{-st}dt \tag{15.7.2}$$

Lemma 15.9. *If the laplace transform of a function $f(t)$ at s is 0 for all $s > 0$, then the function $f(t) = 0$ almost everywhere in $(0,\infty)$.*

Definition 43 (Complete Statistic)**.** *The statistic T is said to be complete for the distribution of X if, for every measurable function g, if*

$$E(g(T)) = 0\ \forall\ \theta \Rightarrow \Pr(g(T) = 0) = 1 \tag{15.7.3}$$

Definition 44 (Sufficient Statistic)**.** *A statistic $t = T(X)$ is sufficient for underlying parameter θ precisely if the conditional probability distribution of the data X, given the statistic $t = T(X)$, does not depend on the parameter θ.*

Theorem 15.8 (Fischer-Neymann Factorisation theorem)**.** *If the probability density function is $f_\theta(x)$, then T is sufficient for θ if and only if nonnegative functions g and h can be found such that*

$$f_\theta(x) = h(x)g_\theta(T(x)) \tag{15.7.4}$$

Lemma 15.10.

$$T = \sum_{i=1}^{n} X_i \tag{15.7.5}$$

is a complete and sufficient statistic.

Proof. a) *(Sufficiency)*

$$f_X(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{15.7.6}$$

$$= \prod_{i=1}^{n} \frac{1}{\theta}\exp\left(-\frac{x_i}{\theta}\right) \tag{15.7.7}$$

$$= 1 \cdot \frac{1}{\theta^n}\exp\left(-\frac{T}{\theta}\right) \tag{15.7.8}$$

$$= h(X) \cdot g(T,\theta) \tag{15.7.9}$$

with

$$g(T,\theta) = \frac{1}{\theta^n}\exp\left(-\frac{T}{\theta}\right) \tag{15.7.10}$$

$$h(X) = 1 \tag{15.7.11}$$

Therefore, T is a sufficient statistic.

b) *(Completeness)* From Lemma 15.8, $X_i \sim \Gamma\left(1, \frac{1}{\theta}\right) \implies T = \sum_{i=1}^{n} X_i \sim \Gamma\left(n, \frac{1}{\theta}\right)$.

$$\therefore E(g(T)) = \int_0^\infty g(t)\frac{t^{n-1}e^{-\frac{t}{\theta}}}{\theta^n(n-1)!}dt \tag{15.7.12}$$

$$= \frac{1}{\theta^n(n-1)!}\int_0^\infty g(t)t^{n-1}e^{-\frac{t}{\theta}}dt \tag{15.7.13}$$

$$= \frac{1}{\theta^n(n-1)!}F\left(\frac{1}{\theta}\right) \tag{15.7.14}$$

where,

$$f(t) = t^{n-1}g(t) \overset{\mathcal{L}}{\rightleftharpoons} F\left(\frac{1}{\theta}\right) \tag{15.7.15}$$

is the Laplace transform relationship. Then

$$E\left(g(T)\right) = 0 \forall\theta > 0 \implies F\left(\frac{1}{\theta}\right) = 0 \forall\theta > 0 \tag{15.7.16}$$

$$\implies f(t) = 0 \text{ a.e } (0, \infty) \tag{15.7.17}$$

$$\text{or, } g(t) = 0 \text{ a.e } (0, \infty) \tag{15.7.18}$$

$$\implies \Pr\left(g(t) = 0\right) = 1 \tag{15.7.19}$$

from Lemma 15.9. $\therefore T$ is a complete statistic.

□

Theorem 15.9 (Lehmann–Scheffé theorem)**.** *If T is a complete sufficient statistic for θ and*

$$E(g(T)) = \tau(\theta) \tag{15.7.20}$$

then $g(T)$ is the uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(\theta)$.

Proposition 15.1. *$E(X_{(1)}|T)$ is the uniformly minimum-variance unbiased estimator (UMVUE) of $E(X_{(1)})$*

Proof. By the law of total expectation,

$$E\left(E(X_{(1)}|T)\right) = E(X_{(1)}) \tag{15.7.21}$$

We know that $T = \sum_{i=1}^{n} X_i$ is a complete and sufficient statistic by Lemma 15.10. By Lehmann–Scheffé theorem, with

$$\theta = X_{(1)}, \tag{15.7.22}$$

$$\tau(x) = E(x), \tag{15.7.23}$$

$$g(T) = E(X_{(1)}|T). \tag{15.7.24}$$

it follows from (15.7.21) that $E(X_{(1)}|T)$ is the UMVUE of $E(X_{(1)})$. □

Proposition 15.2. *$\frac{T}{n^2}$ is uniformly minimum-variance unbiased estimator (UMVUE) of $E(X_{(1)})$*

Proof. a) Let's find the probability distribution function and the expectation value of $X_{(1)}$:

$$\Pr\left(X_{(1)} > x\right) = \Pr\left(X_1 > x\right)\dots\Pr\left(X_n > x\right) \tag{15.7.25}$$

$$= (1 - F_{X_1}(x))\dots(1 - F_{X_n}(x)) \tag{15.7.26}$$

$$= (1 - F_{X_1}(x))^n \tag{15.7.27}$$

$$= \exp\left(-\frac{nx}{\theta}\right) \tag{15.7.28}$$

$$F_{X_{(1)}}(x) = 1 - \exp\left(-\frac{nx}{\theta}\right) \tag{15.7.29}$$

$$f_{X_{(1)}}(x) = \frac{n}{\theta}\exp\left(-\frac{nx}{\theta}\right) \tag{15.7.30}$$

Therefore, $X_{(1)}$ follows an exponential distribution with mean $\frac{\theta}{n}$.

$$E(X_{(1)}) = \frac{\theta}{n} \tag{15.7.31}$$

b) $\frac{T}{n^2}$ is uniformly minimum-variance unbiased estimator (UMVUE) of $E(X_{(1)})$, since $E\left(\frac{T}{n^2}\right) = E(X_{(1)})$:
Note that,

$$E\left(\frac{T}{n^2}\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n^2}\right) \tag{15.7.32}$$

$$= \frac{E(\sum_{i=1}^{n} X_i)}{n^2} \tag{15.7.33}$$

$$= \sum_{i=1}^{n} \frac{E(X_i)}{n^2} \tag{15.7.34}$$

$$= \sum_{i=1}^{n} \frac{\theta}{n^2} \tag{15.7.35}$$

$$= \frac{\theta}{n} \tag{15.7.36}$$

$$= E(X_{(1)}) \tag{15.7.37}$$

Therefore, by Lehmann–Scheffé theorem, with

$$\theta = X_{(1)}, \tag{15.7.38}$$

$$\tau(x) = E(x), \tag{15.7.39}$$

$$g(T) = \frac{T}{n^2}, \tag{15.7.40}$$

it follows that $\frac{T}{n^2}$ is UMVUE of $E(X_{(1)})$.

□

Since there exists a unique UMVUE for $E(X_{(1)})$, it follows from Proposition 15.1 and

Proposition 15.2 that

$$E(X_{(1)}|T) = \frac{T}{n^2} \tag{15.7.41}$$

Hence, option A is correct.

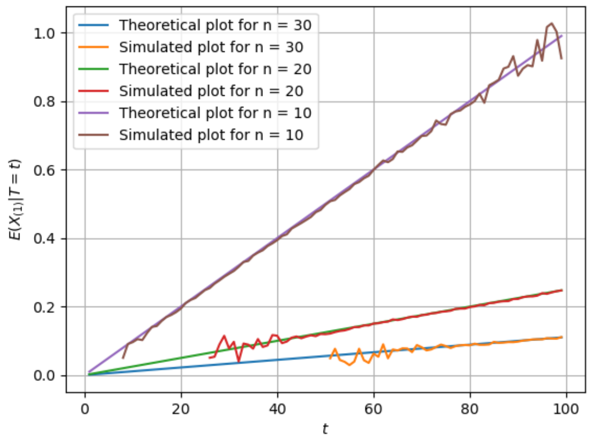


Fig. 15.7.1: Theory vs Simulated plot of $E(X_{(1)}|T)$

15.8. Let $X_1, X_2, ..., X_n$ be independent and identically distributed Bernoulli(θ), where $0 < \theta < 1$ and $n > 1$. Let the prior density of θ be proportional to $\frac{1}{\sqrt{\theta(1-\theta)}}$, $0 < \theta < 1$. Define $S = \sum_{i=1}^{n} X_i$.

