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),

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

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)

$$= \Pr(X_1 = X_2) \quad (2.1.1)$$
$$= \sum_{i=1}^{3} \Pr(X_2 = i|X_1 = i)\Pr(X_1 = i) \quad (2.1.2)$$
$$= \frac{6}{90} + \frac{12}{90} + \frac{6}{90} \quad (2.1.3)$$
$$= \frac{4}{15} \quad (2.1.4)$$

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

$$\frac{\text{(placing a B at last)} \times \text{(arranging other letters)}}{\text{arranging 100 letters}}$$
$$= \frac{60 \times 99!}{100!} = \frac{3}{5} \quad (2.2.1)$$

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

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

TABLE 2.3.1: Description

information. The desired probability is

$$\Pr(X=0, Y=0) = \Pr(X=0 \mid Y=0)\Pr(Y=0) \quad (2.3.1)$$
$$= .12 \quad (2.3.2)$$

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

3 Transformation of Variables

3.1. Let X be a random variable with pdf

$$f_X(x) = \begin{cases} \frac{2x}{\pi^2} & 0 < x < \pi \\ 0 & \text{otherwise} \end{cases} \quad (3.1.1)$$

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,

$$F_X(x) = \Pr(X \le x) \quad (3.1.2)$$
$$= \begin{cases} 0 & x \le 0 \\ \frac{x^2}{\pi^2} & 0 < x < \pi \\ 1 & x \ge \pi \end{cases} \quad (3.1.3)$$

after integration. Consequently,

$$F_Y(y) = \Pr(Y \le y) \quad (3.1.4)$$
$$= \Pr(\sin X \le y) \quad (3.1.5)$$

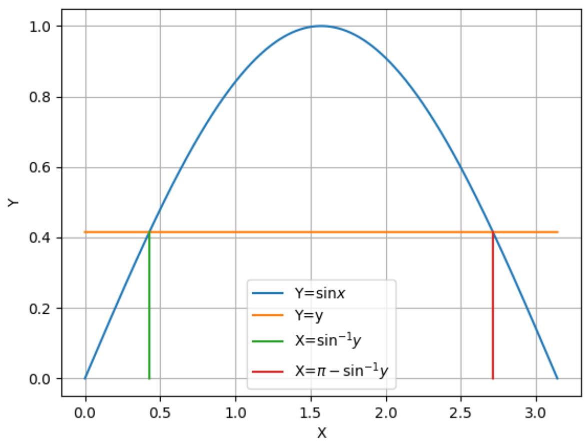

Fig. 3.1.1: $Y = \sin X$ plot

From Fig. 3.1.1,

$$\sin X \leq y$$
$$\implies \left\{X \leq \sin^{-1} y \cup X \geq \pi - \sin^{-1} y\right\} \quad (3.1.6)$$

$$\implies F_Y(y) = \Pr\left(X \leq \sin^{-1} y\right) + \Pr\left(X \geq \pi - \sin^{-1} y\right) \quad (3.1.7)$$
$$= F_X(\sin^{-1} y) + 1 - \Pr\left(X \leq \pi - \sin^{-1} y\right) \quad (3.1.8)$$
$$\implies F_Y(y) = 1 + F_X(\sin^{-1} y) - F_X(\pi - \sin^{-1} y) \quad (3.1.9)$$

Substituting from (3.1.3) in (3.1.9)

$$F_Y(y) = \frac{\left(\sin^{-1} y\right)^2}{\pi^2} + 1 - \frac{\left(\pi - \sin^{-1} y\right)^2}{\pi^2} \quad (3.1.10)$$
$$= \frac{2\sin^{-1} y}{\pi} \quad (3.1.11)$$

$$\therefore f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} \quad (3.1.12)$$
$$= \frac{2}{\pi\sqrt{1-y^2}} \quad (3.1.13)$$

Hence, option(D) is correct.

4 Independent Random Variables

4.1. 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:

4.2. 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:

4.3. 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 4.1. *If $X \sim \mathcal{N}(0, 1)$ then $Y = -X$ also follows standard normal distribution.*

Proof.

$$P(Y \leq u) = P(-X \leq u) \quad (4.3.1)$$
$$= P(X > -u) \quad (4.3.2)$$
$$= 1 - P(X \leq -u) \quad (4.3.3)$$
$$= 1 - (1 - P(X \leq u) \quad (4.3.4)$$
$$= P(X \leq u) \quad (4.3.5)$$

As the distribution is symmetric,

$$P(X \leq -u) = P(X \geq u) = 1 - P(X \leq u) \quad (4.3.6)$$

□

Lemma 4.2. *If n is an even number and $g(x)$ is an odd function, then,* *a)*

$$\Pr\left(g(X_1) > \sum_{k=2}^{n} g(X_k)\right) = \Pr\left(g(X_1) < \sum_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \quad (4.3.7)$$

b)

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \Pr\left(g(X_1) < \prod_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \quad (4.3.8)$$

Proof. a)

$$\Pr\left(g(X_1) > \sum_{k=2}^{n} g(X_k)\right) = \Pr\left(g(-X_1) < \sum_{k=2}^{n} g(-X_k)\right) = \Pr\left(g(X_1) < \sum_{k=2}^{n} g(X_k)\right) \quad (4.3.9)$$

As the cases

$$g(X_1) > \sum_{k=2}^{n} g(X_k) \quad (4.3.10)$$

and

$$g(X_1) < \sum_{k=2}^{n} g(X_k) \quad (4.3.11)$$

are complementary to each other,

$$\Pr\left(g(X_1) > \sum_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \quad (4.3.12)$$

b) Similarly,

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \Pr\left(g(-X_1) < \prod_{k=2}^{n} g(-X_k)\right) = \Pr\left(g(X_1) < \prod_{k=2}^{n} g(X_k)\right) \quad (4.3.13)$$

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

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \Pr\left(g(X_1) < \prod_{k=2}^{n} g(X_k)\right) \quad (4.3.14)$$

