Some More Results: For the independent events E and F

1. Probability that one of E and F occur

$$= P[(E \cap \bar{F}) \cup (\bar{E} \cap F)] = P(E \cap \bar{F}) + P(\bar{E} \cap F)$$
$$= P(E) \cdot P(\bar{F}) + P(\bar{E}) \cdot P(F)$$
$$= P(E)\,[1 - P(F)] + P(F)\,[1 - P(E)]$$
$$= P(E) + P(F) - 2P(E)\,P(F).$$

2. Probability that none of these, occur

$$= P(\bar{E} \cap \bar{F}) = P(\bar{E})\,P(\bar{F})$$
$$= [1 - P(E)]\,[1 - P(F)] = 1 - P(E) - P(F) + P(E)\,P(F).$$

3. Probability that atleast one of these occur

$$= P(E \text{ or } F) = P(E \cup F) = P(E) + P(F) - P(E)\,P(F).$$

4. If E and F are mutually exclusive events, then

$$P(E \cup F) = P(E) + P(F) \text{ since } P(E \cap F) = 0.$$

EXAMPLE 30.11. A bag A has 2 white and 4 red balls and another bag B has 5 white and 7 red balls. A ball is transferred from bag A to bag B and then a ball is drawn from bag B. What is the probability that it is white ?

SOLUTION: The probability of drawing a white ball from bag B will depend upon whether the transferred ball is white or red.

If a red ball is transferred from bag A its probability is $\frac{4}{6}$. There are now 5 white and 8 red balls in bag B hence the probability of drawing a white ball from bag B is $\frac{5}{5+8} = \frac{5}{13}$. Thus, the probability of drawing a white ball from the bag B if the transferred ball is red $= \frac{4}{6} \cdot \frac{5}{13} = \frac{10}{39}$.

Similarly, the probability of drawing a white ball from the bag B if the transferred ball is white $= \frac{2}{6} \cdot \frac{6}{13} = \frac{2}{13}$

Therefore, the required probability $= \frac{10}{39} + \frac{2}{13} = \frac{16}{39}$. **Ans.**

EXAMPLE 30.12. The odds in favour of a particular candidate getting selected in an interview for a post by three experts are 5 to 2, 4 to 3 and 3 to 4. Find the probability that the candidate will be selected by the majority of the experts.

SOLUTION: The probability that the particular candidate will be favoured by the first expert is 5/7, by the second is 4/7 and by the third is 3/7.

A majority of the three experts will favour him when two or three are favourable.

Probability that the first two are favourable and third unfavourable $= \frac{5}{7} \cdot \frac{4}{7}\left(1 - \frac{3}{7}\right) = \frac{80}{343}$.

Probability that the first and third are favourable and second not $= \frac{5}{7} \cdot \frac{3}{7} \cdot \left(1 - \frac{4}{7}\right) = \frac{45}{343}$

Probability that first is unfavourable and last two are favourable $= \left(1 - \frac{5}{7}\right) \cdot \frac{4}{7} \cdot \frac{3}{7} = \frac{24}{343}$

Probability that all the three are favourable $= \frac{5}{7} \cdot \frac{4}{7} \cdot \frac{3}{7} = \frac{60}{343}$

Since these are mutually exclusive, the required probability $= \frac{80}{343} + \frac{45}{343} + \frac{24}{343} + \frac{60}{343} = \frac{209}{343}$. **Ans.**

EXAMPLE 30.13. (*a*) A player thinks there are 40% chances that the tournament will be held in the college this year. If it is so then he is 70% sure that he will be the champion. What is the probability that this player will be champion in the tournament?

(*b*) A bag contains 5 white and 3 black balls. Two balls are drawn at random one after the other without replacement. What is the probability that both balls drawn are black ?

SOLUTION: (*a*) Let T denote the event that there will be tournament in the college and C denote the event that the will be the champion then the desired probability $P(TC)$ is given by

$$P(TC) = P(T)\,P(C/T)$$

P(T) P(C∩T) / P(T)

$$= \frac{40}{100} \times \frac{70}{100} = 0.28 \quad \textbf{Ans.}$$

(*b*) Let A be the event of drawing a black ball in the first attempt then

$$\text{P (A)} = \frac{3}{5+3} = \frac{3}{8}.$$

Let B be the event of drawing a black ball in the second attempt then the probability of the second ball drawn being black, given that the first ball drawn is black, is

3/8 × 2/7

$$= \text{P (B/A)} = \frac{2}{5+2} = \frac{2}{7}$$

Hence the required probability that both the balls drawn in succession are black, is

$$\text{P (AB)} = \text{P (A) P (B/A)} = \frac{3}{8} \cdot \frac{2}{7} = \frac{3}{28}. \quad \textbf{Ans.}$$

EXAMPLE 30.14. A box contains 10 balls of which 3 are coloured and dotted, 1 is coloured and striped, 2 are gray and dotted and 4 are gray and striped. If we draw a coloured ball from the box, what is the probability that (*i*) it a dotted ball (*ii*) it is a striped ball? Also, find the probabilities of other possible events.

SOLUTION:

Number of balls	*Coloured* C	*Dotted* D	*Gray* G	*Striped* S
3	✓	✓		
1	✓			✓
2		✓	✓	
4			✓	✓

(1) The probability that the coloured ball is dotted = $P(D/C) = \frac{P(DC)}{P(C)} = \frac{3}{4}$.

(2) The probability that the coloured ball is striped = $P(S/C) = \frac{P(SC)}{P(C)} = \frac{1}{4}$.

(3) The probability that a gray ball is dotted = $P(D/G) = \frac{P(DG)}{P(G)} = \frac{2}{6} = \frac{1}{3}$.

(4) The probability that a gray ball is striped = $P(S/G) = \frac{P(SG)}{P(G)} = \frac{4}{6} = \frac{2}{3}$.

(5) The probability that a dotted ball is coloured = $P(C/D) = \frac{P(CD)}{P(D)} = \frac{3}{5}$.

(6) The probability that a dotted ball is gray = $P(G/D) = \frac{P(GD)}{P(D)} = \frac{2}{5}$. **Ans.**

BAYE'S THEOREM

If $A_1, A_2, \ldots, A_n$ are exhaustive events of a sample space and an event B corresponds to all the events $A_1, A_2, \ldots, A_n$, that is, the event B has some part common in all the events $A_1, A_2, \ldots, A_n$, and if $P(A_i)$ and $P(B/A_i)$ for $i = 1, 2, \ldots, n$, are known, then

$$\boxed{P(A_i/B) = \frac{P(A_i)\,P(B/A_i)}{\sum_1^n P(A_i)\,P(B/A_i)}, \quad i = 1, 2, \ldots, n.}$$

$\frac{P(A_i)\,P(B/A_i)}{\sum_1 P(A_i)\,P(B/A_i)}$

Proof: By the multiplication theorem of probability, we have

$$P(B\,A_i) = P(B)\,P(A_i/B) = P(A_i)\,P(B/A_i)$$

$$\therefore \quad P(A_i/B) = \frac{P(A_i)\,P(B/A_i)}{P(B)} \qquad \ldots(1)$$

As the event B corresponds to $A_1, A_2, \ldots, A_n$, we have, by addition of probability,

$$P(B) = P(BA_1) + P(BA_2) + \ldots + P(BA_n)$$

$$= \sum_1^n P(BA_i) = \sum_1^n P(A_i)\,P(B/A_i) \qquad \ldots(2)$$

Using (2) in (1) we get $P(A_i/B) = \frac{P(A_i)\,P(B/A_i)}{\sum_1^n P(A_i)\,P(B/A_i)}, \quad i = 1, 2, \ldots, n.$

EXAMPLE 30.15. A factory has two plants. Records show that the plant I produces 30% of the items of the output whereas plant II produces 70% of the items. Further, 5% of the items produced by plant I are defective while 1% produced by the plant II are defective. If a defective item is drawn at random, find the probability that the defective item was produced by (*i*) plant I, (*ii*) by plant II.

SOLUTION: Let A_1 denote the event of drawing an item produced by plant I, A_2 denote the event of drawing an item produced by plant II and B denote the event of drawing a defective item produced by the plant I or by plant II. We are given that

$$P(A_1) = 0.30, \; P(A_2) = 0.70 \text{ and } P(B/A_1) = 0.05, \; P(B/A_2) = 0.01.$$

The required probabilities are $P(A_1/B)$ and $P(A_2/B)$.

$P(B/A_1) = \frac{P(B \cap A_1)}{P(A_1)}$

By Bayes' theorem $P(A_1/B) = \frac{P(A_1)\,P(B/A_1)}{P(A_1)\,P(B/A_1) + P(A_2)\,P(B/A_2)}$

$$= \frac{0.30 \times 0.05}{0.30 \times 0.05 + 0.70 \times 0.01} = \frac{15}{22} = 0.682$$

Similarly $P(A_2/B) = \frac{P(A_2)\,P(B/A_2)}{P(A_1)\,P(B/A_1) + P(A_2)\,P(B/A_2)}$

$$= \frac{0.70 \times 0.01}{0.30 \times 0.05 + 0.70 \times 0.01} = \frac{7}{22} = 0.318. \qquad \textbf{Ans.}$$

EXAMPLE 30.16. A survey was conducted to find the supplies of the consumer durables for the market. It was found that the three major companies A, B and C have market share of 35%, 25% and 40% respectively out of which 2%, 1% and 3% are not upto the satisfaction. A consumer buys a product and is dissatisfied with it. What is the probability that it is from the company C? [*GGSIPU IV Sem I Term 2015*]

SOLUTION: Let X be the event of drawing a dissatisfied item from any company and the event that an item drawn was produced by companies A, B or C be Y_1, Y_2, Y_3 respectively.

$$\therefore \quad P(Y_1) = 0.35, \quad P(Y_2) = 0.25, \quad P(Y_3) = 0.40.$$

and $\quad P(X/Y_1) = 0.02, \quad P(X/Y_2) = 0.01, \quad P(X/Y_3) = 0.03.$

$$P(Y_1)\,P(X/Y_1) = 0.007, \quad P(Y_2)\,P(X/Y_2) = 0.0025, \quad P(Y_3)\,P(X/Y_3) = 0.012.$$

$$\therefore \text{ Required probability} = \frac{0.012}{0.007 + 0.0025 + 0.012} = \frac{0.012}{0.0215} = 0.558. \qquad \textbf{Ans.}$$

EXAMPLE 30.17. An insurance company insured 2000 scooter drivers, 4000 car drivers and 6000 truck drivers. The probability of accidents is 0.01, 0.03 and 0.15 respectively. One of the insured person meets with an accident. What is the probability that he is a scooter driver?