Then valid statements among the following are:

1. The posterior mean of θ does not exist;
2. The posterior mean of θ exists;
3. The posterior mean of θ exists and it is larger than the maximum likelihood estimator for all values of S.
4. The posterior mean of θ exists and it is larger than the maximum likelihood estimator for some values of S.

Solution:

Definition 45 (Posterior Mean). *Posterior mean is the mean of the posterior distribution of θ, i.e.,*

$$E(\theta|X) = \int \theta \, f(\theta|X) \, d\theta \tag{15.8.1}$$

Definition 46 (Beta Function). *The beta function, $B(x, y)$, is defined by the integral*

$$\begin{aligned} B(x,y) &= \int_0^1 t^{x-1} (1-t)^{y-1} \, dt \\ &= \frac{x+y}{xy} \times \frac{1}{{}^{x+y}C_x} \end{aligned} \tag{15.8.2}$$

where $Re(x) > 0$ and $Re(y) > 0$.

Lemma 15.11. *Let $f(\theta)$ be the prior density of θ. Then*

$$f(\theta) = \frac{\theta^{-\frac{1}{2}} (1-\theta)^{-\frac{1}{2}}}{B\left(\frac{1}{2}, \frac{1}{2}\right)} \tag{15.8.3}$$

Proof.

$$\begin{aligned} f(\theta) &\propto \frac{1}{\sqrt{\theta(1-\theta)}} \\ \implies f(\theta) &= \frac{K}{\sqrt{\theta(1-\theta)}} \end{aligned} \tag{15.8.4}$$

where K is the proportionality constant.

$$\begin{aligned} \int_0^1 f(\theta) \, d\theta &= 1 \\ \implies K \int_0^1 \frac{1}{\sqrt{\theta(1-\theta)}} \, d\theta &= 1 \end{aligned} \tag{15.8.5}$$

From (15.8.2) we get,

$$K \times B\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$\implies K = \frac{1}{B\left(\frac{1}{2}, \frac{1}{2}\right)} \tag{15.8.6}$$

$$\therefore f(\theta) = \frac{\theta^{-\frac{1}{2}} (1-\theta)^{-\frac{1}{2}}}{B\left(\frac{1}{2}, \frac{1}{2}\right)} \tag{15.8.7}$$

□

Definition 47 (Likelihood Function). *The likelihood function is defined as*

$$f(\mathbf{X}|\theta) = \prod_{i=1}^{n} f(X_i|\theta) \tag{15.8.8}$$

Lemma 15.12.

$$f(\mathbf{X}|\theta) = \theta^S (1-\theta)^{n-S} \tag{15.8.9}$$

Proof. From (15.8.8) we get,

$$\begin{aligned} f(\mathbf{X}|\theta) &= \prod_{i=1}^{n} \theta^{X_i} (1-\theta)^{1-X_i} \\ &= \theta^S (1-\theta)^{n-S} \end{aligned} \tag{15.8.10}$$

□

Definition 48 (Maximum Likelihood Esitmator). *The maximum likelihood estimator is the value which maximizes the likelihood function,* *i.e.,*

$$MLE = arg_{\theta\in(0,1)}\ max\,(f\,(\mathbf{X}|\theta)) \tag{15.8.11}$$

Lemma 15.13.

$$MLE = \frac{S}{n} \tag{15.8.12}$$

Proof. Using log of likelihood function in (15.8.9) and differentiating, we get,

$$\ln(f\,(\mathbf{X}|\theta)) = S\,\ln(\theta) + (n-S)\ln(1-\theta) \tag{15.8.13}$$

$$\frac{\partial \ln(f\,(X|\theta))}{\partial\theta} = \frac{S}{\theta} + \frac{S-n}{1-\theta} = 0$$
$$\therefore \text{MLE} = \frac{S}{n} \tag{15.8.14}$$

□

Definition 49 (Posterior Density)**.** *The posterior density of θ is defined as*

$$f\,(\theta|\mathbf{X}) = \frac{f\,(\mathbf{X},\theta)}{f\,(\mathbf{X})} \tag{15.8.15}$$

Lemma 15.14.

$$f\,(\theta|\mathbf{X}) = \frac{\theta^{S-\frac{1}{2}}\,(1-\theta)^{n-S-\frac{1}{2}}}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)} \tag{15.8.16}$$

Proof. Using Bayes' theorem,

$$\begin{aligned} f\,(\theta|X) &= \frac{f\,(\mathbf{X},\theta)}{\int_0^1 f\,(\mathbf{X},\theta)\ d\theta} \\ &= \frac{f\,(\mathbf{X}|\theta)\ f\,(\theta)}{\int_0^1 f\,(\mathbf{X}|\theta)\ f\,(\theta)\ d\theta} \end{aligned} \tag{15.8.17}$$

Using (15.8.3) and (15.8.9) in (15.8.17) we get,

$$f\,(\theta|\mathbf{X}) = \frac{\theta^{S-\frac{1}{2}}\,(1-\theta)^{n-S-\frac{1}{2}}}{\int_0^1 \theta^{S-\frac{1}{2}}\,(1-\theta)^{n-S-\frac{1}{2}}\ d\theta} \tag{15.8.18}$$

Using (15.8.2) we get,

$$f\,(\theta|\mathbf{X}) = \frac{\theta^{S-\frac{1}{2}}\,(1-\theta)^{n-S-\frac{1}{2}}}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)} \tag{15.8.19}$$

□

Corollary 15.10.

$$E\,(\theta|\mathbf{X}) = \frac{S+\frac{1}{2}}{n+1} \tag{15.8.20}$$

Proof. From (15.8.1) we get,

$$\begin{aligned} E\,(\theta|\mathbf{X}) &= \int_0^1 \theta\ f\,(\theta|\mathbf{X})\ d\theta \\ &= \int_0^1 \frac{\theta^{S+\frac{1}{2}}\,(1-\theta)^{n-S-\frac{1}{2}}}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)}\,d\theta \\ &= \frac{B\left(S+\frac{3}{2}, n-S+\frac{1}{2}\right)}{B\left(S+\frac{1}{2}, n-S+\frac{1}{2}\right)} \end{aligned} \tag{15.8.21}$$

Using (15.8.2) in (15.8.21) we get

$$E\,(\theta|\mathbf{X}) = \frac{S+\frac{1}{2}}{n+1} \tag{15.8.22}$$

□

Corollary 15.11. *For $E\,(\theta|\mathbf{X})$ to be greater than MLE,*

$$n > 2\,S \tag{15.8.23}$$

Proof.

$$\frac{S+\frac{1}{2}}{n+1} > \frac{S}{n}$$
$$\therefore n > 2\,S \tag{15.8.24}$$

□

a) From (15.8.20) we get

$$E\,(\theta|\mathbf{X}) = \frac{S+\frac{1}{2}}{n+1}$$

$\implies$ for $E\,(\theta|\mathbf{X})$ to exist,

$$n \neq -1$$

Given in the question,

$$n > 1$$

$\implies$ $E\,(\theta|\mathbf{X})$ exists.
$\therefore$ Option 2 is correct and option 1 is incorrect.

b) From (15.8.23) we get

$$E\,(\theta|\mathbf{X}) > \text{ MLE}$$

when

$$n > 2\,S$$

$\implies$ $E\,(\theta|\mathbf{X}) >$ MLE for some values of S.
$\therefore$ Option 4 is correct and option 3 is incorrect.

$\therefore$ Option 2 and 4 are correct.

15.9. Let X_1, X_2, X_3, X_4, X_5 be independent and identically distributed random variables each following a uniform distribution on (0,1) and M denote their median. Then which of the following statements are true?

a) $\Pr\left(M < \frac{1}{3}\right) = \Pr\left(M > \frac{2}{3}\right)$

b) M is uniformly distributed on (0,1)

c) $E(M) = E(X_1)$

d) $V(M) = V(X_1)$

Solution:

Theorem 15.12 (Uniform distribution)**.** *A random variable X is said to be uniformly distributed in $a \leq x \leq b$ if its density function is*

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases} \tag{15.9.1}$$

and the distribution is called uniform distribution. The mean and variance are respectively,

$$\mu = \frac{a+b}{2} \tag{15.9.2}$$

$$\sigma^2 = \frac{(b-a)^2}{12} \tag{15.9.3}$$

Theorem 15.13 (Beta distribution)**.** *The Beta distribution is a continuous distribution defined on the range (0, 1) where*

$$f_X(x) = \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1} \tag{15.9.4}$$

$$F_X(x) = \int_0^x \frac{1}{B(r,s)} x^{r-1}(1-x)^{s-1} dx = \frac{B_x(r,s)}{B(r,s)} \tag{15.9.5}$$

$$B(r,s) = \int_0^1 x^{r-1}(1-x)^{s-1} dx = \frac{(r-1)!(s-1)!}{(r+s-1)!} \tag{15.9.6}$$

$$B_x(r,s) = \int_0^x x^{r-1}(1-x)^{s-1} dx \tag{15.9.7}$$

$$E(X) = \frac{r}{r+s} \tag{15.9.8}$$

$$Var(X) = \frac{rs}{(r+s)^2(r+s+1)} \tag{15.9.9}$$

and $B(r,s)$ is the beta function.