As the cases

$$g(X_1) > \prod_{k=2}^{n} g(X_k) \quad (4.3.15)$$

and

$$g(X_1) < \prod_{k=2}^{n} g(X_k) \quad (4.3.16)$$

are complementary to each other and from (4.3.7) we have

$$\Pr\left(g(X_1) > \prod_{k=2}^{n} g(X_k)\right) = \frac{1}{2} \quad (4.3.17)$$

(A) From (4.3.12) , taking $g(x) = x$,

$$\Pr(X_1 > X_2 + \ldots + X_{10}) = \frac{1}{2} \quad (4.3.18)$$

(B) From (4.3.17) taking $g(x) = x$

$$\Pr(X_1 > X_2 X_3 \ldots X_{10}) = \frac{1}{2} \quad (4.3.19)$$

(C) From (4.3.12) taking $g(x) = \sin x$

$$\Pr(\sin X_1 > \sin X_2 + \ldots + \sin X_{10}) = \frac{1}{2} \quad (4.3.20)$$

(D)

$$\Pr(\sin X_1 > \sin(X_2 + \ldots + X_{10})) = \Pr(\sin(-X_1) < \sin(-X_2 - \ldots - X_{10})) = \Pr(\sin X_1 < \sin(X_2 + \ldots + X_{10})) \quad (4.3.21)$$

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

$$\Pr(\sin X_1 > \sin(X_2 + \ldots + X_{10})) = \Pr(\sin X_1 < \sin(X_2 + \ldots + X_{10})) \quad (4.3.22)$$

Also, as X_1 is a continuous random variable

$$\Pr(\sin X_1 = \sin(X_2 + \ldots + X_{10})) = 0 \quad (4.3.23)$$

As the cases

$$X_1 > X_2 + ... + X_{10} \tag{4.3.24}$$

and

$$X_1 < X_2 + ... + X_{10} \tag{4.3.25}$$

are complementary to each other

$$\Pr(\sin X_1 > \sin(X_2 + ... + X_{10})) = \frac{1}{2} \tag{4.3.26}$$

□

4.4. 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*

$$\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B) \tag{4.4.1}$$

Alternatively,

$$F_{X,Y}(a,b) = F_X(a)F_Y(b) \tag{4.4.2}$$

Lemma 4.3. *From* (4.4.2)*, it follows that*

$$\implies f_{X,Y}(a,b) = f_X(a)f_Y(b) \tag{4.4.3}$$

Proof. From (4.4.2),

$$\frac{\partial^2 F_{X,Y}(a,b)}{\partial b \partial a} = \frac{\partial F_X(a)}{\partial a}\frac{\partial F_Y(b)}{\partial b} \tag{4.4.4}$$

yielding (4.4.3). □

a) From the given information,

$$\Pr(X > a, Y > a) = \Pr(X > a)\Pr(Y > a) \tag{4.4.5}$$

$$= [1 - F_X(a)]\ [1 - F_Y(a)] \tag{4.4.6}$$

$$\begin{aligned}
&\because \Pr(X > a) - \Pr(Y < a)\\
&= \Pr(X > a, Y > a) + \Pr(X > a, Y < a)\\
&- \Pr(X > a, Y < a) - \Pr(X < a, Y < a),\\
&\Pr(X > a, Y > a) = 1 - F_X(a) - F_Y(a)\\
&\qquad + F_{X,Y}(a,a)
\end{aligned} \tag{4.4.7}$$

which, upon substituting from (4.4.6) yields

$$\implies F_{X,Y}(a,a) = F_X(a)F_Y(a) \tag{4.4.8}$$

which is a special case of (4.4.1) 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,

$$\Pr(X > a|Y < b) = \Pr(X > a) \tag{4.4.9}$$

$$\begin{aligned}
\implies &\Pr(X > a, Y < b)\\
&= \Pr(X > a)\Pr(Y < b)
\end{aligned} \tag{4.4.10}$$

for all $a, b \in R$.

$$\begin{aligned}
\because F_Y(b) &= \Pr(X > a, Y < b)\\
&+ \Pr(X < a, Y < b),
\end{aligned} \tag{4.4.11}$$

$$\begin{aligned}
&\Pr(X > a, Y < b)\\
&= F_Y(b) - F_{X,Y}(a,b)\\
&\implies F_Y(b) - F_{X,Y}(a,b)\\
&= (1 - F_X(a))F_Y(b)\\
&\text{or, } F_{X,Y}(a,b) = F_X(a)F_Y(b)
\end{aligned} \tag{4.4.12}$$

upon substituting from (4.4.10) 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.,*

$$cov\left(X,Y\right)=E\left(XY\right)-E\left(X\right)E\left(Y\right)=0 \tag{4.4.13}$$

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

$$f_X\left(x\right)=\begin{cases}\frac{1}{2} & -1\le x\le 1\\ 0 & otherwise\end{cases} \tag{4.4.14}$$

$$\mathrm{E}\left(X\right)=\int_{-1}^{1}xf\left(x\right)dx=0 \tag{4.4.15}$$

Let

$$Y=X^2. \tag{4.4.16}$$

so that X and Y are dependent. Then,

$$\begin{aligned}\mathrm{cov}\left(X,Y\right)&=\mathrm{E}\left(XY\right)-\mathrm{E}\left(X\right)\mathrm{E}\left(Y\right) && (4.4.17)\\ &=\mathrm{E}\left(X^3\right)-0\times\mathrm{E}\left(Y\right) && (4.4.18)\\ &=\int_{-1}^{1}x^3f\left(x\right)dx=0 && (4.4.19)\end{aligned}$$

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

d) Given that,

$$\mathrm{E}\left(\left(X-a\right)\left(Y-b\right)\right)=\mathrm{E}\left(X-a\right)\mathrm{E}\left(Y-b\right) \tag{4.4.20}$$

$$\begin{aligned}\implies \mathrm{cov}\left(X-a,Y-b\right)=&\\ \mathrm{E}\left(\left(X-a\right)\left(Y-b\right)\right)&\\ -\mathrm{E}\left(X-a\right)\mathrm{E}\left(Y-b\right)& && (4.4.21)\end{aligned}$$

$$\text{or, }\mathrm{cov}\left(X-a,Y-b\right)=0=\mathrm{cov}\left(X,Y\right) \tag{4.4.22}$$

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

5 Binomial Distribution

5.1. 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

$$\begin{aligned}\Pr\left(X\ge 2\right)&=1-\Pr\left(X<2\right) && (5.1.1)\\ &=1-\Pr\left(X=0\right)-\Pr\left(X=1\right) && (5.1.2)\\ &=1-{}^{15}C_0p^0q^{15}-{}^{15}C_1p^1q^{14} && (5.1.3)\\ &=0.1709 && (5.1.4)\end{aligned}$$

where

$$p=0.0.5,q=1-p=0.95 \tag{5.1.5}$$

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

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

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

Solution:

6 Poisson Distribution

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

$$P(X=k)=\begin{cases}\frac{e^{-\lambda}\lambda^k}{k!}, & k=0,1,2,...;\lambda>0\\ 0 & \text{otherwise}\end{cases} \tag{6.1.1}$$

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:

$$\begin{aligned}Y&=X^2+3 && (6.1.2)\\ \implies X&=\sqrt{Y-3} && (6.1.3)\end{aligned}$$

Substituting $k=\sqrt{y-3}$ in (6.1.1),

$$p_Y(y)=\begin{cases}\frac{e^{-\lambda}\lambda^{\sqrt{y-3}}}{\sqrt{(y-3)!}}, & y=3,4,7,12,...\\ 0 & \text{otherwise}\end{cases} \tag{6.1.4}$$

Hence, the correct option is (A).

7 Gaussian Distribution

7.1. 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,

$$U = \mathcal{N}\left(0, \frac{1}{4}\right) V \qquad = \mathcal{N}\left(0, \frac{1}{9}\right) \tag{7.1.1}$$

Let $Y = 3V - 2U$. Then,

$$\mathrm{E}(Y) = 3\mathrm{E}(V) - 2\mathrm{E}(U) = 0 \tag{7.1.2}$$

$$\mathrm{var}(Y) = 3^2\mathrm{var}(V) + 2^2\mathrm{var}(U) = 2 \tag{7.1.3}$$

$$\therefore Y = \mathcal{N}(0, 2) \tag{7.1.4}$$

Thus,

$$\Pr(3V \geq 2U) = \Pr(3V - 2U \geq 0) \tag{7.1.5}$$

$$= \Pr(Y \geq 0) = \frac{1}{2} \tag{7.1.6}$$

$\because$ Y is symmetric about the origin.

7.2. (X, Y) follows bivariate normal distribution $N_2(0,0,1,1,\rho)$, -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

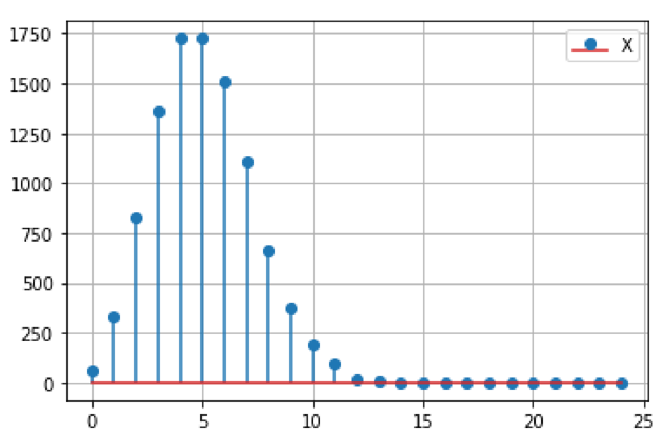

Fig. 6.1.1: Poisson stem plot for X ($\lambda = 5$)

Solution: Given that

$$\mathbf{M} = \begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{7.2.1}$$

where

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{7.2.2}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{7.2.3}$$

Also,

$$X + Y = \mathbf{A}^\top\mathbf{M} \tag{7.2.4}$$

$$X - Y = \mathbf{B}^\top\mathbf{M} \tag{7.2.5}$$

where

$$\mathbf{A} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{B} \qquad = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{7.2.6}$$

Thus,

$$Cov(X + Y, X - Y) = \mathbf{A}^\top\boldsymbol{\Sigma}\mathbf{B} \tag{7.2.7}$$

$$= 0 \tag{7.2.8}$$

$\therefore$ X+Y and X-Y are uncorrelated irrespective of value of ρ where $\rho \in (-1, 1)$.

8 Geometric Distribution

8.1. Suppose X has density

$$f(x|\theta) = \frac{1}{\theta}e^{-x/\theta}, x > 0 \tag{8.1.1}$$

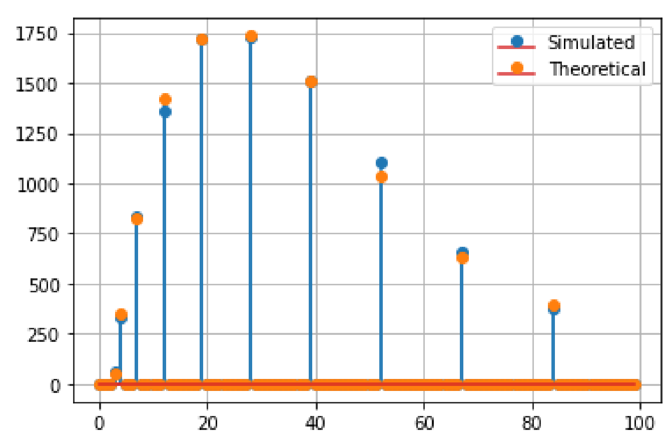

Fig. 6.1.2: Stem plot for Y (Simulated and Theoretical) ($\lambda = 5$)

Define

$$Y = k, \quad k \le X < k+1, \quad k = 0, 1, 2 \dots \tag{8.1.2}$$

Then the distribution of Y is

a) Normal
b) Binomial
c) Poisson
d) Geometric

Solution:

$$\Pr(Y = k) = \Pr(k \le X < k+1) \tag{8.1.3}$$
$$= \int_k^{k+1} f(x|\theta)\, dx \tag{8.1.4}$$
$$= \int_k^{k+1} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \tag{8.1.5}$$
$$= \left[-e^{-\frac{x}{\theta}}\right]_k^{k+1} \tag{8.1.6}$$
$$= e^{-\frac{k}{\theta}}\left(1 - e^{-\frac{1}{\theta}}\right) \tag{8.1.7}$$
$$\implies \Pr(Y = k) = (1-p)^k\, p\, k = 0, 1, 2 \dots \tag{8.1.8}$$

where

$$p = 1 - e^{-\frac{1}{\theta}} \tag{8.1.9}$$

Therefore, the distribution of Y is 4) Geometric.