SOLUTION: Let E_1, E_2, E_3 denote the events that a driver selected at random is a scooter, car or truck driver respectively and let E denote the event of a driver meeting with accident. Then

$$\therefore \quad P(E_1) = \frac{2000}{12000} = \frac{1}{6}, \quad P(E_2) = \frac{4000}{12000} = \frac{1}{3}, \quad P(E_3) = \frac{6000}{12000} = \frac{1}{2}$$

and $\quad P(E/E_1) = 0.01, \quad P(E/E_2) = 0.03, \quad P(E/E_3) = 0.15.$

$\therefore$ Required probability $= P(E_1/E) = \dfrac{P(E_1)\,P(E/E_1)}{P(E_1)\,P(E/E_1) + P(E_2)\,P(E/E_2) + P(E_3)\,P(E/E_3)}$

$$= \frac{\frac{1}{6}\times 0.01}{\frac{1}{6}\times 0.01 + \frac{1}{3}\times 0.03 + \frac{1}{2}\times 0.15} = \frac{0.01}{0.01 + 0.06 + 0.45} = \frac{1}{52} \text{ Ans.}$$

EXAMPLE 30.18. In a bolt factory there are four machines A, B, C, D manufacturing 20%, 15%, 25%, 40% of the total output respectively. Of their outputs 5%, 4%, 3% and 2% in the same order are defective bolts. A bolt is chosen at random from the factory's production and is found to be defective. What is the probability that the bolt was manufactured by machine A or machine D?

SOLUTION: Let E_1, E_2, E_3, E_4 denote the events that a bolt selected at random is manufactured by machine A, B, C and D respectively and E denote the event of it being defective. Then

$$P(E_1) = 0.2,\quad P(E_2) = 0.15,\quad P(E_3) = 0.25,\quad P(E_4) = 0.4$$

Also given are the conditional probabilities

$$P(E/E_1) = 0.05,\quad P(E/E_2) = 0.04,\quad P(E/E_3) = 0.03,\quad P(E/E_4) = 0.02$$

By Baye's theorem, we wave

$$\therefore \quad P(E_1/E) = \frac{P(E_1)\,P(E/E_1)}{P(E_1)\,P(E/E_1) + P(E_2)\,P(E/E_2) + P(E_3)\,P(E/E_3) + P(E_4)\,P(E/E_4)}$$

$$= \frac{0.2(0.05)}{0.2(0.05) + 0.15(0.04) + 0.25(0.03) + 0.40(0.02)}$$

$$= \frac{0.01}{0.01 + 0.006 + 0.0075 + 0.008} = \frac{0.01}{0.0315} = 0.317$$

$$P(E_4/E) = \frac{0.4\,(0.02)}{P(E_1)\,P(E/E_1) + P(E_2)\,P(E/E_2) + P(E_3)\,P(E/E_3) + P(E_4)\,P(E/E_4)} = \frac{0.008}{0.0315}$$

$$= \frac{8}{31.5} = 0.25396.$$

Thus, the probabilities that the bolt was manufactured by machine A or machine D are 0.317 and 0.254 respectively. Required probability = 0.317 + 0.254 = 0.571. **Ans.**

EXAMPLE 30.19. In a test an examinee either guesses or copies or knows the answer to multiple choice questions with 4 choices. The probability that he makes a guess is 1/3 and that he copies the answer is 1/6. The probability that his answer is correct, given that he copies it, is 1/8. Find the probability that he knew the answer to the question, given that he correctly answered it.

SOLUTION: Consider the following events:

A : the examinee guesses the answer

B : the examinee copies the answer

C : the examinee knows the answer

D : the examinee answers correctly.

We are given that $P(A) = 1/3$, $P(B) = 1/6$, $P(D/B) = 1/8$.

$$\Rightarrow \quad P(C) = 1 - P(A) - P(B) = 1 - \frac{1}{3} - \frac{1}{6} = \frac{1}{2}.$$

Obviously $P(D/C) = 1$ (since he knows the answer correctly).

$P(D/A) = 1/4$ (since there are four choices to choose from).

Then by Bay's theorem

$$P(C/D) = \frac{P(D/C)\,P(C)}{P(D/A)\,P(A) + P(D/B)\,P(B) + P(D/C)\,P(C)}$$

$$= \frac{1\left(\frac{1}{2}\right)}{\left(\frac{1}{4}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{8}\right)\left(\frac{1}{6}\right) + (1)\left(\frac{1}{2}\right)} = \frac{24}{29}. \quad \textbf{Ans.}$$

$1 - \frac{1}{3} - \frac{1}{6}$

$\frac{6 - 2 - 1}{6} = \frac{3}{6} = \frac{1}{2}$

$1 - \frac{1}{2} - \frac{1}{8} = \frac{8 - 4 - 1}{8} = \frac{3}{8}$

CHAPTER

31 Probability Distributions: Binomial, Poisson and Normal Distribution

Random Variable, Mathematical Expectation, Moment Generating Function, Discrete and Continuous Probability Distributions; Binomial, Poisson and Normal Distributions.

RANDOM VARIABLE (OR STOCHASTIC VARIABLE)

Consider an experiment having sample space S. A random variable X on a sample space S is a function which assigns a real value to each outcome in the sample space S. Random variables are generally denoted by the capital letters X, Y, etc. and their possible values by corresponding small letters x, y, etc.

For example, consider an experiment of tossing a pair of identical coins and let the random variable X be the number of heads obtained. The sample space S is as follows.

Outcome :	TT	HT	TH	HH
Value of X:	0	1	1	2

Therefore, we have $P(X = 0) = \frac{1}{4}$, $P(X = 1) = \frac{2}{4} = \frac{1}{2}$ and $P(X = 2) = \frac{1}{4}$.

Random variables are either **discrete** or **continuous**. A random variable is said to be **discrete** if its set of possible values is either finite or countably infinite.

The random variable gives rise to **Discrete Probability Distribution** in which there is **probability mass function** $p(x) = P(X = x)$ and if $x_1, x_2, \ldots$ are the possible values of X, then

$$\sum_{i=1}^{\infty} p(x_i) = p(x_1) + p(x_2) + \ldots \infty = 1.$$

In case of **continuous random variable**, the probability is not concentrated on specific points, rather X can take each and every value between two specific limits. Here, we definite a function $f(x)$ called **probability density function**, such that

$$P\left(x - \frac{1}{2}dx \le X \le x + \frac{1}{2}dx\right) = f(x)\,dx \quad \text{and} \quad \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

MATHEMATICAL EXPECTATION

Expectation of a random variable is one of the most important concept in probability theory.

If X is a **discrete random variable** which takes values x_i with probability $P(X = x_i) = p_i$, $i = 1, 2, \ldots,$ then the mathematical expectation (or expected value) of X, denoted by $E(X)$, is defined by

$$\boxed{E(X) = \sum x_i p_i} = \Sigma x_i P(x_i)$$

In other words, the expected value of a random variable X is a weighted average of the possible values of X, each value being weighted by the probability that X assumes.

For example, if a random variable X can take the values 0, 1 and 2 with probability 1/2, 1/4 and 1/4 respectively, then

$$E(X) = 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{4}\right) + 2\left(\frac{1}{4}\right) = \frac{3}{4}$$

However, if X **is a continuous random variable** with probability density function (p.d.f.) $f(x)$, then the expected value of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Observe here that $E(X)$ is the mean or average of the random variable X and is generally denoted by μ.

Further the variance of X in terms of expectation, is given by

$$\sigma^2 = E(X - \mu)^2 = \begin{cases} \sum (x_i - \mu)^2 p_i & \text{in case of discrete r.v.} \\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx & \text{in case of continuous r.v.} \end{cases}$$

In general the r^{th} moment about the mean μ, denoted by μ_r, is given by

$$\mu_r = E(X - \mu)^r = \begin{cases} \sum (x_i - \mu)^r p_i & \text{in case of discrete r.v.} \\ \int_{-\infty}^{\infty} (x-\mu)^r f(x)\, dx & \text{in case of continuous r.v.} \end{cases}$$

Properties of Expectation:

1. **Addition Property.** If X and Y are two random variables, then
$$E(X + Y) = E(X) + E(Y).$$
This can be extended to more than two variables, as
$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$$
2. **Linearity Property.** If X and Y are two random variables and a and b are constants, then $E(aX + bY) = aE(X) + bE(Y)$
which again can be extended to more than two variables.
As a particular case, we have $E(aX + b) = aE(X) + b$.

3. **Multiplication Property.** If X and Y are two independent random variables then
$$E(X, Y) = E(X)\, E(Y).$$
This again can be extended to more than two variables, as
$$E(X_1, X_2 \ldots X_n) = E(X_1)\, E(X_2) \ldots E(X_n)$$

EXAMPLE 31.1. A sales executive recalls his part sales records per phone calls as follows.

Sales in unit (x_i)	0	1	2	3	4	5
Probability (p_i)	0.15	0.20	0.10	0.05	0.30	0.20

Calculate his average number of units sold per call.

SOLUTION: The sales executive can calculate the expected sales units per phone call, as

$$E(X) = p_1X_1 + p_2X_2 + p_3X_3 + p_4X_4 + p_5X_5 + p_6X_6 \quad = \Sigma x_i p_i$$
$$= 0(0.15) + 1(0.20) + 2(0.10) + 3(0.05) + 4(0.30) + 5(0.20)$$
$$= 0.20 + 0.20 + 0.15 + 1.20 + 1.00 = 2.75.$$

This means he can expect to sell 2.75 units per phone call. **Ans.**

EXAMPLE 31.2. (a) A packet contains 4 good blades and 3 defective blades. A customer picks 3 blades from it randomly. What is the expected value of the number of good blades he picks?