Definition 50 (Order statistics)**.** *For a given statistical sample $\{X_1, X_2, \cdots X_n\}$, the order statistics is obtained by sorting the sample in ascending order. It denoted as $\{X_{(1)}, X_{(2)}, \cdots X_{(n)}\}$.*

Definition 51 (Median of order statistics)**.** *Median is defined as the middle number of a sorted sample. It is denoted by M and defined using order statistics of a sample as*

$$M = \begin{cases} X_{((n+1)/2)}, & \text{if } n \text{ is odd,} \\ \dfrac{X_{(n/2)} + X_{(n/2+1)}}{2}, & \text{if } n \text{ is even,} \end{cases} \tag{15.9.10}$$

Remark 15.14. *The order statistics of the uniform distribution on the unit interval have marginal distributions belonging to the Beta distribution family.*

$$X_{(k)} \sim B(k, n+1-k) \tag{15.9.11}$$

a) From definition (51) median M is given by

$$M = X_{((5+1)/2)} \tag{15.9.12}$$

$$= X_{(3)} \tag{15.9.13}$$

From remark (15.14)

$$X_{(3)} \sim B(3,3) \tag{15.9.14}$$

From (15.9.6)

$$B(3,3) = \frac{(3-1)!(3-1)!}{(3+3-1)!} = \frac{1}{30} \tag{15.9.15}$$

From (15.9.4)

$$f(x) = 30x^2(1-x)^2 \tag{15.9.16}$$

From (15.9.5)

$$F(x) = \int_0^x 30x^2(1-x)^2 dx \tag{15.9.17}$$

$$= 30x^3\left(\frac{1}{3} + \frac{x^2}{5} - \frac{x}{2}\right) \tag{15.9.18}$$

$$\Pr\left(M < \frac{1}{3}\right) = F\left(\frac{1}{3}\right) = 0.20987 \quad (15.9.19)$$

$$\Pr\left(M > \frac{2}{3}\right) = F(1) - F\left(\frac{2}{3}\right) = 0.20987 \quad (15.9.20)$$

$$\therefore \Pr\left(M < \frac{1}{3}\right) = \Pr\left(M > \frac{2}{3}\right) \quad (15.9.21)$$

Hence **Option 1 is true.**

b) From (15.9.13), median M is a third order statistic. Clearly from remark (15.14), M is not an uniform distribution.
Hence **Option 2 is false.**

c) From (15.9.1)

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \quad (15.9.22)$$

From (15.9.2)

$$E(X_1) = \frac{1}{2} \quad (15.9.23)$$

From (15.9.8)

$$E(M) = \frac{3}{3+3} = \frac{1}{2} \quad (15.9.24)$$

$$\therefore E(M) = E(X_1) \quad (15.9.25)$$

Hence **Option 3 is true.**

d) From (15.9.3)

$$V(X_1) = \frac{1}{12} \quad (15.9.26)$$

From (15.9.9)

$$V(M) = \frac{1}{28} \quad (15.9.27)$$

$$\therefore V(M) \neq V(X_1) \quad (15.9.28)$$

Hence **Option 4 is false.**

15.10. Suppose n units are drawn from a population of N units sequentially as follows. A random sample

$$U_1, U_2, ...U_N \text{ of size N, drawn from } U(0,1) \quad (15.10.1)$$

The k-th population unit is selected if

$$U_k < \frac{n - n_k}{N - k + 1}, k = 1, 2, ..N.\text{where, } n_1 = 0, n_k = \quad (15.10.2)$$

number of units selected out of first k-1 units for each k = 2, 3, ..N. Then,

a) The probability of inclusion of the second unit in the sample

$$\text{is } \frac{n}{N} \quad (15.10.3)$$

b) The probability of inclusion of the first and the second unit in the sample

$$\text{is } \frac{n(n-1)}{N(N-1)} \quad (15.10.4)$$

c) The probability of not including the first and including the second unit in the sample

$$\text{is } \frac{n(N-n)}{N(N-1)} \quad (15.10.5)$$

d) The probability of including the first and not including the second unit in the sample

$$\text{is } \frac{n(n-1)}{N(N-1)} \quad (15.10.6)$$

Solution:

$$\text{Defining random variable } X \in \{0, 1, 2, ..N\} \quad (15.10.7)$$

$$\text{Where, } X = i \text{ when ith unit is included.} \quad (15.10.8)$$

The first unit in the sample is included if

$$U_1 < \frac{n - n_1}{N - 1 + 1} \quad (15.10.9)$$

$$\text{Here, } n_1 = 0 \text{ is given in the qn.} \quad (15.10.10)$$

$$\therefore \Pr(X = 1) = \frac{n}{N} \quad (15.10.11)$$

a) For k=2,

$$n_2 = 1 \text{ when, first unit is included.} \quad (15.10.12)$$

$$U_2 < \frac{n - n_2}{N - 2 + 1}\left(= \frac{n-1}{N-1}\right) \quad (15.10.13)$$

$$\therefore \Pr(X = 2 \mid X = 1) = \frac{n-1}{N-1} \quad (15.10.14)$$

$$\Pr(X = 1, X = 2)$$
$$= \Pr(X = 2 \mid X = 1) \times \Pr(X = 1) \quad (15.10.15)$$
$$\therefore \Pr(X = 1, X = 2) = \frac{n(n-1)}{N(N-1)} \quad (15.10.16)$$

$n_2 = 0$ when, first unit is not included.

$$(15.10.17)$$
$$U_2 < \frac{n - n_2}{N - 2 + 1}\left(= \frac{n}{N-1}\right) \quad (15.10.18)$$
$$\therefore \Pr(X = 2 \mid X \neq 1) = \frac{n}{N-1} \quad (15.10.19)$$

$$\Pr(X \neq 1, X = 2)$$
$$= \Pr(X = 2 \mid X \neq 1) \times \Pr(X \neq 1) \quad (15.10.20)$$
$$\therefore \Pr(X \neq 1, X = 2) = \left(1 - \frac{n}{N}\right) \times \frac{n}{N-1} \quad (15.10.21)$$
$$\therefore \Pr(X \neq 1, X = 2) = \frac{n(N-n)}{N(N-1)} \quad (15.10.22)$$

From (15.10.16) and (15.10.22)

$$\Pr(X = 2) = \frac{n(n-1)}{N(N-1)} + \frac{n(N-n)}{N(N-1)} = \frac{n}{N} \quad (15.10.23)$$

Hence, option 1 is correct.

b) From (15.10.16)

$$\Pr(X = 1, X = 2) = \frac{n(n-1)}{N(N-1)} \quad (15.10.24)$$

Hence, option 2 is correct.

c) From (15.10.22)

$$\Pr(X \neq 1, X = 2) = \frac{n(N-n)}{N(N-1)} \quad (15.10.25)$$

Hence, option 3 is correct.

d)

$$\Pr(X = 1, X \neq 2) = \frac{n}{N} \times \left(1 - \frac{n}{N}\right) = \frac{n(N-n)}{N^2} \quad (15.10.26)$$

Hence, option 4 is incorrect.

Therefore, Options 1, 2, 3 are correct

15.11. Let $X_1, X_2, X_3, .., X_n$ be independent random variables follow a common continuous distribution **F**, which is symmetric about 0. For i=1,2,3,..n, define

$$S_i = \begin{cases} 1 & if\ X_i > 0 \\ -1 & if\ X_i < 0\ and \\ 0 & if\ X_i = 0 \end{cases} \quad (1.1)$$

R_i=rank of $|X_i|$ in the set$\{|X_1|, |X_2|, .., |X_n|\}$.Which of the following statements are correct?

a) $S_1, S_2, .., S_n$ are independent and identically distributed.
b) $R_1, R_2, .., R_n$ are independent and identically distributed.
c) $S = (S_1, S_2, .., S_n)$ and $R = (R_1, R_2, .., R_n)$ are independent.