9 Two Dimensions

9.1. Let $c \in \mathbb{R}$ be a constant. Let X, Y be random variables with joint probability density function

$$f(x, y) = \begin{cases} cxy & 0 < x < y < 1, \\ 0 & \text{otherwise} \end{cases} \tag{9.1.1}$$

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,

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \tag{9.1.2}$$
$$= \int_0^y cxy\, dx \tag{9.1.3}$$
$$= cy\left(\frac{x^2}{2}\right)\Bigg|_0^y \tag{9.1.4}$$
$$= \frac{cy^3}{2} \tag{9.1.5}$$

$$\implies f_Y(y) = \begin{cases} \frac{cy^3}{2}, & 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{9.1.6}$$

$\because$ the area under the pdf is 1, from (9.1.6),

$$\implies \int_{-\infty}^{\infty} f_Y(y)\, dy = 1 \tag{9.1.7}$$
$$\implies \int_0^1 c\frac{y^3}{2} = 1 \tag{9.1.8}$$
$$\implies \frac{c}{8} = 1 \tag{9.1.9}$$
$$\text{or, } c = 8 \tag{9.1.10}$$

Also,

$$f_Y(y) = \begin{cases} 4y^3 & \text{, if } 0 < y < 1 \\ 0 & \text{, otherwise} \end{cases} \tag{9.1.11}$$

c)

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \tag{9.1.12}$$
$$= \int_x^1 cxy\, dy \tag{9.1.13}$$
$$= cx\left(\frac{y^2}{2}\right)\Bigg|_x^1 \tag{9.1.14}$$
$$= cx\left(\frac{1-x^2}{2}\right) \tag{9.1.15}$$

$$\implies f_X(x) = \begin{cases} 4x\left(1-x^2\right), & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{9.1.16}$$

From (9.1.16) and (9.1.11)

$$f_X(x) \times f_Y(y) = \begin{cases} 16xy^3\left(1 - x^2\right) & \text{, if } 0 < x, y < 1 \\ 0 & \text{, otherwise} \end{cases} \tag{9.1.17}$$

$$\neq f(x, y) \tag{9.1.18}$$

Hence, X and Y are not independent.

d)

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx \tag{9.1.19}$$

$$= \int_0^x 4x(1 - x^2)\, dx \tag{9.1.20}$$

$$= \int_0^x 4x - 4x^3\, dx \tag{9.1.21}$$

$$= 2x^2 - 4x^4 \text{ for } 0 < x < 1 \tag{9.1.22}$$

yielding

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 2x^2 - 4x^4 & 0 < x < 1 \\ 1 & x \geq 1 \end{cases} \tag{9.1.23}$$

From (9.1.23),

$$\Pr(Y - \epsilon < X < Y + \epsilon)$$
$$= F_X(Y + \epsilon) - F_X(Y - \epsilon)$$
$$= 8\epsilon Y\left(1 - Y^2 - \epsilon^2\right) \tag{9.1.24}$$

upon simplification. Letting

$$g(Y) = 8\epsilon Y(1 - Y^2 - \epsilon^2), \tag{9.1.25}$$

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y) f_Y(y)\, dy \tag{9.1.26}$$

$$= \int_0^1 (4y^3)(8\epsilon y)(1 - y^2 - \epsilon^2)\, dy \tag{9.1.27}$$

$$\implies \Pr(Y - \epsilon < X < Y + \epsilon)$$
$$= 32\epsilon\left(\frac{2 - 7\epsilon^2}{35}\right) \tag{9.1.28}$$

Now substituting $\epsilon = 0$ in the above,

$$\Pr(X = Y) = 0 \tag{9.1.29}$$

9.2. Let X and Y be random variables having the joining probability density function

$$f_{XY}(x, y) = \begin{cases} \frac{1}{\sqrt{2\pi y}} e^{\frac{-1}{2y}(x-y)^2} & x \in (-\infty, \infty), \\ & y \in (0, 1) \\ 0 & \text{otherwise} \end{cases} \tag{9.2.1}$$

The covariance between the random variables X and Y is

Solution:

9.3. Let a random variable X follow exponential distribution with mean 2. Define $Y = [X-2|X > 2]$. The value of $\Pr(Y \geq t)$ is ...

Solution: From the given information,

$$\Pr(Y \geq t) = \frac{\Pr(X - 2 \geq t, X > 2)}{\Pr(X > 2)} \tag{9.3.1}$$

$$= \frac{\Pr(X \geq t + 2, X > 2)}{\Pr(X > 2)} \tag{9.3.2}$$

$\because$ X has an exponential distribution with parameter $\lambda = \frac{1}{2}$,

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } 0 < x < \infty \\ 0, & \text{otherwise} \end{cases} \tag{9.3.3}$$

and

$$\Pr(X > 2) = 1 - F_X(2) = e^{-2\lambda} \tag{9.3.4}$$

Also,

$$\Pr(X \geq t + 2, X > 2) = \begin{cases} \Pr(X \geq t + 2) & t \geq 0 \\ \Pr(X > 2) & t < 0 \end{cases} \tag{9.3.5}$$

Substituting (9.3.5) in (9.3.2), using (9.3.4) and simplifying,

$$\Pr(Y \geq t) = \begin{cases} e^{-\frac{t}{2}} & t \geq 0 \\ 1 & t < 0 \end{cases} \tag{9.3.6}$$

10 Markov Chain

10.1. **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)...... **So-**

lution: The given problem can be represented

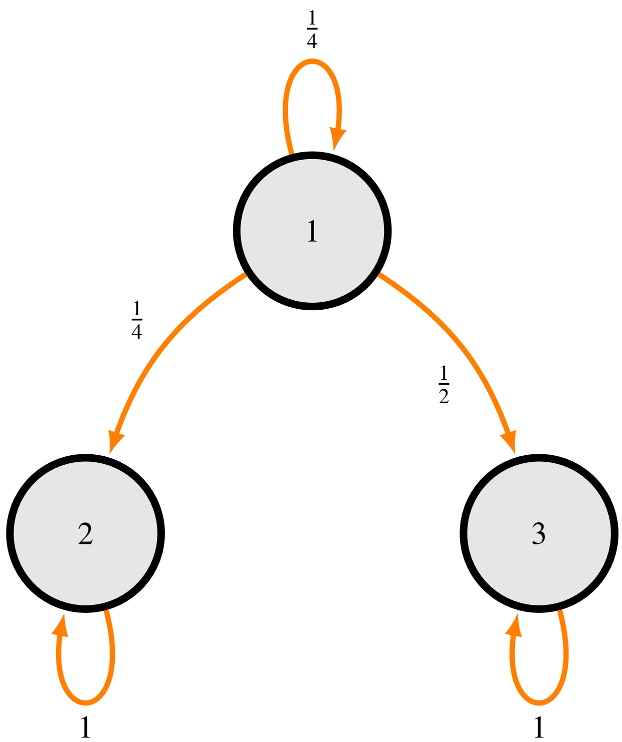

Fig. 10.1.1

using Table 10.1.1 and the Markov chain in Fig. 10.1.1. The state transition matrix for the