(b) What is the expectation of number of failures proceeding the first success in an infinite series of independent trials with constant probability (2/3) of success in each trial. [GGSIPU IV Sem End Term 2015]

SOLUTION: (a) Let X be the number of good blades picked by the customer out of the sample. The probability function $p(x)$ is given by

$$p(x) = \frac{{}^4C_x\, {}^3C_{3-x}}{{}^7C_3}, \quad x = 0, 1, 2, 3.$$

Then
$$p(0) = \frac{{}^4C_0 \cdot {}^3C_3}{{}^7C_3} = \frac{3!}{7\cdot 6\cdot 5} = \frac{1}{35}, \quad p(1) = \frac{{}^4C_1\, {}^3C_2}{{}^7C_3} = \frac{12}{35},$$
$$p(2) = \frac{{}^4C_2 \cdot {}^3C_1}{{}^7C_3} = \frac{18}{35} \quad \text{and} \quad p(3) = \frac{{}^4C_3 \cdot {}^3C_0}{{}^7C_3} = \frac{4}{35}.$$

Therefore, $E(X) = \sum x_i p_i = 0\left(\frac{1}{35}\right) + 1\left(\frac{12}{35}\right) + 2\left(\frac{18}{35}\right) + 3\left(\frac{4}{35}\right) = \frac{60}{35} = \frac{12}{7}.$ **Ans.**

(b) Let X denote the random variable representing the number of failures proceeding the first success. Thus X may take on the values $0, 1, 2, 3, \ldots$ each with probability $p, q p, q^2p, q^3p, \ldots$ where $q = 1 - p$.

$\therefore$ By definition
$$E(X) = \sum p_i x_i = qp + 2q^2p + 3q^3p + \ldots$$
$$= qp\,[1 + 2q + 3q^2 + \ldots]$$

Let
$$S = 1 + 2q + 3q^2 + 4q^3 + \ldots$$
$\therefore$
$$qS = q + 2q^2 + 3q^3 + 4q^4 + \ldots$$

$$\Rightarrow \quad S(1-q) = 1 + q + q^2 + q^3 + \dots = \frac{1}{1-q}$$

$$\Rightarrow \quad S = \frac{1}{(1-q)^2}, \quad \text{Given } p = \frac{2}{3} \text{ hence } q = \frac{1}{3}$$

$$\therefore \quad E(X) = qpS = \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{9}{4} = \frac{1}{2}. \qquad \textbf{Ans.}$$

EXAMPLE 31.3. *If X represents the outcome when a pair of dice is thrown, find the expectation $E(X)$ and variance $V(X)$.*

SOLUTION: Here $X = i$, $i = 1, 2, \dots, 6$ and $P(X = i) = \frac{1}{6}$.

Therefore $$E(X) = \sum i p_i = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{21}{6} = \frac{7}{2}$$

and $$E(X^2) = \sum i^2 p_i = 1\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 9\left(\frac{1}{6}\right) + 16\left(\frac{1}{6}\right) + 25\left(\frac{1}{6}\right) + 36\left(\frac{1}{6}\right) = \frac{91}{6}.$$

Hence $$\boxed{V(X) = E(X^2) - [E(X)]^2} = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}. \qquad \textbf{Ans.}$$

COVARIANCE AND VARIANCE

Covariance: If $E(X) = \bar{X}$ and $E(Y) = \bar{Y}$, then the covariance of the random variables X and Y, denoted by cov (X, Y), is given by

$$\begin{aligned}
\text{cov}(X, Y) &= E[(X - \bar{X})(Y - \bar{Y})] = E(XY - X\bar{Y} - Y\bar{X} + \bar{X}\bar{Y}) \\
&= E(XY) - \bar{Y}E(X) - \bar{X}E(Y) + \bar{X}\bar{Y} \\
&= E(XY) - \bar{Y}\bar{X} - \bar{X}\bar{Y} + \bar{X}\bar{Y}
\end{aligned}$$

Thus $$\boxed{\text{cov}(X, Y) = E(XY) - \bar{X}\bar{Y}.}$$

Note that, when X and Y are independent, we have

$$E(XY) = E(X)\,E(Y), \text{ then } \text{cov}(X, Y) = 0.$$

Also, we can easily establish that

1. $\text{cov}(aX, bY) = ab\ \text{cov}(X, Y)$
2. $\text{cov}(X + a, Y + b) = \text{cov}(X, Y)$
3. $\text{cov}\left(\frac{X-a}{h}, \frac{Y-b}{k}\right) = \frac{1}{hk}\text{cov}(X, Y)$
4. $\text{cov}(X, Y, Z) = \text{cov}(X, Y) + \text{cov}(X, Z)$.

Variance: It is easy to verify the following properties:

1. $\text{Var}(X) = \text{cov}(X, X)$
2. $\text{Var}(ax + b) = a^2\ \text{Var}(X)$
3. $\text{Var}(aX + bY) = a^2\ \text{Var}(X) + b^2\ \text{Var}(Y) + 2ab\ \text{cov}(X, Y)$.

MOMENT GENERATING FUNCTION

The moment generating function (m.g.f.) of a random variable X about the origin is defined as

$$\boxed{M_0(t) = E(e^{tX}) = \sum e^{tX} P(X).} \quad M_0(t) = \sum e^{tX} P(x) = E(e^{tx})$$

This function $M_0(t)$ is called moment generating function because all the moments of X can be obtained from $M_0(t)$, as follows:

$$\begin{aligned}
M_0(t) &= \sum\left[1 + tX + \frac{t^2}{2!}X^2 + \frac{t^3}{3!}X^3 + \dots + \frac{t^r X^r}{r!} + \dots\right] p(X) \\
&= 1 + t\,E(X) + \frac{t^2}{2!}E(X^2) + \dots + \frac{t^r}{r!}E(X^r) + \dots \\
&= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \dots + \frac{t^r}{r!}\mu_r' + \dots \qquad \text{where } \mu_r' = E(X^r).
\end{aligned}$$

Obviously, we can write here $$\mu_r' = \left[\frac{d^r}{dt^r}[M_0(t)]\right]_{t=0}.$$

As a general case, the moment generating function of the random variable X about an arbitrary point 'a', which can be $\bar{X}$ (or μ) also, is defined as

$$M_a(t) = E(e^{t(X-a)}) = e^{-at}E(e^{tX}) = e^{-at}M_0(t).$$

Next, the moment generating function of the sum of two independent variables is the product of their moment generating functions.

To establish it, consider two independent random variables X and Y whose m.g.fs. are $M_X(t)$ and $M_Y(t)$ respectively, then m.g.f. of $X + Y$ is given by

$$\begin{aligned}
M_{X+Y}(t) &= E[e^{t(X+Y)}] = E(e^{tX} \cdot e^{tY}) \\
&= E(e^{tX})\,E(e^{tY}) \qquad \text{(since } X \text{ and } Y \text{ are independent)}
\end{aligned}$$

Thus, $$\boxed{M_{X+Y}(t) = M_X(t)\,M_Y(t).}$$

Further, it is easy to verify that

1. $M_{cX}(t) = M_X(ct)$ where c is some constant
2. If $U = \frac{X-a}{h}$, then $M_U(t) = e^{at/h} M_X\left(\frac{t}{h}\right)$.

DISCRETE PROBABILITY DISTRIBUTION

The distribution which arises out of a discrete random variable is called **discrete probability distribution.** Suppose a random variable X can assume set of discrete values $x_1, x_2, \dots, x_n$ with respective probabilities $p_1, p_2, \dots, p_n$ where $p_1 + p_2 + \dots + p_n = 1$, we say that a discrete probability distribution is defined. The function $P(X)$ which has respective values $p_1, p_2, \dots, p_n$ for $X = x_1, x_2, \dots, x_n$, is called the **probability mass function or simply mass function.**

For example, if a pair of dice is rolled and the random variable X denotes the sum of the numbers appearing on the dice, then the probability distribution is as follows.

$X =$	2	3	4	5	6	7	8	9	10	11	12
Mass function $P(X)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

It is worth nothing that a probability distribution is analogous to a relative frequency distribution with relative frequencies replaced by probabilities.

DISTRIBUTION FUNCTION

The distribution function $F(x)$ of the discrete random variable X, is defined as

$$F(x) = P(X = x) = \sum_{i=1}^{n} p(x_i) \quad \text{where } x_1 \le x, x_2 \le x, \ldots, x_n \le x.$$

This distribution function $F(x)$ is better known as **cumulative distribution function.**

This implies that $P(a \le X \le b) = F(b) - F(a)$.

The mean value (μ) of the probability distribution of the variable X, is generally known as **expectation** of X $\therefore$ $E(X) = \sum x_i\, p(x_i)$.

$$\text{Mean deviation about mean} = \sum |x_i - \mu|\, p(x_i)$$

$$\text{Variance } (X) = \sigma^2 = \sum (x_i - \mu)^2\, p(x_i)$$

and the r^{th} moment about the mean μ, is defined as

$$\mu_r = \sum (X - \mu)^r\, p(x_i).$$

UNIFORM DISTRIBUTION

This is the simplest discrete probability distribution where the random variable assumes different values with uniform (or equal) probability.

A random variable X is said to have a discrete uniform distribution over the range $[1, n]$ if its probability density function can be written as

$$P(X = x_j) = \frac{1}{n}, \quad i = 1, 2, \ldots, n.$$

The distribution suits random experiments when all the outcomes are equally likely, for example, the throwing of an unbiased die or the random draw of a card from a well shuffled pack of cards.

Constants of the Uniform Distribution

$$E(X) = \text{Mean} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}(1+2+\ldots+n) = \frac{1}{n}\cdot\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

and

$$E(X^2) = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{1}{n}\cdot(1^2+2^2+\ldots+n^2) = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}.$$

Hence,

$$\text{Variance} = \sigma^2 = E(X^2) - [E(X)]^2 = \frac{1}{6}(n+1)(2n+1) - \left[\frac{(n+1)}{2}\right]^2$$

$$= \frac{n+1}{12}[2(2n+1) - 3(n+1)] = \frac{n^2-1}{12}.$$

Also, the moment generating function about the origin

$$M_0(t) = E(e^{tX}) = \frac{1}{n}\sum_{x=1}^{n} e^{tX}. \quad \text{(it is a geometric series)}$$

$$= \frac{1}{n}\left[\frac{e^t(1-e^{nt})}{1-e^t}\right] = \frac{e^t - e^{(n+1)t}}{n(1-e^t)}.$$

$S_n = \frac{a(1-r^n)}{1-r}$, $r<1$

EXAMPLE 31.4. Two cards are drawn successively with replacement from a well-shuffled pack of 52 playing cards. Find the probability distribution of the number of aces.