Solution:

A sequence $\{X_i\}$ is an Independent and identical if and only if $F_{X_n}(x) = F_{X_k}(x)$ $\forall$ n,k,x and any subset of terms of the sequence is a set of mutually independent random variables. Where F is the probability density function.

a) As the probability distribution function of $\{X_i\}$ is symmetric about origin we can say that

$$F_{X_i}(-x) = F_{X_i}(x) \forall x \in R \quad (2.1)$$

and the mean of the distribution(μ)

$$\mu = 0 \quad (2.2)$$

The sequence S_i depend on X_i as mention in 1.1, as each S_i depend only on X_i we can say that sequence S_i is independent.

$$\Pr(S_1 = 1, S_2 = 1, ..., S_n = 1) = \Pi_{i=1}^{n} \Pr(S_i = 1) \quad (2.3)$$

Any subset of terms of sequence $\{S_i\}$ is a set of mutually independent random variables and its distribution is identical.

$$F_{S_n}(s) = F_{S_k}(s) \quad \forall s, k, n \quad (2.4)$$

So, the sequence $\{S_i\}$ is independent and identical.

b) **Ranking** refers to the data transformation in which the numerical or ordinary values are replaced by the rank of numerical value when compared to a list of other values.Usually we follow increasing order for ranking.

Ranking of a sequence depend on every elements of the sequence.Let $\{R_i\}$ be the output sequence of the ranking function of $\{|X_i|\}$.

$$R_k = \text{rank of } |X_k| \text{ in the set}\{|X_1|, |X_2|, .., |X_n|\} \tag{2.5}$$

As R_k depend not only on $|X_k|$ but on the rest of the elements of the set$\{|X_1|, |X_2|, .., |X_n|\}$. So the sequence R_i is not independent. Hence R_i is not an independent and identical distribution.

c) As the i^{th} element of sequence R depends only on set$\{|X_1|, |X_2|, .., |X_n|\}$, we can say that sequence S and R are independent.

Answer:A,C

15.12. A simple random variable of size n will be drawn from a class of 125 students, and the mean mathematics score of the sample will be computed, If the standard error of the sample mean for "with replacement sampling" is twice as much as the standard error of the sample mean for "without replacement sampling", the value of n is ?

a) 32
b) 63
c) 79
d) 94

Solution: Let N be the population size so, N=120. The given sample size is n. **Notations :** y : student under consideration. y_i : Maths marks of i^{th} student in the sample. Y : student of class. Y_i : Maths marks of i^{th} student in the class. $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$: Average of sample class. $\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$: Average of whole class. $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$: S=Std dev of the class. $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$: Variance of the class. Standard error of sample mean $SE_{mean} = \frac{s}{\sqrt{n}}$.

Where

s = standard deviation of sample mean.

n = sample class size.

Variance of the $\bar{y}$

$$V(\bar{y}) = E(\bar{y} - \overline{Y})^2 \tag{15.12.1}$$

$$= E\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{Y})\right]^2 \tag{15.12.2}$$

$$= E\left[\frac{1}{n^2}\sum_{i=1}^{n}(y_i - \overline{Y})^2 + \frac{1}{n^2}\sum\sum_{1\le i\ne j\le n}(y_i - \overline{Y})(y_j - \overline{Y})\right] \tag{15.12.3}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}E(y_i - \overline{Y})^2 + \frac{1}{n^2}\sum\sum_{1\le i\ne j\le n}E(y_i - \overline{Y})(y_j - \overline{Y}) \tag{15.12.4}$$

$$\text{Let } K = \sum\sum_{1\le i\ne j\le n}E(y_i - \overline{Y})(y_j - \overline{Y}) \tag{15.12.5}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 + \frac{K}{n^2} \tag{15.12.6}$$

$$= \frac{1}{n^2}n\sigma^2 + \frac{K}{n^2} \tag{15.12.7}$$

$$= \frac{N-1}{Nn}S^2 + \frac{K}{n^2} \tag{15.12.8}$$

Finding the value of K in case of Simple random sampling with repetition (SRSWR)and Simple random sampling without repetition(SRSWOR) allows us to calculate the variance of mean. **K value in case of SRSWOR**

$$K = \sum\sum_{1\le i\ne j\le n}E(y_i - \overline{Y})(y_j - \overline{Y})$$

Consider

$$E(y_i - \overline{Y})(y_j - \overline{Y}) = \frac{1}{N(N-1)}\sum\sum_{1\le k\ne l\le n}E(y_k - \overline{Y})(y_l - \overline{Y})$$

Since

$$\left[\sum_{k=1}^{N}(y_k - \overline{Y})\right]^2 = \sum_{i=1}^{N}(y_k - \overline{Y})^2 + \sum\sum_{1\le k\ne l\le n}E(y_k - \overline{Y})(y_l - \overline{Y})$$

$$\implies 0 = (N-1)S^2 + \sum\sum_{1\le k\ne l\le n} E(y_k - \overline{Y})(y_l - \overline{Y})$$
$$\implies E(y_i - \overline{Y})(y_j - \overline{Y}) = \frac{1}{N(N-1)}(N-1)(-S^2)$$
$$\implies K = n(n-1)\frac{(-S^2)}{N}$$

Putting this value in (15.12.8) gives us

$$V(\overline{y})_{WOR} = \frac{N-1}{Nn}S^2 + \frac{n-1(-S^2)}{Nn} \quad (15.12.9)$$
$$= \frac{N-n}{Nn}S^2 \quad (15.12.10)$$

K value in case of SRSWR

$$K = \sum\sum_{1\le i\ne j\le n} E(y_i - \overline{Y})(y_j - \overline{Y})$$

Since we are selecting the samples with replacements choosing i^{th} and j^{th} sample is independent of each other. So,

$$K = \sum\sum_{1\le i\ne j\le n} E(y_i - \overline{Y})E(y_j - \overline{Y})$$
$$= 0$$

(Since deviation about mean is 0)

Putting K=0 in (15.12.8) we get

$$V(\overline{y})_{WR} = \frac{N-1}{Nn}S^2 \quad (15.12.11)$$

From equation (15.12.10) standard error of mean of sample class without repetition

$$SE_{WOR} = \frac{s}{\sqrt{n}} \quad (15.12.12)$$
$$= \sqrt{\frac{V(\overline{y})_{WOR}}{n}} \quad (15.12.13)$$
$$= \sqrt{\frac{N-n}{Nn^2}}S \quad (15.12.14)$$

From equation (15.12.11) standard error of mean of sample class with repetition

$$SE_{WR} = \sqrt{\frac{V(\overline{y})_W R}{n}} \quad (15.12.15)$$
$$= \sqrt{\frac{N-1}{Nn^2}}S \quad (15.12.16)$$

Given to find the value of n if $2 \times SE_{WOR} = SE_{WR}$. From (15.12.14) and (15.12.16) we can write

$$2\sqrt{\frac{N-n}{Nn^2}}S = \sqrt{\frac{N-1}{Nn^2}}S \quad (15.12.17)$$
$$\implies 4(N-n) = N-1 \quad (15.12.18)$$
$$\implies 4N + 1 - N = 4n \quad (15.12.19)$$
$$\implies 4n = 3(125) + 1 \quad (15.12.20)$$
$$\implies n = 94 \quad (15.12.21)$$

Therefore the sample size for the given condition to be met is n=94.(**Option D**)

15.13. Let X_1 and X_2 be i.i.d. with probability mass function $f_\theta(x) = \theta^x(1-\theta)^{1-x};\ x = 0, 1$ where $\theta \in (0, 1)$. Which of the following statements are true?

a) $X_1 + 2X_2$ is a sufficient statistic
b) $X_1 - X_2$ is a sufficient statistic
c) $X_1^2 + X_2^2$ is a sufficient statistic
d) $X_1^2 + X_2$ is a sufficient statistic

Solution: Given that, X_1 and X_2 are i.i.d. with probability mass function

$$f(x) = \begin{cases}(1-\theta) & x = 0\\ \theta & x = 1\end{cases} \quad (15.13.1)$$

A statistic $t = T(X)$ is sufficient for a parameter θ if the conditional probability distribution of the data, given the statistic $t = T(X)$ does not depend on the parameter θ. i.e,

$$P_\theta(X_1 = x_1, X_2 = x_2 | T = t) \quad (15.13.2)$$

is independent of θ for all x_1, x_2 and t

a) Let $T = X_1 + 2X_2$
Consider a case where $x_1 = 0, x_2 = 0$ and $t = 0$

$$\Pr(T = 0) = \Pr(X_1 + 2X_2 = 0) \quad (15.13.3)$$
$$= \Pr(X_1 = 0, X_2 = 0) \quad (15.13.4)$$