State	Description
1	$\{T, T\}$
2	$Y = \{T, H\}$
3	$N = \{\{H, H\}, \{H, T\}\}$

TABLE 10.1.1: States and their notations

Markov chain can be expressed as

$$P = \begin{array}{c c} & \begin{array}{ccc} 2 & 3 & 1 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 1 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} \end{array} \tag{10.1.1}$$

Clearly, the state 1 is transient, while 2, 3 are absorbing. Comparing (10.1.1) with the standard form of the state transition matrix

$$P = \begin{array}{c c} & \begin{array}{cc} A & N \end{array} \\ \begin{array}{c} A \\ N \end{array} & \begin{bmatrix} I & O \\ R & Q \end{bmatrix} \end{array} \tag{10.1.2}$$

TABLE 10.1.2: 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

where, From (10.1.1) and (10.1.2),

$$R = \begin{pmatrix} 0.25 & 0.5 \end{pmatrix}, Q = \begin{pmatrix} 0.25 \end{pmatrix} \tag{10.1.3}$$

The limiting matrix for absorbing Markov chain is

$$\bar{P} = \begin{pmatrix} I & O \\ FR & O \end{pmatrix} \tag{10.1.4}$$

where

$$F = (I - Q)^{-1} = (1 - 0.25)^{-1} = \frac{4}{3} \tag{10.1.5}$$

is called the fundamental matrix of P. Upon substituting from (10.1.3) in (10.1.5),

$$F = (1 - 0.25)^{-1} = \frac{4}{3} \tag{10.1.6}$$

and

$$FR = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \end{pmatrix} \tag{10.1.7}$$

which, upon substituting in (10.1.4) yields

$$\bar{P} = \begin{array}{c c} & \begin{array}{ccc} 2 & 3 & 1 \end{array} \\ \begin{array}{c} 2 \\ 3 \\ 1 \end{array} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 \end{bmatrix} \end{array} \tag{10.1.8}$$

$$\therefore \bar{p}_{12} = \frac{1}{3} \tag{10.1.9}$$

11 Convergence

11.1. Let $X_1, X_2, \dots$ be independent and identically distributed random variables each following a uniform distribution on (0,1). Denote

$$T_n = \max\{X_1, X_2, \dots, X_n\}. \tag{11.1.1}$$

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 3. 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,

$$F_X(x) = F_{X_i}(x), i = 1, 2, \dots, n \forall x \in \mathbb{R} \tag{11.1.2}$$

Definition 4. 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,*

$$X_{(1)} = \min(X_1, X_2, \dots, X_n) \tag{11.1.3}$$
$$X_{(2)} = \text{the } 2^{nd} \text{ smallest of } X_1, X_2, \dots, X_n \tag{11.1.4}$$
$$\vdots \tag{11.1.5}$$
$$X_{(n)} = \max(X_1, X_2, \dots, X_n) \tag{11.1.6}$$

Lemma 11.1. Distribution of the maximum :

$$f_{T_n}(x) = \begin{cases} nx^{n-1}, & 0 < x < 1 \\ 0, & otherwise \end{cases} \tag{11.1.7}$$

$$F_{T_n}(x) = \begin{cases} x^n, & 0 < x < 1 \\ 1, & x \geq 1 \\ 0, & otherwise \end{cases} \tag{11.1.8}$$

Proof:

$$F_{X_{(n)}}(x) = \Pr(X_{(n)} \leq x) \tag{11.1.9}$$
$$= \Pr(X_1 \leq x, X_2 \leq x, \dots, X_n \leq x) \tag{11.1.10}$$
$$= \Pr(X_1 \leq x)\Pr(X_2 \leq x)\dots\Pr(X_n \leq x) \tag{11.1.11}$$
$$= [\Pr(X_1 \leq x)]^n \text{ (i.i.d)} \tag{11.1.12}$$
$$= [F_X(x)]^n \tag{11.1.13}$$

and

$$f_{X_{(n)}}(x) = \frac{d}{dx}\left(F_{X_{(n)}}(x)\right) = \frac{d}{dx}\left([F_X(x)]^n\right) \tag{11.1.14}$$
$$= n\left([F_X(x)]^{n-1}\right)\frac{d}{dx}(F_X(x)) \tag{11.1.15}$$
$$= n[F_X(x)]^{n-1} f_X(x) \tag{11.1.16}$$

$\because$

$$f_{X_i}(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & otherwise \end{cases}, \tag{11.1.17}$$

$$F_{X_i}(x) = \begin{cases} x, & 0 < x < 1 \\ 1, & x \geq 1 \\ 0, & otherwise, \end{cases} \tag{11.1.18}$$

$\forall i \in \mathbb{N}$. Substituting the above in (11.1.16) and (11.1.13) yields (11.1.7) and (11.1.8) respectively. Then, as $T_n = max\{X_1, X_2, \dots, X_n\} = X_{(n)}$,

Lemma 11.2. *If $Y = aX + b$ and $a < 0$, then*

$$F_Y(y) = 1 - F_X\left(\frac{y-b}{a}\right) \tag{11.1.19}$$

Definition 5. 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*

$$\lim_{n\to\infty} \Pr(|X_n - X| \geq \epsilon) = 0, \forall \epsilon > 0 \tag{11.1.20}$$

Definition 6. 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*

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \tag{11.1.21}$$

for all x at which $F_X(x)$ is continuous.