SOLUTION: Let X be the random variable which takes values equal to the number of aces obtained with draw of two cards. Following three cases arise:

Case I: There is no ace.

$$P\text{ (no ace is drawn)} = P(X = 0) = \frac{48}{52}\cdot\frac{48}{52} = \frac{144}{169}.$$

Case II: Of the two drawn cards, one is ace

$$P(X = 1) = \frac{4}{52}\cdot\frac{48}{52} + \frac{48}{52}\cdot\frac{4}{52} = \frac{24}{169}.$$

Case III: Both the cards drawn are aces

$$P(X = 2) = \frac{4}{52}\cdot\frac{4}{52} = \frac{1}{169}.$$

Thus, the probability distribution is as follows

X	0	1	2
$P(X)$	$\frac{144}{169}$	$\frac{24}{169}$	$\frac{1}{169}$

Ans.

EXAMPLE 31.5. Find the probability distribution of the number of green balls drawn when three balls are drawn one by one without replacement from a bag containing three green and five white balls.

SOLUTION: Let X be the random variable which represents the number of green balls drawn when three balls are drawn without replacement.

$$\therefore \quad P(X = 0) = P\text{ (no green ball is drawn)} = P(W, W, W)$$

$$= \frac{5}{8}\cdot\frac{4}{7}\cdot\frac{3}{6} = \frac{5}{28}.$$

$$P(X = 1) = P\text{ (one green ball is drawn)}$$

$$= P(G, W, W) + P(W, G, W) + P(W, W, G)$$

$$= \frac{3}{8}\cdot\frac{5}{7}\cdot\frac{4}{6} + \frac{5}{8}\cdot\frac{3}{7}\cdot\frac{4}{6} + \frac{5}{8}\cdot\frac{4}{7}\cdot\frac{3}{6} = \frac{15}{28}.$$

$$P(X = 2) = P\text{(two green balls are drawn)}$$

$$= P(G, G, W) + P(G, W, G) + P(W, G, G)$$

$$= \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{5}{6} + \frac{3}{8}\cdot\frac{5}{7}\cdot\frac{2}{6} + \frac{5}{8}\cdot\frac{3}{7}\cdot\frac{2}{6} = \frac{15}{56}$$

$$P(X = 3) = P(G\,G\,G) = \frac{3}{8}\cdot\frac{2}{7}\cdot\frac{1}{6}.$$

Probability distribution is

X	0	1	2	3
$P(X)$	$\frac{5}{28}$	$\frac{15}{28}$	$\frac{15}{56}$	$\frac{1}{56}$

EXAMPLE 31.6. In a supply of 10 similar T.V.s by a company 4 are known to be defective. A college purchases 3 TVs from this company. Find the probability distribution for the number of defective TVs purchased and the distribution function.

SOLUTION: If the random variable X denote the number of defective TVs then X can take the values 0, 1, 2, 3, therefore

$$p(0) = P(X = 0) = \frac{{}^4C_0 \cdot {}^6C_3}{{}^{10}C_3} = \frac{20}{120} = \frac{1}{6}$$

$$p(1) = P(X = 1) = \frac{{}^4C_1 \cdot {}^6C_2}{{}^{10}C_3} = \frac{60}{120} = \frac{1}{2}$$

$$p(2) = P(X = 2) = \frac{{}^4C_2 \cdot {}^6C_1}{{}^{10}C_3} = \frac{36}{120} = \frac{3}{10}$$

and $$p(3) = P(X = 3) = \frac{{}^4C_3 \cdot {}^6C_0}{{}^{10}C_3} = \frac{4}{120} = \frac{1}{30}$$

The probability distribution $p(x)$ of X and the distribution function $F(x)$ are given by

X	0	1	2	3
$p(x)$	1/6	1/2	3/10	1/30
$F(x)$	1/6	2/3	29/30	1

Ans.

EXAMPLE 31.7. A random variate X has the following probability distribution

x	0	1	2	3	4
$p(x)$	k	$2k$	$2k$	k^2	$5k^2$

Find the value of k and obtain $P(X < 3)$ and $P(0 < X < 4)$.
Determine the distribution function of X, also mean and variance of X.

SOLUTION: Since $\sum p(x) = 1$ we have $k + 2k + 2k + k^2 + 5k^2 = 1$ or $6k^2 + 4k + k - 1 = 0$ which gives $k = -1$ and $k = 1/6$. Therefore $k = 1/6$ as k cannot be negative. **Ans.**

Next, $$P(X < 3) = p(0) + p(1) + p(2) = \frac{1}{6} + \frac{2}{6} + \frac{2}{6} = \frac{5}{6}.$$

and $$P(0 < X < 4) = p(1) + p(2) + p(3) = \frac{2}{6} + \frac{2}{6} + \frac{1}{36} = \frac{25}{36}$$

Now, the probability distribution and the distribution function is as follows:

x	0	1	2	3	4
$p(x)$	1/6	2/6	2/6	1/36	5/36
$F(x)$	1/6	3/6	5/6	31/36	1

Then $E(x) =$ $$\text{mean} = \mu_x = \sum x_i p_i = 0 + \frac{2}{6} + \frac{4}{6} + \frac{3}{36} + \frac{20}{36} = \frac{59}{36}$$

$$\text{Variance} = \sigma_x^2 = \left[0\left(\frac{1}{6}\right) + 1\left(\frac{2}{6}\right) + 4\left(\frac{2}{6}\right) + 9\left(\frac{1}{36}\right) + 16\left(\frac{5}{36}\right)\right] - \left(\frac{59}{36}\right)^2$$

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

Two most important discrete probability distributions are – Binomial Distribution and Poisson's distribution.

BINOMIAL PROBABILITY DISTRIBUTION

Binomial distribution was introduced by J. Bernoulli. It expresses probability in terms of success and failure. Here an experiment is performed under same conditions n times. If in each trial, all trials being independent, the probability of occurrence of an event, called 'success', is p and of non-occurrence of the event, called 'failure' is $q\ (= 1 - p)$, then the probability of r successes is ${}^nC_r p^r q^{n-r}$. Thus, the probabilities of 0, 1, 2, ..., r, ..., n successes, are

$$q^n,\quad {}^nC_1 p\, q^{n-1},\quad {}^nC_2 p^2 q^{n-2},\ \ldots,\quad {}^nC_r p^r q^{n-r},\ \ldots,\quad p^n.$$

Actually, these probabilities appear as successive terms in the binomial expansion of $(q + p)^n$, that is why it is called binomial distribution.

Here the total probability $= q^n + {}^nC_1 q^{n-1} p + {}^nC_2 q^{n-2} p^2 + \ldots \ldots + p^n = (q + p)^n = 1.$

Therefore, the binomial distribution is a probability distribution.

If we are to obtain the **probable frequencies** of various outcomes when the n independent trials constitute one experiment and this experiment is repeated N times, then the frequency of r successes is $N \cdot {}^nC_r p^r q^{n-r}$. This comes from

$$N(q + p)^n = N q^n + N \cdot {}^nC_1 q^{n-1} p + N \cdot {}^nC_2 q^{n-2} p^2 + \ldots + N \cdot {}^nC_r q^{n-r} p^r + \ldots + N p^n.$$

The frequencies in the above expression constitute the **binomial frequency distribution** and are known as **the expected or theoretical frequencies.** On the other hand the frequencies obtained by actually performing the experiment are called **observed frequencies.** Generally there is some difference between the theoretical and observed frequencies and the difference becomes smaller and smaller as N increases.

THE CONSTANTS OF THE BINOMIAL DISTRIBUTION

$$\text{The arithmetic mean} = \frac{\sum_{x=0}^{n} x\, p(x)}{\sum_{x=0}^{n} p(x)} = \sum_{0}^{n} x\, p(x) \qquad \left(\text{since } \sum_{x=0}^{n} p(x) = 1\right).$$

$$= \sum_{0}^{n} x \cdot {}^nC_x p^x q^{n-x} = \sum_{0}^{n} \frac{x\, n!}{x!(n-x)!} p^x q^{n-x} = np \sum_{0}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$$= np \sum_{1}^{n} {}^{n-1}C_{x-1}\, p^{x-1} q^{n-x} = np\,(q + p)^{n-1} = np.$$

$\therefore$ **Mean of the binomial distribution** $= \boxed{\bar{x} = np.}$ $= \mu_1$

Next, the variance σ^2 of the binomial distribution, is given by

$$\sigma^2 = \mu_2 = \frac{\sum_{x=0}^{n} (x - \bar{x})^2 p(x)}{\sum_{x=0}^{n} p(x)} = \sum_{0}^{n} x^2 p(x) - \left(\sum_{0}^{n} x\, p(x)\right)^2 \qquad \left(\text{as } \sum_{0}^{n} p(x) = 1\right)$$

$$= \sum_{0}^{n} x^2\, {}^nC_x\, p^x q^{n-x} - (np)^2 = \sum_{0}^{n} \frac{x^2\, n!\, p^x q^{n-x}}{x!(n-x)!} - (np)^2 \qquad (\text{as } \bar{x} = np)$$

$$= np \sum_{1}^{n} \frac{x\,(n-1)!\,p^{x-1}\,q^{n-x}}{(x-1)!\,(n-x)!} - (np)^2 = np \sum_{1}^{n} \frac{(x-1+1)\,(n-1)!\,p^{x-1}\,q^{n-x}}{(x-1)!\,(n-x)!} - (np)^2$$

$$= np \sum_{2}^{n} \frac{(n-1)!\,p^{x-1}\,q^{n-x}}{(x-2)!\,(n-x)!} + np \sum_{1}^{n} \frac{(n-1)!\,p^{x-1}\,q^{n-x}}{(x-1)!\,(n-x)!} - (np)^2$$

$$= np^2\,(n-1) \sum_{2}^{n} {}^{n-2}C_{x-2}\, p^{x-2}\, q^{n-x} + np \sum_{1}^{n} {}^{n-1}C_{x-1}\, p^{x-1}\, q^{n-x} - (np)^2$$

$$= np^2\,(n-1)\,(q+p)^{n-2} + np\,(q+p)^{n-1} - (np)^2$$

$$= n\,(n-1)\,p^2 \cdot 1 + np \cdot 1 - n^2 p^2 = np\,(1-p) = npq.$$

Thus, the **variance of the binomial distribution** $\boxed{= \sigma^2 = \mu_2 = npq.}$

Also, we can easily show that

$$\mu_3 = npq\,(q-p), \qquad \mu_4 = npq\,[1 + 3\,(n-2)\,pq]$$

$$\therefore \qquad \beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{n^2p^2q^2\,(q-p)^2}{n^3p^3q^3} = \frac{(q-p)^2}{npq}.$$

If $\beta_1 = 0$ the distribution is symmetrical, *i.e.*, there is no skewness.