As X_1 and X_2 are independent

$$\Pr(T = 0) = \Pr(X_1 = 0)\Pr(X_2 = 0)$$
$$= (1-\theta)^2 \quad (15.13.5)$$

The conditional probability,

$$\Pr(X_1 = 0, X_2 = 0 | T = 0)$$
$$= \frac{\Pr((X_1 = 0, X_2 = 0) \cap (T = 0))}{\Pr(T = 0)} \quad (15.13.6)$$

From (15.13.4), $(X_1 = 0, X_2 = 0) \subseteq (T = 0)$

$$= \frac{\Pr(X_1 = 0, X_2 = 0)}{\Pr(T = 0)} = \frac{(1-\theta)^2}{(1-\theta)^2} = 1 \tag{15.13.7}$$

Similarly, conditional probabilities for other values of x_1, x_2 and t are given in table 15.13.1

x_1	x_2	t $t = X_1 + 2X_2$	Conditional probability $P_\theta(X_1 = x_1, X_2 = x_2 \mid T = t)$
0	0	0 otherwise	1 0
1	0	1 otherwise	1 0
0	1	2 otherwise	1 0
1	1	3 otherwise	1 0

TABLE 15.13.1: Conditional Probabilities

From table 15.13.1, all the conditional probabilities are independent of θ
$\therefore X_1 + 2X_2$ is a sufficient statistic.

b) Let $T = X_1 - X_2$
Consider a case where $x_1 = 0, x_2 = 0$ and $t = 0$

$$\Pr(T = 0) = \Pr(X_1 - X_2 = 0)$$
$$= \Pr(X_1 = 0, X_2 = 0) + \Pr(X_1 = 1, X_2 = 1) \tag{15.13.8}$$

As X_1 and X_2 are independent

$$= \Pr(X_1 = 0)\Pr(X_2 = 0)$$
$$+ \Pr(X_1 = 1)\Pr(X_2 = 1) = (1-\theta)^2 + \theta^2 \tag{15.13.9}$$

The conditional probability,

$$\Pr(X_1 = 0, X_2 = 0 \mid T = 0)$$
$$= \frac{\Pr((X_1 = 0, X_2 = 0) \cap (T = 0))}{\Pr(T = 0)} \tag{15.13.10}$$

From (15.13.8), $(X_1 = 0, X_2 = 0) \subseteq (T = 0)$

$$= \frac{\Pr(X_1 = 0, X_2 = 0)}{\Pr(T = 0)} = \frac{(1-\theta)^2}{(1-\theta)^2 + \theta^2} \tag{15.13.11}$$

depends on θ.
$\therefore X_1 - X_2$ is not a sufficient statistic.

c) Let $T = X_1^2 + X_2^2$
Consider a case where $x_1 = 1, x_2 = 0$ and $t = 1$

$$\Pr(T = 1) = \Pr\left(X_1^2 + X_2^2 = 1\right)$$
$$= \Pr(X_1 = 1, X_2 = 0) + \Pr(X_1 = 0, X_2 = 1)$$
$$= \theta(1-\theta) + (1-\theta)\theta = 2\theta(1-\theta) \tag{15.13.12}$$

The conditional probability,

$$\Pr(X_1 = 1, X_2 = 0 \mid T = 1)$$
$$= \frac{\Pr((X_1 = 1, X_2 = 0) \cap (T = 0))}{\Pr(T = 1)} \tag{15.13.13}$$

From (15.13.12), $(X_1 = 1, X_2 = 0) \subseteq (T = 1)$

$$= \frac{\Pr(X_1 = 1, X_2 = 0)}{\Pr(T = 1)} = \frac{\theta(1-\theta)}{2\theta(1-\theta)} = \frac{1}{2} \tag{15.13.14}$$

Similarly, conditional probabilities for other values of x_1, x_2 and t are given in table 15.13.2

x_1	x_2	t $t = X_1^2 + X_2^2$	Conditional probability $P_\theta(X_1 = x_1, X_2 = x_2 \mid T = t)$
0	0	0 otherwise	1 0
1	0	1 otherwise	$\frac{1}{2}$ 0
0	1	1 otherwise	$\frac{1}{2}$ 0
1	1	2 otherwise	1 0

TABLE 15.13.2: Conditional Probabilities

From table 15.13.2, all the conditional probabilities are independent of θ
$\therefore X_1^2 + X_2^2$ is a sufficient statistic.

d) Let $T = X_1^2 + X_2$
Consider a case where $x_1 = 1, x_2 = 0$ and

$t = 1$

$$\Pr(T = 1) = \Pr\left(X_1^2 + X_2 = 1\right)$$
$$= \Pr(X_1 = 1, X_2 = 0) + \Pr(X_1 = 0, X_2 = 1)$$
$$= \theta(1-\theta) + (1-\theta)\theta = 2\theta(1-\theta) \tag{15.13.15}$$

The conditional probability,

$$\Pr(X_1 = 1, X_2 = 0 | T = 1) = \frac{\Pr((X_1 = 1, X_2 = 0) \cap (T = 0))}{\Pr(T = 1)} \tag{15.13.16}$$

From (15.13.15), $(X_1 = 1, X_2 = 0) \subseteq (T = 1)$

$$= \frac{\Pr(X_1 = 1, X_2 = 0)}{\Pr(T = 1)} = \frac{\theta(1-\theta)}{2\theta(1-\theta)} = \frac{1}{2} \tag{15.13.17}$$

Similarly, conditional probabilities for other values of x_1, x_2 and t are given in table 15.13.3

x_1	x_2	t $t = X_1^2 + X_2$	Conditional probability $P_\theta(X_1 = x_1, X_2 = x_2 \mid T = t)$
0	0	0 otherwise	1 0
1	0	1 otherwise	$\frac{1}{2}$ 0
0	1	1 otherwise	$\frac{1}{2}$ 0
1	1	2 otherwise	1 0

TABLE 15.13.3: Conditional Probabilities

From table 15.13.3, all the conditional probabilities are independent of θ
$\therefore X_1^2 + X_2$ is a sufficient statistic.

Answer : Options 1,3,4

15.14. Let X_1 and X_2 be a random sample of size two from a distribution with probability density function

$$f_\theta(x) = \theta\left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}x^2} + (1-\theta)\left(\frac{1}{2}\right)e^{-|x|},$$

$-\infty < x < \infty$,
where $\theta \in \left\{0, \frac{1}{2}, 1\right\}$. If the observed values of X_1 and X_2 are 0 and 2, respectively, then the maximum likelihood estimate of θ is

a) 0
b) $\frac{1}{2}$
c) 1
d) not unique

Solution: Given $X_1 = 0$, $X_2 = 2$, n=2 and

$$f_\theta(x) = \theta\left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}x^2} + (1-\theta)\left(\frac{1}{2}\right)e^{-|x|} \tag{15.14.1}$$

Then log of likelihood function is given by

$$l(\theta) = \sum_{i=1}^{i=n} \log f_\theta(x_i) \tag{15.14.2}$$
$$= \log f_\theta(x_1) + \log f_\theta(x_2) \tag{15.14.3}$$
$$= \log\left(\theta\left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}0^2} + (1-\theta)\left(\frac{1}{2}\right)e^{-|0|}\right) + \log\left(\theta\left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{1}{2}2^2} + (1-\theta)\left(\frac{1}{2}\right)e^{-|2|}\right) \tag{15.14.4}$$

$$= \log\left(\theta\left(\frac{1}{\sqrt{2\pi}}\right) + (1-\theta)\left(\frac{1}{2}\right)\right) + \log\left(\theta\left(\frac{1}{\sqrt{2\pi}}\right)e^{-2} + (1-\theta)\left(\frac{1}{2}\right)e^{-2}\right) \tag{15.14.5}$$
$$= 2\log\left(\theta\left(\frac{1}{\sqrt{2\pi}}\right) + (1-\theta)\left(\frac{1}{2}\right)\right) - 2 \tag{15.14.6}$$

Since likelihood $L(\theta) = e^{l(\theta)}$.