a)

$$\lim_{n\to\infty} \Pr(|T_n - 1| \geq \epsilon) = \lim_{n\to\infty} \Pr(1 - T_n \geq \epsilon)$$
$$= \lim_{n\to\infty} \Pr(T_n \leq 1 - \epsilon) = \lim_{n\to\infty} F_{T_n}(1-\epsilon) \tag{11.1.22}$$

$$\because F_{T_n}(1-\epsilon) = \begin{cases} (1-\epsilon)^n, & 0 < \epsilon < 1 \\ 0, & \epsilon \geq 1 \end{cases} \tag{11.1.23}$$

and

$$\because \lim_{n\to\infty}(1-\epsilon)^n = 0 \text{ for } 0 < \epsilon < 1 \tag{11.1.24}$$

$$\tag{11.1.25}$$

from (11.1.24), (11.1.23) and (11.1.22),

$$\lim_{n\to\infty} \Pr\left(|T_n - 1| \geq \epsilon\right) = 0, \forall \epsilon > 0 \tag{11.1.26}$$

$\therefore T_n$ converges to 1 in probability.

b) Substituting $a = -n, b = n$ in (11.1.19),

$$F_{n(1-T_n)}(x) = 1 - F_{T_n}\left(1 - \frac{x}{n}\right) \tag{11.1.27}$$

$$\tag{11.1.28}$$

where

$$F_{T_n}\left(1-\frac{x}{n}\right) = \begin{cases} \left(1-\frac{x}{n}\right)^n, & 0 < x < n \\ 1, & x \leq 0 \\ 0, & x \geq n \end{cases} \tag{11.1.29}$$

$$from (11.1.8) \tag{11.1.30}$$

$$\because \lim_{n\to\infty}\left(1-\frac{y}{n}\right)^n = e^{-y}, \tag{11.1.31}$$

$$\tag{11.1.32}$$

$$\therefore \lim_{n\to\infty} F_{T_n}\left(1-\frac{x}{n}\right) = \begin{cases} e^{-x}, & x > 0 \\ 1, & x \leq 0 \end{cases} \tag{11.1.33}$$

$$\implies \lim_{n\to\infty} F_{n(1-T_n)}(x) = 1 - \lim_{n\to\infty} F_{T_n}\left(1-\frac{x}{n}\right) \tag{11.1.34}$$

which can be expressed as

$$\therefore \lim_{n\to\infty} F_{n(1-T_n)}(x) = \begin{cases} 1-e^{-x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{11.1.35}$$

$\therefore n(1-T_n)$ converges in distribution to the random variable $X \sim Exponential(1)$.

c) Substituting $a = -n^2, b = n^2$ in (11.1.19),

$$F_{n^2(1-T_n)}(x) = 1 - F_{T_n}\left(1 - \frac{x}{n^2}\right) \tag{11.1.36}$$

$$F_{T_n}\left(1-\frac{x}{n^2}\right) = \begin{cases} \left(1-\frac{x}{n^2}\right)^n, & 0 < x < n^2 \\ 1, & x \leq 0 \\ 0, & x \geq n^2 \end{cases} \tag{11.1.37}$$

$$= \begin{cases} 1, & x > 0 \\ 1, & x \leq 0 \end{cases} \tag{11.1.38}$$

$$\because \lim_{n\to\infty}\left(1-\frac{x}{n^2}\right)^n = 1 \tag{11.1.39}$$

yielding

$$\lim_{n\to\infty} F_{n^2(1-T_n)}(x) = \begin{cases} 0, & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{11.1.40}$$

which is not a valid CDF. Hence, $n^2(1-T_n)$ does not converge in distribution.

d)

$$\begin{aligned} &\lim_{n\to\infty} \Pr\left(|\sqrt{n}(1-T_n) - 0| \geq \epsilon\right) \\ &= \lim_{n\to\infty} \Pr\left(1 - T_n \geq \frac{\epsilon}{\sqrt{n}}\right) \\ &= \lim_{n\to\infty} \Pr\left(T_n \leq 1 - \frac{\epsilon}{\sqrt{n}}\right) \\ &= \lim_{n\to\infty} F_{T_n}\left(1 - \frac{\epsilon}{\sqrt{n}}\right) \end{aligned}$$

$$= \begin{cases} \left(1-\frac{\epsilon}{\sqrt{n}}\right)^n, & 0 < \epsilon < \sqrt{n} \\ 0, & \epsilon \geq \sqrt{n} \end{cases} \tag{11.1.41}$$

$$\because \lim_{n\to\infty}\left(1-\frac{\epsilon}{\sqrt{n}}\right)^n = 0 \text{ for } 0 < \epsilon < \sqrt{n},$$

$$\lim_{n\to\infty} \Pr\left(|\sqrt{n}(1-T_n) - 0| \geq \epsilon\right) = 0, \forall \epsilon > 0 \tag{11.1.42}$$

$\therefore \sqrt{n}(1-T_n)$ converges to 0 in probability.

Hence, options 1), 2), 4) are correct.

11.2. Let $\{X_i\}_{i\geq 1}$ be a sequence of i.i.d. random variables with $E(X_i) = 0$ and $V(X_i) = 1$. 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:

Lemma 11.3. Strong Law of Large Numbers: *Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables, each having finite mean $E(X_i)$. Then for any $\epsilon > 0$,*

$$\lim_{n\to\infty} \Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E(X_i)\right| \geq \epsilon\right) = 0 \quad (11.2.1)$$

Or, $\frac{1}{n}\sum_{i=1}^{n} X_i$ converges in probability to $E(X_i)$.

11.3. 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

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (11.3.1)$$

$$S_n = \frac{\sum_{i=1}^{n} \frac{1}{i} X_i}{\sum_{i=1}^{n} \frac{1}{i}} \quad (11.3.2)$$

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 $i = 1, 2, \dots, n$ are independent random variables we can use this property to state

$$Var\left(\sum_{i=1}^{n} g(X_i)\right) = \sum_{i=1}^{n} Var(g(X_i)) \quad (11.3.3)$$

Definition 7. *Random Sample: The random variables $X_1, X_2, X_3, \dots, X_n$ 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

$$Var(X_1) = 1 \quad (11.3.4)$$
$$Var(X_2) = 2 \quad (11.3.5)$$
$$Var(X_1) \neq Var(X_2) \quad (11.3.6)$$

The equation (11.3.6) doesn't follow point(11.3d) in definition(7) and hence the random variables are not a random sample.