Measure of skewness $= \gamma_1 = \sqrt{\beta_1}$.

$3 + \frac{1-6pq}{npq}$

Kurtosis $= \beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{3n^2p^2q^2 + npq\,(1-6\,pq)}{n^2p^2q^2} = \boxed{3 + \dfrac{1-6\,pq}{npq}}.$

The above constants can also be calculated by employing the **moment generating function.**

The m.g.f. of Binomial distribution about the origin denoted by $M_o\,(t)$, is given by

$$M_o\,(t) = E\,(e^{tx}) = \Sigma\, {}^nC_x\, p^x\, q^{n-x} \cdot e^{tx}$$

$$= \Sigma\, {}^nC_x\, (pe^t)^x\, q^{n-x} = (q + pe^t)^n$$

$\beta_1 = \frac{(q-p)^2}{npq}$

For example, $\mu_1' = \left[\frac{d}{dt} M_o\,(t)\right]_{t=0} = [n\,(q+pe^t)^{n-1} \cdot pe^t]_{t=0} = np.$

EXAMPLE 31.8. (*a*) A coin is tossed six times. Calculate the probability of obtaining four or more heads.

(*b*) If the sum of the mean and variance of a Binomial distribution of 5 trials is $\frac{9}{5}$, find $P(X \geq 1)$.

SOLUTION: (*a*) Here $p = q = 1/2$ and $n = 6$.

$n = 5$

$\bar{x} = np$

Probability of getting 4 heads $= P\,(4) = {}^6C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^2 = \frac{6 \times 5}{2}\frac{1}{2^6} = \frac{15}{64} = 0.234.$

$\bar{x} + \sigma^2$

Probability of getting 5 heads $= P\,(5) = {}^6C_5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^1 = \frac{6}{64} = 0.095.$

$= \frac{9}{5}$

Probability of getting 6 heads $= P\,(6) = \left(\frac{1}{2}\right)^6 = \frac{1}{64} = 0.016.$

$\therefore$ Probability of getting 4 or more heads $= P\,(4) + P\,(5) + P\,(6) = 0.345$ **Ans.**

$\sigma^2 = npq$

$\bar{x} = np$

(*b*) We know that mean and variance of a Binomial distribution are np and npq respectively. Here $n = 5$.

$\therefore \quad 5p + 5pq = \frac{9}{5}$. Since $q = 1 - p$ we have $25p^2 - 50p + 9 = 0$

$\therefore \quad p = \frac{50 \pm \sqrt{2500 - 900}}{50} = \frac{50 \pm 40}{50} = \frac{1}{5}$ only. $\therefore\ P\,(X = 0) = q^5 = \left(\frac{4}{5}\right)^5$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \left(\frac{4}{5}\right)^5 = 1 - (0.8)^5 = 0.6724 \quad \textbf{Ans.}$$

EXAMPLE 31.9. If one out of every 10 bulbs is defective, find (*i*) mean and standard deviation for the distribution of defective bulbs in a total of 500 bulbs (*ii*) the coefficient of skewness γ_1 and coefficient of kurtosis β_2.

SOLUTION: (*i*) Here $p = 1/10 = 0.1$, $n = 500$, $q = 0.9$.

Mean $= np = 50$ therefore we can expect 50 bulbs to be defective out of a total of 500.

Standard deviation, $\sigma = \sqrt{npq} = \sqrt{500 \times \frac{1}{10} \times \frac{9}{10}} = \sqrt{45} = 6.71.$ **Ans.**

(*ii*) The coefficient of skewness γ_1 is given by

$$\gamma_1 = \sqrt{\beta_1} = \sqrt{\frac{(q-p)^2}{npq}} = \frac{q-p}{\sqrt{npq}} = \frac{0.9 - 0.1}{6.71} = 0.119$$

The coefficient of kurtosis γ_2 is given by

$$\beta_2 = 3 + \frac{1 - 6\,pq}{npq} = 3 + \frac{1 - 6\,(0.1)\,(0.9)}{45} = 3 + \frac{1 - 0.54}{45} = 3.01. \quad \textbf{Ans.}$$

EXAMPLE 31.10. Compute the variance of the probability distribution of the number of doublets in four throws of a pair of dice.

SOLUTION: Probability of getting a doublet in a throw of a pair of dice $= \frac{6}{36} = \frac{1}{6} = p$, $\therefore\ q = \frac{5}{6}$.

Let X denote the number of doublets obtained in four throws of pair of dice.

Then $P(X = 0) = \left(\frac{5}{6}\right)^4 = \frac{625}{1296}$, $\quad P(X = 1) = {}^4C_1 \frac{1}{6}\left(\frac{5}{6}\right)^3 = \frac{500}{1296}$

$P(X = 2) = {}^4C_2 \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2 = \frac{150}{1296}$, $\quad P(X = 3) = {}^4C_3 \left(\frac{1}{6}\right)^3 \cdot \frac{5}{6} = \frac{20}{1296}$

and $\quad P(X = 4) = \left(\frac{1}{6}\right)^4 = \frac{1}{1296}$.

To calculate mean and variance of the distribution, we have

X	P	PX	X^2	PX^2
0	625/1296	0	0	0
1	500/1296	500/1296	1	500/1296
2	150/1296	300/1296	4	600/1296
3	20/1296	60/1296	9	180/1296
4	1/1296	4/1296	16	16/1296

$$\text{Mean} = \Sigma PX = 864/1296 = \frac{2}{3}, \qquad \Sigma PX^2 = 1296/1296 = 1$$

$$\text{Variance } \sigma^2 = \Sigma PX^2 - (\Sigma PX)^2 = 1 - \left(\frac{2}{3}\right)^2 = \frac{5}{9}.$$

EXAMPLE 31.11. (*a*) The probability that a bomb dropped from a plane will hit the target is $\frac{1}{5}$. If 6 bombs are dropped, find the probability that (*i*) exactly two will strike the target (*ii*) at least two will hit the target.

(*b*) The incidence of occupational disease in an industry is such that the workmen have 10% chances of suffering from it. What is the probability that in a group of 7, 5 or more will suffer from it?

SOLUTION: (*a*) Probability of a bomb hitting the target $= p = \frac{1}{5}$ hence $q = 1 - p = \frac{4}{5}$

(*i*) The probability that exactly two bombs will strike the target $= {}^6C_2\left(\frac{1}{5}\right)^2\left(\frac{4}{5}\right)^4$.

$$= \frac{6.5}{1.2}\cdot\frac{1}{25}\cdot\frac{256}{625} = \frac{3\times 256}{5\times 625} = \frac{1536}{625} = 0.246.$$

(*ii*) The probability that atleast two will hit the target

$$= 1 - [\text{Prob. that no bomb will strike} + \text{Prob. that one will strike}]$$

$$= 1 - \left[{}^6C_0\left(\frac{1}{6}\right)^0\left(\frac{4}{5}\right)^6 + {}^6C_1\left(\frac{1}{5}\right)^1\left(\frac{4}{5}\right)^5\right]$$

$$= 1 - \left(\frac{4}{5}\right)^5\left(\frac{4}{5}+\frac{6}{5}\right) = 1 - 2(0.8)^5 = 0.345. \qquad \textbf{Ans.}$$

(*b*) The probability p of a workman suffering $= 0.1$, $\therefore q = 0.9$

This follows a Binomial distribution.

Probability that 5 or more will be suffering from the disease, is equal to

$$= {}^7C_5(0.1)^5(0.9)^2 + {}^7C_6(0.1)^6(0.9)^1 + {}^7C_7(0.1)^7(0.9)^0$$

$$= 21(0.1)^5(0.9)^2 + 7(0.1)^6(0.9) + (0.1)^7 = (0.1)^5[21\times 0.81 + 0.1\times 6.3 + 0.01]$$

$$= 0.00001\,(17.01 + 0.63 + 0.01) = 0.001765 \qquad \textbf{Ans.}$$

EXAMPLE 31.12. (*a*) A die is tossed thrice. Getting '5' or '6' on the die in a toss is take as success. Find the mean and variance of number of success.

(*b*) Assuming that half the population are consumers of chocolate, so that the chance of an individual being a consumer is 1/2 and assuming that each of the 100 investigators takes 10 individuals to see whether they are consumers. How many investigators would you expect to report that three people or less were consumers.

SOLUTION: (*a*) Let X denote the number of successes then it can take values 0, 1, 2, or 3.

Probability of success $= p = \frac{2}{6} = \frac{1}{3}$ hence $q = 1 - \frac{1}{3} = \frac{2}{3}$. It follows a binomial distribution.

$$P(X=0) = \left(\frac{2}{3}\right)^3 = \frac{8}{27}, \qquad P(X=1) = {}^3C_1\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 = \frac{12}{27},$$

$$P(X=2) = {}^3C_2\left(\frac{1}{3}\right)^2\cdot\left(\frac{2}{3}\right)^1 = \frac{6}{27}, \qquad P(X=3) = {}^3C_3\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^0 = \frac{1}{27}.$$

$\therefore$ Probability distribution of the random variable X is as under.

X	0	1	2	3
P(X)	8/27	12/27	6/27	1/27

For mean and variance of the above distribution, we have

X	P	PX	PX^2
0	8/27	0	0
1	12/27	12/27	12/27
2	6/27	12/27	24/27
3	1/27	3/27	9/27

$$\text{Mean} = \mu = \Sigma PX = \frac{27}{27} = 1$$

$$\text{Variance} = \sigma^2 = \Sigma PX^2 - \mu^2 = \frac{45}{27} - 1 = \frac{2}{3}. \qquad \textbf{Ans.}$$

(*b*) Chance of an individual to be a consumer $= p = \frac{1}{2}$ hence $q = \frac{1}{2}$.