Likelihood function $L(\theta)$ at $\theta = 0, \frac{1}{2}, 1$ is given by

a) At $\theta = 0$ $\quad L(\theta = 0) = \frac{1}{4}e^{-2} = 0.0338$

b) At $\theta = 1$ $\quad L(\theta = 1) = \frac{1}{2\pi}e^{-2} = 0.0215$

c) At $\theta = \frac{1}{2}$ $\quad L(\theta = \frac{1}{2}) = \left(\frac{1}{2\sqrt{2\pi}} + \frac{1}{4}\right)^2 e^{-2} = 0.0273$

Hence the maximum likelihood estimate of θ is at $\theta = 0$

15.15. Let $\{X_n, n \geq 1\}$ be i.i.d. uniform (-1,2) random variables. Which of the following statements are true?

a) $\frac{1}{n}\sum_{i=1}^{n} X_i \to 0$ almost surely
b) $\left\{\frac{1}{2n}\sum_{i=1}^{n} X_{2i} - \frac{1}{2n}\sum_{i=1}^{n} X_{2i-1}\right\} \to 0$ almost surely
c) $\sup\{X_1, X_2, \ldots\} = 2$ almost surely
d) $\inf\{X_1, X_2, \ldots\} = -1$ almost surely

Solution: We using convergence in almost surely and Strong law of large number (SLLN)

a) *Almost sure convergence :* Let $X_1, X_2, \ldots$ be an infinite sequence of random variables. We shall say that the sequence $\{X_i\}$ converges with probability 1 (or converges almost surely (a.s.)) to a random variable Y, if

$$\Pr\left(\lim_{n\to\infty} X_n = Y\right) = 1 \tag{15.15.1}$$

$$\text{and we write}, X_n \xrightarrow{a.s.} Y \tag{15.15.2}$$

b) *SLLN :* Let X_n be i.i.d with $\mathbf{E}[|X_1|] < \infty$. Then, as $n \to \infty$, we have

$$\frac{S_n}{n} \xrightarrow{\text{a.s.}} \mathbf{E}[X_1] \implies \frac{S_n}{n} \xrightarrow{\text{P}} \mathbf{E}[X_1] \tag{15.15.3}$$

$$, \text{where}\, S_n = X_1 + \cdots + X_n \tag{15.15.4}$$

also,

$$X_i \xrightarrow{a.s.} X \implies g(X_i) \xrightarrow{a.s.} g(X) \tag{15.15.5}$$

a)

$$\frac{1}{n}(X_1 + \cdots + X_n) \to E(X) \in (-1, 2) \tag{15.15.6}$$

$$\text{as}\, n \to \infty, \tag{15.15.7}$$

according to strong law of large numbers (SLLN).
So, option (A) is incorrect.

b) using this 15.15.5 , we solve as

$$\left\{\frac{1}{2n}\sum_{i=1}^{n} X_{2i} - \frac{1}{2n}\sum_{i=1}^{n} X_{2i-1}\right\} \xrightarrow{a.s.} \left\{\frac{nX}{2n} - \frac{nX}{2n}\right\} \tag{15.15.8}$$

$$= 0 \tag{15.15.9}$$

option (B) is correct.

c) Similarly, Let $M = \sup(S)$. Then,

$$x \le M, \quad \forall x \in S \tag{15.15.10}$$

$$\forall \epsilon > 0, \quad (M - \epsilon, M] \cap S \neq \emptyset \tag{15.15.11}$$

where, S be a nonempty subset of $\mathbb{R}$ with an upper bound. Using $X_i \xrightarrow{a.s.} X$ this , we conclude that

$$\sup\{X_1, X_2, \ldots\} = 2 \textit{ almost surely} \tag{15.15.12}$$

d) Let $m = \inf(S)$. Then

$$x \ge m, \quad \forall x \in S \tag{15.15.13}$$

$$\forall \epsilon > 0, \quad [m, m + \epsilon] \cap S \neq \emptyset \tag{15.15.14}$$

where, S be a nonempty subset of $\mathbb{R}$ with an lower bound. Again using $X_i \xrightarrow{a.s.} X$ this , we conclude that

$$\inf\{X_1, X_2, \ldots\} = -1 \textit{ almost surely} \tag{15.15.15}$$

Hence (B), (C) and (D) are correct options.

15.16. $X_1, X_2, \cdots$ are independent identically distributed random variables having common density f. Assume $f(x) = f(-x)$ for all $x \in \mathbb{R}$. Which of the following statements is correct?

a) $\frac{1}{n}(X_1 + \cdots + X_n) \to 0$ in probability
b) $\frac{1}{n}(X_1 + \cdots + X_n) \to 0$ almost surely
c) $\Pr\left(\frac{1}{\sqrt{n}}(X_1 + \cdots + X_n) < 0\right) \to \frac{1}{2}$
d) $\sum_{i=1}^{n} X_i$ has the same distribution as $\sum_{i=1}^{n}(-1)^i X_i$

Solution: We using

a) (1) *Convergence in probability :* Let $X_1, X_2, \ldots$ be an infinite sequence of random variables, and let Y be another random variable. Then the sequence $\{X_n\}$ converges in probability to Y, if

$$\forall \epsilon > 0, \lim_{n\to\infty} \Pr(|X_n - Y| \ge \epsilon) = 0, \tag{15.16.1}$$

and we write

$$X_n \xrightarrow{P} Y. \tag{15.16.2}$$

(2) *Convergence in almost surely :* Let $X_1, X_2, \ldots$ be an infinite sequence of random variables. We shall say that the

sequence $\{X_i\}$ converges with probability 1 (or converges almost surely (a.s.)) to a random variable Y, if

$$\Pr\left(\lim_{n\to\infty} X_n = Y\right) = 1 \tag{15.16.3}$$

and we write

$$X_n \overset{a.s.}{\to} Y \tag{15.16.4}$$

(3) *Strong law of large number(SLLN) :* Let $X_1, X_2, \dots$ be an infinite sequence of random variables, If $\mathbf{E}[|X_1|] < \infty$. Then, as $n \to \infty$, we have

$$\frac{S_n}{n} \overset{a.s.}{\to} \mathbf{E}[X_1] \implies \frac{S_n}{n} \overset{P}{\to} \mathbf{E}[X_1], \tag{15.16.5}$$

$$\text{where, } S_n = X_1 + \dots + X_n \tag{15.16.6}$$

using SLLN, (B) are incorrect option.

b) *Relation between in probability and almost surely :* Let $Z, Z_1, Z_2, \dots$ be random variables. Suppose $Z_n \to Z$ with probability 1. Then ,we say

$$Z_n \overset{a.s.}{\to} Z \implies Z_n \overset{P}{\to} Z. \tag{15.16.7}$$

(15.16.5), also in probability also hold this equation. Hence (A) is incorrect option.

c) *Central Limit Theorem :* Let $X_1, X_2, \dots$ be i.i.d. with finite mean μ and finite variance σ^2. Let $Z \sim N(0, 1)$. Set $S_n = X_1 + \dots + X_n$, and

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \tag{15.16.8}$$

Then as $n \to \infty$, the sequence $\{Z_n\}$ converges in distribution to the Z, i.e., $Z_n \overset{D}{\to} Z$.

Consider,

$$Y = \frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}} \tag{15.16.9}$$

So,

$$E(Y) = E\left(\frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}}\right) = 0 \tag{15.16.10}$$

$$V(Y) = V\left(\frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}}\right) = \frac{1}{n} 2n = 2 \tag{15.16.11}$$

$$Y \sim N[0, 2] \tag{15.16.12}$$

we know that,

$$f(x) = f(-x) \implies \text{Symmetry about Zero,} \tag{15.16.13}$$

So,

$$\Pr(Y < 0) = \frac{1}{2} \tag{15.16.14}$$

$$\Pr\left(\frac{1}{\sqrt{n}}(X_1 + \dots + X_n) < 0\right) = \frac{1}{2} \tag{15.16.15}$$

Hence, (C) is incorrect option.

d) *Characteristic function :* For a scalar random variable X the characteristic function is defined as the expected value of $e^{it\times}$, where i is the imaginary unit, and $t \in \mathbf{R}$ is the argument of the characteristic function:

$$\begin{cases} \varphi_X : \mathbb{R} \to \mathbb{C} \\ \varphi_X(t) = \mathrm{E}\left[e^{itX}\right] = \int_{\mathbb{R}} e^{itx} dF_X(x) \\ = \int_{\mathbb{R}} e^{itx} f_X(x) dx = \int_0^1 e^{itQ_X(p)} dp \end{cases} \tag{15.16.16}$$

Here F_X is the cumulative distribution function of X,Consider, $\phi_x(t)$ is characteristic function of $X_i, i = 1, \dots, n$.

$$f(x) = f(-x) \implies \phi_x(t) = \phi_{-x}(t) \tag{15.16.17}$$

Therefore,

$$\phi_{\sum_{i=1}^n X_i}(t) = \phi_{X_1 + \ldots + X_n}(t) = \phi_{X_1}(t) \cdots \phi_{X_n}(t) \tag{15.16.18}$$

$$= [\phi_x(t)]^n \tag{15.16.19}$$

similarly,

$$\phi_{\sum_{i=1}^n (-1)^i X_i}(t) = \phi_{-X_1} + \phi_{X_2} + \ldots + \phi_{(-1)^n X_n}(t) \tag{15.16.20}$$

$$= \phi_{-X_1}(t) \cdot \phi_{X_2}(t) \cdots \phi_{(-1)^n X_n}(t) \tag{15.16.21}$$

$$= [\phi_x(t)]^n \tag{15.16.22}$$

$$\phi_{\sum_{i=1}^n X_i}(t) = \phi_{\sum_{i=1}^n (-1)^i X_i}(t) \tag{15.16.23}$$

$\therefore$ $\sum_{i=1}^n X_i$ has same distribution as $\sum_{i=1}^n (-1)^i X_i$.