a) Expectation of $\bar{X}_n$ and S_n

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \quad (11.3.7)$$
$$= \frac{1}{n}\sum_{i=1}^{n} E(X_i) \quad (11.3.8)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \mu \quad (11.3.9)$$
$$= \mu \quad (11.3.10)$$
$$E(S_n) = E\left(\frac{\sum_{i=1}^{n} \frac{1}{i} X_i}{\sum_{i=1}^{n} \frac{1}{i}}\right) \quad (11.3.11)$$
$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{i}\right)}\sum_{i=1}^{n} E\left(\frac{1}{i} X_i\right) \quad (11.3.12)$$
$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{1}{i}\right)}\sum_{i=1}^{n} \frac{\mu}{i} \quad (11.3.13)$$
$$= \mu \quad (11.3.14)$$

From (11.3.10) and (11.3.14) we get option(11.3a) is correct.

b) Variance of $\bar{X}_n$ and S_n using (11.3.3)

$$Var(\bar{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \quad (11.3.15)$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} Var(X_i)\right) \quad (11.3.16)$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} i\right) \quad (11.3.17)$$
$$= \frac{1}{2} + \frac{1}{2n} \quad (11.3.18)$$
$$Var(S_n) = Var\left(\frac{\sum_{i=1}^{n} \frac{1}{i} X_i}{\sum_{i=1}^{n} \frac{1}{i}}\right) \quad (11.3.19)$$

$$= \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{i}\right)^2}\sum_{i=1}^{n}\frac{1}{i^2}Var(X_i) \quad (11.3.20)$$
$$= \frac{1}{\left(\sum_{i=1}^{n}\frac{1}{i}\right)^2}\sum_{i=1}^{n}\frac{1}{i^2}i \quad (11.3.21)$$
$$= \frac{1}{\sum_{i=1}^{n}\frac{1}{i}} \quad (11.3.22)$$

As n is sufficiently large

$$Var(\bar{X}_n) = \frac{1}{2} \quad (11.3.23)$$
$$Var(S_n) = 0 \quad (11.3.24)$$
$$Var(S_n) < Var(\bar{X}_n) \quad (11.3.25)$$

from (11.3.25) we get option(11.3b) as correct.

c)

Definition 8. *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.,*

$$\hat{\Theta} = h(X_1, X_2, \cdots, X_n) \quad (11.3.26)$$

Definition 9. *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*

$$\lim_{n\to\infty} P(|\hat{\Theta}_n - \theta| \geq \epsilon) = 0, \textit{ for all } \epsilon > 0. \quad (11.3.27)$$

From (11.3.6) as given data is not a random sample we don't define point estimator and hence option(11.3c) is incorrect.

d)

Definition 10. *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 11. *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 (11.3.6) suggests that given data is not a random sample we don't define statistic and hence option(11.3d) is incorrect.

Hence option(11.3a) and option(11.3b) are correct.

11.4. Let $X_1, X_2, \ldots$ be i.i.d. $N(0,1)$ random variables. Let

$$S_n = X_1^2 + X_2^2 + \cdots + X_n^2 \forall n \geq 1. \quad (11.4.1)$$

Which of the following statements are correct?

a)
$$\frac{S_n - n}{\sqrt{2}} \sim N(0,1) \quad \forall n \geq 1 \quad (11.4.2)$$

b)
$$\forall \epsilon > 0, \Pr\left(\left|\frac{S_n}{n} - 2\right| > \epsilon\right) \to 0, n \to \infty \quad (11.4.3)$$

c) $\frac{S_n}{n} \to 1$ with probability 1

d)
$$\Pr\left(S_n \leq n + \sqrt{n}x\right) \to \Pr\left(Y \leq x\right) \quad \forall x \in \mathbb{R}, Y \sim N(0,2) \quad (11.4.4)$$

12 Statistics

12.1. Let Y_1 denote the first order statistic in a random sample of size n from a distribution that has the pdf

$$f(x) = \begin{cases} e^{-(x-\theta)} & \text{when } \theta < x < \infty \\ 0 & \text{otherwise} \end{cases} \quad (12.1.1)$$

Obtain the distribution of $Z_n = n(Y_1 - \theta)$.
Solution: From the given information

$$Y_1 = \min\{X_1, X_2, ...X_n\} \quad (12.1.2)$$

and

$$F_{Z_n}(z) = \Pr\left(n(Y_1 - \theta) \leq z\right) \quad (12.1.3)$$
$$= \Pr\left(Y_1 \leq \frac{z}{n} + \theta\right) \quad (12.1.4)$$
$$= 1 - \Pr\left(Y_1 > \frac{z}{n} + \theta\right) \quad (12.1.5)$$

Let

$$\left(\frac{z}{n} + \theta\right) = z' \quad (12.1.6)$$

Then

$$F_{Z_n}(z) = 1 - \prod_{i=1}^{n} \Pr(X_i > z') \tag{12.1.7}$$

$$= 1 - (1 - F(z'))^n \tag{12.1.8}$$

$$\implies F_{Z_n}(z) = 1 - \left(1 - F\left(\frac{z}{n} + \theta\right)\right)^n \tag{12.1.9}$$

where

$$F(x) = \int_{-\infty}^{x} f(t)\, dt \tag{12.1.10}$$

$$= \begin{cases} 1 - e^{-(x-\theta)} & \text{when } \theta < x < \infty \\ 0 & \text{otherwise} \end{cases} \tag{12.1.11}$$

Substituting from (12.1.11) in (12.1.9),

$$F_{Z_n}(z) = \begin{cases} 1 - e^{-n\left(\frac{z}{n}+\theta-\theta\right)} & \theta < \frac{z}{n} + \theta < \infty \\ 0 & \text{otherwise} \end{cases} \tag{12.1.12}$$

$$= \begin{cases} 1 - e^{-z} & \text{when } 0 < z < \infty \\ 0 & \text{otherwise} \end{cases} \tag{12.1.13}$$

and

$$f_{Z_n}(z) = \frac{d}{dz} F_{Z_n}(z) \tag{12.1.14}$$

$$= \begin{cases} e^{-z} & 0 < z < \infty \\ 0 & \text{otherwise} \end{cases} \tag{12.1.15}$$

The plots for the cdf in (12.1.13) and the pdf in (12.1.15) are shown in Fig. 12.1.1 and Fig. 12.1.2 respectively:

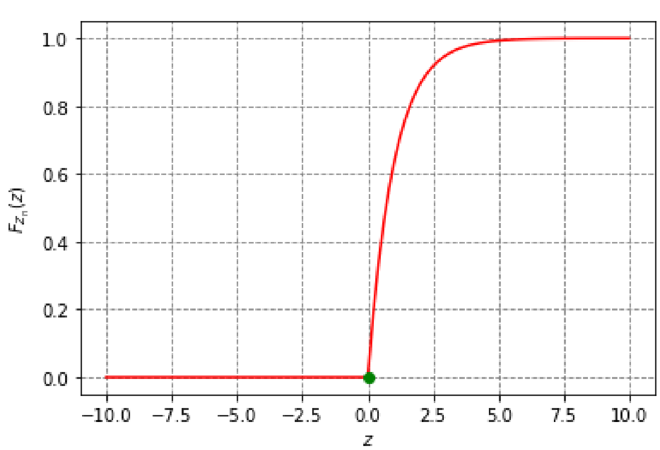

Fig. 12.1.1: cdf of Z_n

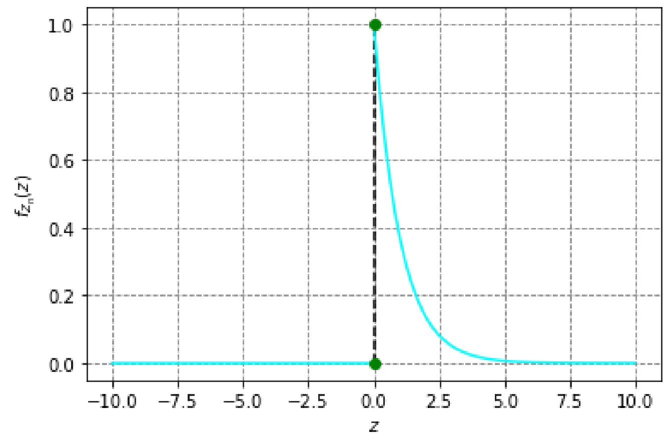

Fig. 12.1.2: pdf of Z_n

12.2. Let $X_1, X_2, \dots, X_n$ be a random sample of size n (≥ 2) from a distribution having the probability density function

$$f(x;\theta) = \begin{cases} \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) & x > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{12.2.1}$$

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) $\dfrac{T}{n^2}$

(B) $\dfrac{T}{n}$

(C) $\dfrac{(n+1)T}{2n}$

(D) $\dfrac{(n+1)^2T}{4n^2}$

Solution:

Lemma 12.1. Lehmann–Scheffé theorem : *If T is a complete sufficient statistic for θ and*

$$E(g(T)) = \tau(\theta) \tag{12.2.2}$$

then $g(T)$ is the uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(\theta)$.

We know that

$$T = \sum_{i=1}^{n} X_i \tag{12.2.3}$$

is a complete and sufficient statistic. By the law of total expectation,

$$E\left(E(X_{(1)}|T)\right) = E(X_{(1)}) \tag{12.2.4}$$

By Lehmann–Scheffé theorem, with

$$\theta = X_{(1)}, \tag{12.2.5}$$
$$\tau(x) = E(x), \tag{12.2.6}$$
$$g(T) = E(X_{(1)}|T). \tag{12.2.7}$$

it follows from (12.2.4) that $E(X_{(1)}|T)$ is the UMVUE of $E(X_{(1)})$.

$$\Pr(X_{(1)} > x) = \Pr(X_1 > x) \dots \Pr(X_n > x) \tag{12.2.8}$$
$$= (1 - F_{X_1}(x)) \dots (1 - F_{X_n}(x)) \tag{12.2.9}$$
$$= (1 - F_{X_1}(x))^n \tag{12.2.10}$$
$$= \exp\left(-\frac{nx}{\theta}\right) \tag{12.2.11}$$
$$F_{X_{(1)}}(x) = 1 - \exp\left(-\frac{nx}{\theta}\right) \tag{12.2.12}$$
$$f_{X_{(1)}}(x) = \frac{n}{\theta}\exp\left(-\frac{nx}{\theta}\right) \tag{12.2.13}$$

Therefore, $X_{(1)}$ follows an exponential distribution with mean $\frac{\theta}{n}$.

$$E(X_{(1)}) = \frac{\theta}{n} \tag{12.2.14}$$

Note that,

$$E\left(\frac{T}{n^2}\right) = E\left(\frac{\sum_{i=1}^{n} X_i}{n^2}\right) \tag{12.2.15}$$
$$= \frac{E(\sum_{i=1}^{n} X_i)}{n^2} \tag{12.2.16}$$
$$= \sum_{i=1}^{n} \frac{E(X_i)}{n^2} \tag{12.2.17}$$
$$= \sum_{i=1}^{n} \frac{\theta}{n^2} \tag{12.2.18}$$
$$= \frac{\theta}{n} \tag{12.2.19}$$
$$= E(X_{(1)}) \tag{12.2.20}$$

Therefore, by Lehmann–Scheffé theorem, with

$$\theta = X_{(1)}, \tag{12.2.21}$$
$$\tau(x) = E(x), \tag{12.2.22}$$
$$g(T) = \frac{T}{n^2}, \tag{12.2.23}$$

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

$$E(X_{(1)}|T) = \frac{T}{n^2} \tag{12.2.24}$$

Hence, option A is correct.

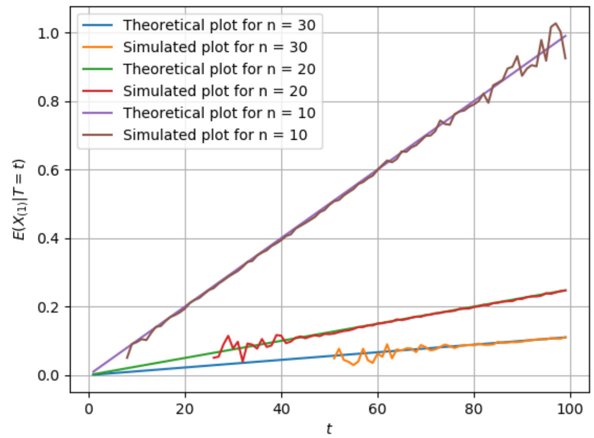

Fig. 12.2.1: Theory vs Simulated plot of $E(X_{(1)}|T)$