Probability that less than or equal to three people, are consumers

$$= P(r \le 3) = P(r=0) + P(r=1) + P(r=2) + P(r=3)$$

We know that $P(r) = {}^nC_r\, p^r q^{n-r} = {}^{10}C_r\left(\frac{1}{2}\right)^r\left(\frac{1}{2}\right)^{n-r} = {}^{10}C_r\left(\frac{1}{2}\right)^{10}$.

$$\therefore \quad \text{Required probability} = \left(\frac{1}{2}\right)^{10}\left[{}^{10}C_0 + {}^{10}C_1 + {}^{10}C_2 + {}^{10}C_3\right]$$

$$= \left(\frac{1}{2}\right)^{10}\left[1 + 10 + \frac{10.9}{1.2} + \frac{10.9.8}{1.2.3}\right] = \frac{176}{32\times 32} = \frac{11}{64} = 0.172$$

The required number of investigators reporting that, out of 10, at the most three are consumers $= 100 \times 0.172 = 17.2$. **Ans.**

EXAMPLE 31.13. (*a*) Fit a binomial distribution to the following frequency data

x :	0	1	3	4
f :	28	62	10	4

(*b*) Fit a binomial distribution to the following data and compare the theoretical frequencies with the actual ones.

x	0	1	2	3	4	5	6	7	8	9
f	6	20	28	12	8	6	0	0	0	0

SOLUTION: (*a*) The given distribution can be put in tabular form as

x	f	xf
0	28	0
1	62	62
3	10	30
4	4	16

Here $\Sigma f = 104$ and $\Sigma xf = 108$

Hence mean $\bar{x} = \frac{108}{104} = 1.038 = np$ for binomial distribution.

$\therefore$ $p = \frac{1.038}{4} = 0.2595$ and $q = 1 - 0.2595 = 0.7405$.

Therefore, binomial distribution is

$$N(q+p)^n = 104(0.7405 + 0.2595)^4$$
$$= N[(0.7405)^4 + {}^4C_1(0.7405)^3(0.2595) + {}^4C_2(0.7405)^2(0.2595)^2 + {}^4C_1(0.7405)(0.2595)^3 + (0.2595)^4].$$

Ans.

(*b*) Here $n = 10$, $\Sigma f_i = 80$

$$\text{Mean} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{(0+20+56+36+32+30+0)}{80} = \frac{174}{80} = 2.175.$$

Thus, mean of the binomial distribution $= np = 2.175$

$\therefore$ $p = \frac{2.175}{10} = 0.2175$ and $q = 1 - 0.2175 = 0.7825$.

Hence the binomial distribution to be fitted here, is

$$N(q+p)^n = 80(0.7825 + 0.2175)^{10}$$
$$= 80[{}^{10}C_0(0.7825)^{10} + {}^{10}C_1(0.7825)^9(0.2175)^1 + {}^{10}C_2(0.7825)^8(0.2175)^2 + \dots + {}^{10}C_{10}(0.7825)^{10}(0.2175)^0]$$

Successive terms in the above expansion, give the theoretical frequencies as

x	0	1	2	3	4	5	6	7	8	9
f	6.9	19.1	24.0	17.8	8.6	2.9	7	1	0	0

approximately.

Ans.

POISSON'S DISTRIBUTION

This distribution is a limiting form of the binomial distribution. In the binomial distribution let us take n as very large and p as very small such that np = constant = m, say, then

$$P(r) = {}^nC_r\, p^r q^{n-r} = \frac{n(n-1)(n-2)\dots\dots(n-r+1)}{r!}p^r q^{n-r}$$

$$= \frac{np(np-p)(np-2p)\dots\dots(np-pr+p)}{r!}(1-p)^{n-r} \qquad (\text{since } q = 1-p)$$

$$= \frac{m(m-p)(m-2p)\dots\dots(m-pr+p)}{r!}\left(1-\frac{m}{n}\right)^{n-r}$$

$$= \frac{m(m-p)(m-2p)\dots\dots(m-pr+p)\left(1-\frac{m}{n}\right)^n}{r!}\left(1-\frac{m}{n}\right)^{-r}, \quad \text{where } m = np.$$

$$\therefore \quad \underset{n\to\infty}{\text{Lt}}\, P(r) = \frac{m(m-0)(m-0)\dots\dots(m-0)}{r!}(1-0)^{-r}\underset{n\to\infty}{\text{Lt}}\left(1-\frac{m}{n}\right)^n = \frac{m^r}{r!}e^{-m}.$$

Therefore, in Poisson's distribution the probabilities of $0, 1, 2, \dots, r, \dots$ successes are

$$e^{-m},\ m\,e^{-m},\ \frac{m^2}{2!}e^{-m},\ \dots,\ \frac{m^r}{r!}e^{-m},\ \dots.$$

$$\text{The sum of all the probabilities} = e^{-m}\left(1 + m + \frac{m^2}{2!} + \frac{m^3}{3!} + \dots\right) = e^{-m}e^m = 1,$$

Thus, Poisson's distribution is a probability distribution. There is only one constant of this distribution which is m (= np). Here $P(r) = \frac{e^{-m}m^r}{r!}$

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

Next Mean $= \Sigma r\, p(r) = 0\,p(0) + 1p(1) + 2p(2) + \dots \infty$

$$= \sum_0^\infty r\frac{e^{-m}m^r}{r!} = m\sum_{r=1}^\infty \frac{e^{-m}m^{r-1}}{(r-1)!} = m\,e^{-m}\cdot e^m = m. \Rightarrow \sum_0^\infty \lambda \frac{e^{-m}m^{\lambda}}{\lambda!}$$

Thus, **Mean of the Poisson's distribution = m.**

Also, **the variance = $\sigma^2 = \mu_2 = npq = m$** since $q \to 1$ as $p \to 0$.

Next, $\mu_3 = npq(q-p) = m$ in the limit

and $\mu_4 = npq[1 + 3(n-2)pq] = npq[1 + 3npq - 6pq]$

$= m(1 + 3m - 0) = m + 3m^2$ in the limit.

Further, the coefficient of skewness $= \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1}{\sqrt{m}}$ in the limit.

and kurtosis $= \beta_2 = 3 + \frac{1-6pq}{npq} = 3 + \frac{1}{m}$ in the limit. **Ans.**

Also the moment generating function of Poisson's distribution is given by

$$\boxed{M_x = \sum_0^\infty e^{tx}e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_0^\infty \frac{(\lambda e^t)^x}{x!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t-1)}.}$$

EXAMPLE 31.14. (*a*) Suppose that on an average, 1 house in 1000 houses gets fire in a year in a district. If there are 2000 houses in that district, find the probability that exactly 5 houses will have fire during that year.

(*b*) If the probability of a bad reaction from a certain injection is 0.001, determine the chance that out of 1000 individuals more than two will have bed reactions. *[GGSIPU IV Sem II Term 2015]*

(*c*) Define the Poisson's distribution and show that for a Poisson distribution with mean m, we have

$$\mu_{r+1} = rm\,\mu_{r-1} + m\frac{d}{dm}\mu_r \quad \text{where} \quad \mu_r = \sum_{x=0}^{\infty} (x-m)^r \frac{e^{-m}\, m^x}{x!}$$

SOLUTION: (*a*) It is a Poisson distribution where $n = 2000$, $p = \dfrac{1}{1000}$ therefore, mean $= m = np = 2$. Further

$$P(r) = e^{-m}\frac{m^r}{r!} = e^{-2}\frac{2^r}{r!} \qquad (e = 2.7183)$$

$$\therefore \quad P(5) = 2.7183^{-2}\frac{2^5}{5!} = \frac{32}{120\,(2.7183)^2} = \frac{32}{120\,(7.389)} = 0.036. \qquad \textbf{Ans.}$$

(*b*) If is follows a Poisson's distribution with $p = 0.001, n = 1000$ hence mean $= np = 1 = m$

$$\text{Prob. } (X=0) = \frac{e^{-m}m^0}{0!}, \quad \text{Prob. } (X=1) = \frac{e^{-m}m^1}{1!} \quad \text{and Prob. } (X=2) = e^{-m}\frac{m^2}{2!}$$

$$\therefore \quad P(X = 0, 1, 2) = e^{-1}\left[1+1+\frac{1}{2}\right] = \frac{5}{2e} = \frac{2.5}{2.7183}$$

and
$$P(X>2) = 1 - \frac{2.5}{2.7182} = \frac{0.2183}{2.7183} \qquad \textbf{Ans.}$$

(*c*) We know that $\mu_r = \displaystyle\sum_{x=0}^{\infty} (x-m)^r \frac{e^{-m}\, m^x}{x!}$

$$\therefore \quad \frac{d}{dm}\mu_r = \sum_{x=0}^{\infty} \frac{1}{x!}\frac{d}{dm}[(x-m)^r e^{-m} m^x]$$

$$= \sum_{x=0}^{\infty}\frac{1}{x!} r(x-m)^{r-1}(-1)[e^{-m} m^x] + \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^r \frac{d}{dm}[e^{-m} m^x]$$

$$= -r\sum_{x=0}^{\infty}(x-m)^{r-1}\frac{e^{-m} m^x}{x!} + \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^r [e^{-m} x\, m^{x-1} - m^x e^{-m}]$$

$$= -r\sum_{x=0}^{\infty}(x-m)^{r-1}\frac{e^{-m} m^x}{x!} + \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^r [e^{-m}(x-m+m)\, m^{x-1} - m^x e^{-m}]$$

$$= -r\sum_{x=0}^{\infty}(x-m)^{r-1}\frac{e^{-m} m^x}{x!} + \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^{(r+1)} e^{-m}\cdot m^{x-1} + \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^r e^{-m} m^x - \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^r e^{-m}\cdot m^x$$

$$\text{or} \quad m\frac{d}{dm}\mu_r = -mr\sum_{x=0}^{\infty}(x-m)^{r-1}\frac{e^{-m} m^x}{x!} + \sum_{x=0}^{\infty}\frac{1}{x!}(x-m)^{r+1} e^{-m}\cdot m^x$$

$$\text{or} \quad m\frac{d}{dm}\mu_r = -mr\,\mu_{r-1} + \mu_{r+1} \qquad \textbf{Hence Proved.}$$

EXAMPLE 31.15. (*a*) A screw making machine produces on an average 2 defective screws out of 100 and packs them in tin boxes of 500. Find the probability that a box contains 15 defective screws.