Hence, only (D) is correct option.

15.17. $X_1, X_2, \dots, X_n$ are independent and identically distributed as $N(\mu, \sigma^2)$, $-\infty < \mu < \infty$, $\sigma^2 > 0$. Then

a) $\sum_1^n \frac{(X_i-\bar{X})^2}{n-1}$ is the Minimum Variance Unbiased Estimate of σ^2

b) $\sqrt{\sum_1^n \frac{(X_i-\bar{X})^2}{n-1}}$ is the Minimum Variance Unbiased Estimate of σ

c) $\sum_1^n \frac{(X_i-\bar{X})^2}{n}$ is the Maximum Likelihood Estimate of σ^2

d) $\sqrt{\sum_1^n \frac{(X_i-\bar{X})^2}{n}}$ is the Maximum Likelihood Estimate of σ

Solution: The pdf for each random variable is same as they are all identical and independent Normal Distributions with same μ and σ^2.

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(x-\mu)^2}{2\sigma^2} \tag{15.17.1}$$

Let us take our maximum likelihood function for given random variable X_i

$$L(\mu;\sigma|X_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(X_i-\mu)^2}{2\sigma^2} \tag{15.17.2}$$

Since all the random variables are i.i.d

$$L(\mu;\sigma|X_1, X_2, \dots, X_n) = \prod_{i=1}^{n} L(\mu;\sigma|X_i) \tag{15.17.3}$$

Let us denote:

$$L_m : L(\mu;\sigma|X_1, X_2, \dots, X_n) \tag{15.17.4}$$

Substituting (15.17.2) for each Random Variable in (15.17.3)

$$L_m = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{(X_i-\mu)^2}{2\sigma^2} \tag{15.17.5}$$

Taking natural log on both sides and simplifying

$$\ln L_m = \frac{-n}{2} \ln 2\pi - n \ln \sigma - \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{2\sigma^2} \tag{15.17.6}$$

In order to find Maximum Likelihood we need to maximise μ and σ w.r.t. all Random variables. Taking partial derivative w.r.t μ and taking σ as constant

$$\frac{\partial \ln L_m}{\partial \mu} = \sum_{i=1}^{n} \frac{(X_i-\mu)}{\sigma^2} \tag{15.17.7}$$

The value for μ at which L_m achieves maximum value is same in $\ln L_m$

$$\because \frac{\partial \ln L_m}{\partial \mu} = 0 \tag{15.17.8}$$

$$\therefore \sum_{i=1}^{n} \frac{(X_i-\mu)}{\sigma^2} = 0 \tag{15.17.9}$$

On simplifying the expression we get:

$$n\mu = \sum_{i=1}^{n} X_i \tag{15.17.10}$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{15.17.11}$$

Let us denote the value achieved in (15.17.11) as $\bar{X}$. Taking partial derivative w.r.t σ and taking μ as constant

$$\frac{\partial \ln L_m}{\partial \sigma} = \frac{-n}{\sigma} + \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma^3} \tag{15.17.12}$$

The value for σ at which L_m achieves maximum value is same in $\ln L_m$

$$\frac{\partial \ln L_m}{\partial \sigma} = 0 \tag{15.17.13}$$

$$\frac{-n}{\sigma} + \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma^3} = 0 \tag{15.17.14}$$

Upon simplifying the expression

$$\frac{n}{\sigma} = \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{\sigma^3} \tag{15.17.15}$$

$$\sigma^2 = \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{n} \tag{15.17.16}$$

Substituting (15.17.11) in (15.17.16)

$$\sigma^2 = \sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{n} \tag{15.17.17}$$

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{n}} \tag{15.17.18}$$

Hence **Option 3** and **Option 4** are correct

15.18. Suppose X_1 and X_2 are independent and identically distributed random variables each following an exponential distribution with mean θ, i.e., the common pdf is given by $f_\theta(x) = \frac{1}{\theta}e^{\frac{-x}{\theta}}, 0 < x < \infty, 0 < \theta < \infty$. Then which of the following is true? Conditional distribution of X_2 given $X_1 + X_2 = t$ is

a) exponential with mean $\frac{t}{2}$ and hence $X_1 + X_2$ is sufficient for θ
b) exponential with mean $\frac{t\theta}{2}$ and hence $X_1 + X_2$ is not sufficient for θ
c) uniform$(0, t)$ and hence $X_1 + X_2$ is sufficient for θ
d) uniform$(0, t\theta)$ and hence $X_1 + X_2$ is not sufficient for θ

Solution: Let $f_{X_1,X_2}(x_1, x_2)$ denote the joint probability distribution of random variables X_1 and X_2. Let Z be another random variable such that $Z = X_1 + X_2$. Let $\Phi_{X_1}(\omega)$ and $\Phi_Z(\omega)$ be the characteristic functions of the probability density functions $f_{X_1}(x)$ and $f_Z(x)$ respectively. The conditional probability density function of X_2 can be defined by:

$$f_{X_2|(X_1+X_2=t)}(x_2) = \begin{cases} \frac{f_{X_1,X_2}(x_1,x_2)}{f_{(X_1+X_2)}(t)} & \text{if } x_1 + x_2 = t \\ 0 & \text{otherwise} \end{cases} \tag{15.18.1}$$

$$x_1 + x_2 = t \tag{15.18.2}$$

$$0 < x_1, x_2 < \infty \tag{15.18.3}$$

$$x_1 = t - x_2 \tag{15.18.4}$$

$$t - x_2 > 0 \tag{15.18.5}$$

$$x_2 < t \tag{15.18.6}$$

From equations (15.18.3) and (15.18.6), we can conclude that $x_2 \in (0, t)$ if $x_1 + x_2 = t$. Also, given in the question,

$$0 < \theta < \infty \tag{15.18.7}$$

$$f_{X_1}(x_1) = \frac{1}{\theta}e^{\frac{-x_1}{\theta}}, 0 < x_1 < \infty \tag{15.18.8}$$

$$f_{X_2}(x_2) = \frac{1}{\theta}e^{\frac{-x_2}{\theta}}, 0 < x_2 < \infty \tag{15.18.9}$$

Since X_1 and X_2 are independent,

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) \times f_{X_2}(x_2) \tag{15.18.10}$$

$$= \frac{1}{\theta}e^{\frac{-x_1}{\theta}} \times \frac{1}{\theta}e^{\frac{-x_2}{\theta}} \tag{15.18.11}$$

$$= \frac{1}{\theta^2}e^{\frac{-(x_1+x_2)}{\theta}} \tag{15.18.12}$$

$$\Phi_{X_1}(\omega) = \frac{1}{\theta}\int_0^\infty e^{i\omega x}e^{\frac{-x}{\theta}}\,dx \tag{15.18.13}$$

$$= \frac{1}{\theta} \times \frac{1}{i\omega - \frac{1}{\theta}}\left(e^{x(i\omega-\frac{1}{\theta})}\right)\Big|_0^\infty \tag{15.18.14}$$

$$= \frac{1}{1 - i\omega\theta} - \frac{\lim_{x\to\infty}\left(e^{x(i\omega-\frac{1}{\theta})}\right)}{1 - i\omega\theta} \tag{15.18.15}$$

$$= \frac{1}{1 - i\omega\theta} - 0 = \frac{1}{1 - i\omega\theta} \tag{15.18.16}$$

$$\Phi_Z(\omega) = \left(\frac{1}{1 - i\omega\theta}\right)^2 \tag{15.18.17}$$

$$f_Z(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-i\omega x}}{\left(\frac{1}{1-i\omega\theta}\right)^2}\,d\omega \tag{15.18.18}$$