(*b*) A manufacturer knows that the condensers he makes, contain on an average 1% defective. He packs them in boxes of 100. What is the probability that a box selected at random will contain 3 or more defective condensers.

SOLUTION: (*a*) Probability of a defective screw $= \dfrac{2}{100} = 0.02 = p$

Since $N = 500$ it follows a Poisson's distribution in which mean $= m = Np = 500 \times 0.02 = 10$

$\therefore$ Probability of getting 15 defective screws out of 500

$$= \frac{m^r e^{-m}}{r!} = \frac{10^{15} e^{-10}}{15!} \quad \text{which is the required probability.} \qquad \textbf{Ans.}$$

(*b*) Here the probability p that a condenser is defective = 0.01 which is very small and $n = 100$ is very large. It follows the Poisson's distribution.

$$\therefore \quad \text{Mean } m = np = 100 \times \frac{1}{100} = 1$$

Probability that three or more condensers are defective

= 1 – [Prob. of no defective + Prob. of 1 defective + Prob. of 2 defective]

$$= 1 - \left[e^{-m} + e^{-m}\cdot m + e^{-m}\cdot\frac{m^2}{2!}\right]_{m=1} = 1 - e^{-1}\left(1+1+\frac{1^2}{2}\right) = 1 - \frac{5}{2e} = 0.08 \text{ approx.}$$

EXAMPLE 31.16. (*a*) It has been found that 2% of the tools produced by a certain machine are defective. What is the probability that in a shipment of 400 such tools, 3% or more will be defective?

(*b*) In a factory manufacturing razor blades there is a small chance of 1/50 for any blade to be defective. The blades are placed in packets of 10 blades. Using Poisson distribution calculate the approximate number of packets containing not more than 2 defective blades in a consignment of 10,000 packets.

SOLUTION: (*a*) $p = \dfrac{2}{100} = 0.02$, $n = 400$ hence $\lambda = np = 8$. Here 3% of 400 = 12.

It is a case of Poisson's distribution.

$$P(X = r) = e^{-\lambda}\frac{\lambda^r}{r!} = e^{-8}\left(\frac{8^r}{r!}\right)$$

$$\therefore \quad P(X \geq 12) = 1 - [P(X=0) + P(X=1) + \ldots + P(X=11)]$$

$$= 1 - e^{-8}\left[1 + \frac{8}{1!} + \frac{8^2}{2!} + \ldots + \frac{8^{11}}{11!}\right] \qquad \textbf{Ans.}$$

(*b*) $N = 10{,}000$, $p = \frac{1}{50}$, $n = 10$. Hence mean $m = np = 0.2$.

Since $P(r) = e^{-m}\frac{m^r}{r!} = e^{-0.2}\frac{(0.2)^r}{r!}$ we have $P(r = 0) = e^{-0.2} = 0.8187$.

$\therefore$ Number of packets having no defective blade $= N[P(r = 0)] = 0.8187 \times 10{,}000 = 8187$.

Next, number of packets having one defective blade

$$= NP(r = 1) = N e^{-0.2}\frac{(0.2)}{1!} = 10{,}000 \times 0.2 \times e^{-0.2} = 2000 \times 0.8187 = 1637.4.$$

And the number of packets having two defective blades

$$= NP(r = 2) = N\left[\frac{e^{-0.2}(0.2)^r}{r!}\right]_{r=2} = \frac{e^{-0.2}\,0.04}{2!}(10{,}000) = 163.74.$$

$\therefore$ The approximate number of packets containing not more than two defective blades in a consignment of 10,000 packets, is equal to

$$10{,}000 - (8187 + 1637.4 + 163.74) = 11.86 = 12 \text{ approx.} \quad \textbf{Ans.}$$

EXAMPLE 31.17. (*a*) A manufacturer who produces medicine bottles, finds that 0.1% of the bottles are defective. The bottles are packed in boxes containing 500 bottles. Find that in 100 such boxes how many boxes are expected to contain (*i*) no deflective (*ii*) at least two defectives. *[GGSIPU IV Sem End Term 2015]*

(*b*) The distribution of the number of road accidents per day in a city is Poisson with mean 4. Find the number of days out of 100 days when there will be

(*i*) no accident,

(*ii*) atleast 2 accidents

(*iii*) atmost 3 accidents

(*iv*) between 2 and 5 accidents.

SOLUTION: (*a*) (*i*) $np = 500 \times 0.01 = 0.5 = m$. It is a case of Poisson's distribution.

Number of boxes containg no deflective bottles out of two boxes $= e^{-m} \times 100 = e^{-5}\,100 = 60.65$

(*ii*) Number of boxes containing at least two defective bottles out of 100 boxes

$$= 100\left[1 - e^{-0.5}\left(\frac{0.5^0}{0!} + \frac{0.5^1}{1!}\right)\right] = 100\left[1 - e^{-0.5}(1 + 0.5)\right] \qquad e^{-m}\frac{m^r}{r!}$$

$$= 100[1 - 0.6065(1.5)] = 100[1 - 0.9098] = 9.02 \quad \textbf{Ans.}$$

(*b*) The mean of Poisson distribution $= m = 4$.

Probability of r accidents in a day $= P(X = r) = \frac{e^{-m}m^r}{r!} = \frac{e^{-4}(4^r)}{r!}$.

(*i*) $P(X = 0) = \frac{e^{-4}(4^0)}{0!} = e^{-4} = 0.0183$

$\therefore$ Number of days on which no accident is there $= 100 \times 0.0183 = 1.83$

(*ii*) $P(X = 1) = \frac{e^{-4}(4^1)}{1!} = 4e^{-4} = 4(0.0183) = 0.0732$

Hence $P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)] = 1 - (0.0183 + 0.0732) = 0.9085$

$\therefore$ Number of days of at least 2 accidents $= 100 \times 0.9085 = 90.85$

(*iii*) $P(X = 3) = \frac{e^{-4}(4)^3}{3!} = \frac{64}{6}e^{-4} = \frac{32}{3}(0.0183) = 0.1952$

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

$$= \frac{e^{-4}4^0}{0!} + e^{-4}\frac{4^1}{1!} + e^{-4}\frac{4^2}{2!} + e^{-4}\frac{4^3}{3!} = e^{-4}\left[1 + 4 + \frac{16}{2} + \frac{64}{6}\right] = 0.0183[13 + 10.67]$$

$$= 0.0183(23.47) = 0.43347$$

$\therefore$ Number of days of atmost three accidents $= 100 \times 0.43347 = 43.347$

(*iv*) $P(2 \leq X \leq 5) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$$= e^{-4}\left[\frac{4^2}{2!} + \frac{4^3}{3!} + \frac{4^4}{4!} + \frac{4^5}{5!}\right] = e^{-4}\left[\frac{568}{15}\right] = 0.6954$$

$\therefore$ Number of days of accidents between 2 and 5 $= 100 \times 0.6954 = 69.54$ **Ans.**

EXAMPLE 31.18. (*a*) Fit a Poisson distribution to the following data to calculate the theoretical frequencies and compare with the actual ones:

x :	0	1	2	3	4
f :	122	60	15	2	1

(*b*) Fit a Poisson's distribution to the following frequency distribution and calculate the theoretical frequencies:

x	0	1	2	3	4
f	192	100	24	3	1

SOLUTION: (*a*) Mean $= \frac{\Sigma fx}{\Sigma f} = \frac{60 + 30 + 6 + 4}{200} = 0.5 = m =$ mean of Poisson's Distribution.

Theoretical frequency for r successes

$$= N\frac{e^{-m}m^r}{r!} = 200\frac{e^{-0.5}(0.5)^r}{r!} = \frac{200 \times 0.61}{2^r \times r!} \quad \text{where } r = 0, 1, 2, 3, 4.$$

$\therefore$ To compare with actual frequencies, the theoretical frequencies are as follows

x	0	1	2	3	4
f	121	60.5	15	2.5	0.31

Ans.

(*b*) Let us first calculate the mean of the given distribution.

x	f	xf
0	192	0
1	100	100
2	24	48
3	3	9
4	1	4

Here $\Sigma f = N = 320$ and $\Sigma xf = 161$ $\quad\therefore$ Mean $\bar{x} = m = 161/320 = 0.5031$

By Poisson's distribution $P(X = r) = \frac{e^{-m}m^r}{r!} = \frac{e^{-0.5031}}{r!}(0.5031)^r$

Theoretical Frequency for $x = 0$ is $Ne^{-m} = 320\,e^{-0.5031} = 193.4823$

For $x = 1$ it is $Ne^{-m}.\ m = 193.4823 \times 0.503 = 97.347$

For $x = 2$ it is $\dfrac{Ne^{-m}}{2!} m^2 = \dfrac{97.347 \times 0.503}{2} = 24.489$

For $x = 3$ it is $\dfrac{Ne^{-m} m^3}{3!} = \dfrac{24.489 \times 0.503}{3} = 4.107$

For $x = 4$ it is $\dfrac{Ne^{-m} m^4}{4!} = \dfrac{4.107 \times 0.503}{4} = 0.5166.$

Therefore, theoretical frequencies under Poisson's distribution are

x	0	1	2	3	4
f	193.4823	97.347	24.489	4.107	0.5166

Ans.

EXAMPLE 31.19. (*a*) An insurance company insures 4000 people against loss of both the eyes in a car accident. Based on previous data it was assumed that 10 persons out of 1,00,000 will have such type of injury in car accident. What is the probability that more than 2 of the insured will collect on their policy in a given year?

(*b*) If the variance of the Poisson's distribution is 2, find the probabilities for $r = 1, 2, 3$ and 4 from the recurrence relation of the Poisson distribution. Also find $P(r \geq 4)$.

$\sigma^2 = nqp = 2 = r$

SOLUTION: (*a*) Here $p = \dfrac{10}{100000} = \dfrac{1}{10000} = 0.0001.$

In this Poisson's distribution $\lambda = 4000\ (0.0001) = 0.4.$

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!} = \frac{e^{-0.4}(0.4)^r}{r!}$$

$$P(r > 2) = 1 - P(0) - P(1) - P(2)$$
$$= 1 - 0.67032 - 0.26812 - 0.05362 = 0.00794 \qquad \textbf{Ans.}$$

(*b*) Mean = Variance of the Poisson distribution = $\lambda = 2$.