The equation (15.18.18) is the characteristic function expression of a gamma random variable with k=2. Thus,

$$f_Z(x) = \frac{x^{k-1}e^{\frac{-x}{\theta}}}{\Gamma(k)\theta^k} \tag{15.18.19}$$

$$= \frac{x^{2-1}e^{\frac{-x}{\theta}}}{\Gamma(2)\theta^2} \tag{15.18.20}$$

$$= \frac{xe^{\frac{-x}{\theta}}}{\theta^2} \tag{15.18.21}$$

$$f_{X_2|(X_1+X_2=t)}(x_2) = \begin{cases} \frac{f_{X_1,X_2}(x_1,x_2)}{f_Z(t)} & x_2 \in (0, t) \\ 0 & \text{otherwise} \end{cases} \tag{15.18.22}$$

Let $x_2 \in (0, t)$.

$$f_{X_2|(X_1+X_2=t)}(x_2) = \frac{f_{X_1,X_2}(x_1, x_2)}{f_Z(t)} \tag{15.18.23}$$
$$= \frac{\frac{1}{\theta^2}e^{\frac{-(x_1+x_2)}{\theta}}}{\frac{1}{\theta^2}e^{\frac{-t}{\theta}}t} \tag{15.18.24}$$
$$= \frac{e^{\frac{-(t)}{\theta}}}{e^{\frac{-t}{\theta}}t} \tag{15.18.25}$$
$$= \frac{1}{t} \quad \forall x_2 \in (0, t) \tag{15.18.26}$$

The obtained pdf is uniform$(0, t)$. Any distribution is sufficient for underlying parameter θ if the conditional probability distribution of the data does not depend on the parameter θ. And since the conditional distribution of X_2 does not depend on θ for any value of t, $X_1 + X_2$ is sufficient for θ. Verifying the pdf,

$$\text{total probability} = \int_0^t f_{X_2|(X_1+X_2=t)}(x_2)\, dx_2 \tag{15.18.27}$$
$$= \int_0^t \frac{1}{t}\, dx_2 \tag{15.18.28}$$
$$= 1 \tag{15.18.29}$$

Hence, the correct answer is option (15.18c)

15.19. Let $X_1, X_2, X_3, ..., X_n$ be i.i.d observations from a distribution with continuous probability density function f which is symmetric around θ i.e

$$f(x-\theta) = f(\theta - x) \tag{15.19.1}$$

for all real x.Consider the test $H_0 : \theta = 0$ vs $H_1 : \theta > 0$ and the sign test statistic

$$S_n = \sum_{i=1}^{n} sign(X_i) \tag{15.19.2}$$

where

$$sign(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \tag{15.19.3}$$

Let z_α be the upper $100(1-\alpha)^{th}$ percentile of the standard normal distribution where $0 < \alpha < 1$. Which of the following is/are correct?

a) If $\theta = 0$ then $\lim_{n\to\infty} P\left\{S_n > \sqrt{n}z_\alpha\right\} = 1$

b) If $\theta = 0$ then $\lim_{n\to\infty} P\left\{S_n > \sqrt{n}z_\alpha\right\} = \alpha$

c) If $\theta > 0$ then $\lim_{n\to\infty} P\left\{S_n > \sqrt{n}z_\alpha\right\} = 1$

d) If $\theta > 0$ then $\lim_{n\to\infty} P\left\{S_n > \sqrt{n}z_\alpha\right\} = \alpha$

Solution: $H_0 : \theta = 0$ Assume hypothesis $H_0 : \theta = 0$ is true.

a) Given X is symmetric around zero.

$$f_X(x) = f_X(-x) \tag{15.19.4}$$
$$\int_0^\infty f_X(x)dx = \int_0^\infty f_X(-x)dx \tag{15.19.5}$$

i) Solving LHS of (15.19.5)

$$\int_0^\infty f_X(x)dx = \Pr(X \geq 0) \tag{15.19.6}$$

ii) Solving RHS of (15.19.5)

$$\int_0^\infty f_X(-x)dx \tag{15.19.7}$$

Changing $-x \to x$ we get

$$\int_0^\infty f_X(-x)dx = \int_{-\infty}^0 f_X(x)dx \tag{15.19.8}$$
$$= \Pr(X \leq 0) \tag{15.19.9}$$

but

$$\int_{-\infty}^0 f_X(x)dx + \int_0^\infty f_X(x)dx = 1 \tag{15.19.10}$$

from (15.19.5) , (15.19.8) and (15.19.10)

$$\int_{-\infty}^0 f_X(x)dx = \int_0^\infty f_X(x)dx = \frac{1}{2} \tag{15.19.11}$$

$$\implies \Pr(X \leq 0) = \Pr(X \geq 0) = \frac{1}{2} \tag{15.19.12}$$

b) Let Y be a random variable such that

$$Y = sign(X) \tag{15.19.13}$$

$$Y = \begin{cases} 1 & X > 0 \\ -1 & X < 0 \end{cases} \tag{15.19.14}$$

From (15.19.12) and (15.19.14) we have

$$\Pr(Y = -1) = \Pr(Y = 1) = \frac{1}{2} \tag{15.19.15}$$

So $Y = sign(X)$ is also symmetric around zero.

$$\implies \mu_y = 0 \tag{15.19.16}$$

and variance is

$$\sigma_y^2 = (-1)^2\left(\frac{1}{2}\right) + (1)^2\left(\frac{1}{2}\right) \tag{15.19.17}$$

$$= 1 \tag{15.19.18}$$

c) Given

$$S_n = \sum_{i=1}^{n} sign(X_i) \tag{15.19.19}$$

$$S_n(\theta = 0) = \sum_{i=1}^{n} Y_i \tag{15.19.20}$$

From central limit theorem

$$Z = \lim_{n\to\infty} \sqrt{n}\left(\frac{\frac{S_n}{n} - \mu_y}{\sigma_y}\right) \tag{15.19.21}$$

$$= \lim_{n\to\infty} \sqrt{n}\left(\frac{S_n}{n}\right) \tag{15.19.22}$$

$$= \lim_{n\to\infty}\left(\frac{S_n}{\sqrt{n}}\right) \tag{15.19.23}$$

where Z is a standard normal variable N(0,1).

i) Given

$$\alpha = P\{Z > z_\alpha\} \tag{15.19.24}$$

So from (15.19.23) and (15.19.24)

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} > z_\alpha\right\} = \alpha \tag{15.19.25}$$

$$\implies \lim_{n\to\infty} P\left\{S_n > \sqrt{n}z_\alpha\right\} = \alpha \tag{15.19.26}$$

$H_1 : \theta > 0$ is true

a) Given X is symmetric around $\theta > 0$.Let us assume $\theta = \theta_0 > 0$.

$$f_X(\theta_0 - x) = f_X(\theta_0 + x) \tag{15.19.27}$$

$$\int_{\theta_0}^{\infty} f_X(\theta_0 - x)dx = \int_{\theta_0}^{\infty} f_X(\theta_0 + x)dx \tag{15.19.28}$$

i) Solving LHS of (15.19.28).Changinng $(\theta_0 - x) \to t$

$$\int_{\theta_0}^{\infty} f_X(\theta_0 - x)dx = \int_{-\infty}^{0} f_X(t)dt \tag{15.19.29}$$

$$= \Pr(X \leq 0) \tag{15.19.30}$$

ii) Solving RHS of (15.19.28).Changing $(\theta_0 + x) \to t$

$$\int_{\theta_0}^{\infty} f_X(\theta_0 + x)dx = \int_{2\theta_0}^{\infty} f_X(t)dt \tag{15.19.31}$$

$$= \int_{0}^{\infty} f_X(t)dt - \int_{0}^{2\theta_0} f_X(t)dt \tag{15.19.32}$$

$$= \Pr(X \geq 0) - k \tag{15.19.33}$$

where

$$k = \int_{0}^{2\theta_0} f_X(t)dt > 0 \tag{15.19.34}$$

From (15.19.28),(15.19.14) and (15.19.33)

$$\Pr(X \geq 0) > \Pr(X \leq 0) \tag{15.19.35}$$

b) So

$$\Pr(Y = 1) > \Pr(Y = -1) \tag{15.19.36}$$

Therefore,if we perform the experiment and find the value of $\left(\frac{S_n}{\sqrt{n}}\right)$,it is most likely to occur on the right side of the distribution of $\left(\frac{S_n}{\sqrt{n}}\right)$.In (15.19.23) it is shown that distrubution of the random variable $\left(\frac{S_n}{\sqrt{n}}\right)$ is $N(0,1)$ when n is very large.So

$$\lim_{n\to\infty} P\left\{\frac{S_n}{\sqrt{n}} > Z_\alpha\right\} = 1 \tag{15.19.37}$$