Since $P(r) = e^{-\lambda}\dfrac{\lambda^2}{r!}$ we have the recurrence relation

$$P(r+1) = \frac{\lambda}{r+1}P(r) = \frac{2}{r+1}P(r). \text{ Here } P(0) = e^{-2} = 0.1353.$$

Since $P(1) = e^{-2}.\dfrac{2^1}{1!} = 2e^{-2} = 2(0.1353) = 0.2706$

$\therefore$ $P(2) = \dfrac{2}{2}P(1) = 0.2706,$

$P(3) = \dfrac{2}{3}P(2) = \dfrac{2}{3}(0.2706) = 0.1804$

and $P(4) = \dfrac{2}{4}P(3) = \dfrac{1}{2}(0.1804) = 0.092$

Now $P(r \geq 4) = 1 - [P(0) + P(1) + P(2) + P(3)]$
$= 1 - [0.1353 + 0.2706 + 0.2706 + 0.1804]$
$= 0.1431.$ **Ans.**

CONTINUOUS PROBABILITY DISTRIBUTIONS

The distributions defined by the continuous variable like heights or weights, are continuous distributions. In continuous distributions, in place of finding the probability that x equals some value, we find the probability of x falling in certain interval.

The probability distribution of a continuous variable x is defined by a function $f(x)$ such that the probability of x falling in the interval $x - \frac{1}{2}dx \le x \le x + \frac{1}{2}dx$ is $f(x)\,dx$, that is,

$$P\left(x - \frac{1}{2}dx \le x \le x + \frac{1}{2}dx\right) = f(x)\,dx.$$

This function $f(x)$ is called the **probability density function** and the continuous curve $y = f(x)$ is called the **probability curve**. The range of the variable may be finite or infinite. Even when the range is finite we consider it as infinite by assuming the density function to be zero outside the given range. Thus, if $\phi(x)$ is the density function in (a, b) then we can write it as

$$f(x) = \begin{cases} 0, & x < a \\ \phi(x), & a \le x \le b \\ 0, & x > b \end{cases}$$

An important point to be noted about the density function, is that

$$f(x) > 0 \quad \text{and} \quad \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

This can be understood as the fact that the total area under the probability curve and the X-axis is unity which can be interpreted as **'the total probability is always one'**.

If $\mathbf{F(x) = P(X \le x)} = \int_{-\infty}^{x} f(x)\,dx$ then $F(x)$ is called the **cumulative distribution function or simply the distribution function** of the continuous variable X.

MATHEMATICAL EXPECTATION IN CASE OF CONTINUOUS DISTRIBUTION

If X is a continuous random variable with probability density function $f(x)$, then the expected value of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx \quad \text{such that} \quad \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Thus, $E(x)$ is nothing but the **MEAN** of the random variable X and is usually denoted by μ.

The **VARIANCE** of X in terms of expectation, is given by

$$\sigma^2 = \begin{cases} E[(x-\mu)^2] = \Sigma(x_i - \mu)^2 . p_i & \text{for discrete distribution} \\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx & \text{for continuous distribution} \end{cases}$$

In general, the r^{th} moment about the mean μ, denoted by μ_r, is given by

$$\mu_r = \begin{cases} E(x-\mu)^r = \Sigma(x_i-\mu)^r p_i & \text{for discrete distribution} \\ \int_{-\infty}^{\infty} (x-\mu)^r f(x)\, dx & \text{for continuous distribution} \end{cases}$$

Some obvious but important properties of the distribution function $F(x)$, are

(*i*) $F'(x) = f(x) \geq 0$ hence $F(x)$ is an increasing function of x

(*ii*) $F(-\infty) = 0$

(*iii*) $F(\infty) = 1$

(*iv*) $P(a \leq x \leq b) = \int_a^b f(x)\, dx = \int_{-\infty}^b f(x)\, dx - \int_{-\infty}^a f(x)\, dx = F(b) - F(a).$

(*v*) If $f(x)$ is the probability density function of a continuous variable X, then the moment generating function of the continuous probability distribution about $x = a$, is

$$M_a(t) = \int_{-\infty}^{\infty} e^{t(x-a)} f(x)\, dx = e^{-at} \int_{-\infty}^{\infty} e^{tx} f(x)\, dx = e^{-at}.M_o(t).$$

EXAMPLE 31.20. (*a*) Show that the function $f(x)$, defined as

$$f(x) = e^{-x} \quad \text{for } x \geq 0$$
$$= 0 \quad \text{for } x < 0$$

is a probability density function and find the probability that the variate X having $f(x)$ as density function will lie in the interval (1, 2). Also find the probability distribution function F (2).

(*b*) Find the moment generating function for triangular distribution defined by

$$f(x) = \begin{cases} x, & 0 \leq x \leq 1 \\ 2-x, & 1 \leq x \leq 2 \end{cases}.$$

SOLUTION: (*a*) $\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{0} 0\, dx + \int_0^{\infty} e^{-x}\, dx = 0 - \left[e^{-x}\right]_0^{\infty} = 1$

Therefore $f(x)$ is the probability density function.

$$P(1 \leq X \leq 2) = \int_1^2 e^{-x}\, dx = \left[-e^{-x}\right]_1^2 = e^{-1} - e^{-2} = 0.233$$

Next, $$F(2) = \int_{-\infty}^{2} f(x)\, dx = \int_{-\infty}^{0} 0\, dx + \int_0^2 e^{-x}\, dx = 0 - \left[e^{-x}\right]_0^2 = 1 - e^{-2}$$

$$= 1 - 0.135 = 0.865.$$

Ans.

(*b*) We know that the m.g.f. of $f(x)$, $x \in (a, b)$, is equal to

$$\int_a^b t^{tx} f(x)\, dx = \int_0^1 e^{tx} x\, dx + \int_1^2 (2-x)\, e^{tx}\, dx$$

Normal Dist.
$z = \frac{x - np}{\sqrt{npq}}$

σ^2 : Variance
σ : $\sqrt{\text{Variance}}$: Standard Deviation

NORMAL DISTRIBUTION

Normal distribution is the limiting form of the binomial distribution when $n \to \infty$ but neither p nor q is very small. It is a continuous distribution of fundamental importance. Let us define a variable z as $\boxed{z = \frac{x - np}{\sqrt{npq}} = \frac{x - \mu}{\sigma}}$ where x is a continuous random variable with two parameters, mean $\mu\,(= np)$ and standard deviation $\sigma\,(= \sqrt{npq})$. Thus, z is a variate with mean 0 and standard deviation unity. In the limit when $n \to \infty$ the distribution of z becomes continuous in $(-\infty, \infty)$. Its form is

$$y = P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} = \boxed{\frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2}$$

where $P(x)$ is the probability density function and this curve $y = P(x)$ is called *normal curve*.

PROPERTIES OF THE NORMAL DISTRIBUTION

Here we shall, analytically obtain the mean, median, mode, mean deviation, standard derivation and other moments of the normal distribution.

1. **Mean:** By definition, we have

$z = \frac{x - np}{\sqrt{npq}}$

$$\text{mean} = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}\, x\, dx \qquad \left(\text{Putting } \frac{x-\mu}{\sigma} = z\right)$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} (\mu + \sigma z)\, \sigma\, dz$$

$z = \frac{x - \mu}{\sigma}$

$$= \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, dz + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2}\, z\, dz$$

$$= \frac{2\mu}{\sqrt{2\pi}} \int_0^{\infty} e^{-z^2/2}\, dz + 0 \quad \text{(by property of definite integrals)}$$

$$= \mu\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-z^2/2}\, dz \quad (\text{now putting } z = \sqrt{2t}\,)$$

$$= \mu\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-t} \sqrt{2} \cdot \frac{1}{2\sqrt{t}}\, dt = \frac{\mu}{\sqrt{\pi}} \int_0^{\infty} e^{-t}\, t^{1/2 - 1}\, dt$$

$$= \frac{\mu}{\sqrt{\pi}} \left\lceil \tfrac{1}{2} \right. = \mu. \qquad \left(\text{as } \left\lceil \tfrac{1}{2} \right. = \sqrt{\pi}\right).$$

2. **Mode:**

We know that mode is that value of variate x for which y is maximum, that is, $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ is maximum. This is attained when $\frac{(x-\mu)^2}{2\sigma^2}$ is minimum, that is, when $x = \mu$. Thus, mode of the normal distribution is at $\boxed{x = \mu}$ and modal ordinate $= \frac{1}{\sigma\sqrt{2\pi}}$.

3. **Median:**

We know that the median divides the total area under the normal curve into two equal ...

$$\text{The total area} = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{(x-\mu)^2}{\sigma}\right)} dx = 1 \quad \text{as witnessed while obtaining mean.}$$

$$\text{Let us calculate} \int_{-\infty}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \frac{1}{2}$$

Thus the ordinate at $x = \mu$ divides the total area under the curve into two equal halves, therefore, the median of the distribution is at $\boxed{x = \mu.}$

4. **Mean Deviation:**

$$\text{Mean deviation} = \int_{-\infty}^{\infty} \frac{|x - \mu|}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad \left(\text{Putting } \frac{x-\mu}{\sigma} = z\right)$$

$$= \int_{-\infty}^{\infty} \frac{\sigma |z|}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \frac{2\sigma}{\sqrt{2\pi}} \int_0^{\infty} z\, e^{-z^2/2}\, dz \qquad \left(\text{Putting } \frac{z^2}{2} = t\right)$$

$$= \sigma\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-t}\, dt = -\sigma\sqrt{\frac{2}{\pi}} \left[e^{-t}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}\, \sigma = 0.7979\sigma = \frac{4}{5}\sigma \text{ approx.}$$

Therefore, the mean deviation of the normal distribution $= \sqrt{\frac{2}{\pi}}\, \sigma = \frac{4}{5}\sigma$ approx.

5. **Standard Derivation:**

$$\text{Variance} = \mu_2 = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot (x - \mu)^2\, dx \qquad \left(\text{Putting } \frac{x-\mu}{\sigma} = z, \text{ we get}\right)$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} \sigma^2 z^2 \cdot \sigma\, dz = \frac{2}{\sqrt{2\pi}} \sigma^2 \int_0^{\infty} z^2 e^{-z^2/2}\, dz \qquad \text{putting } z^2 = 2t,$$

$$= \sqrt{\frac{2}{\pi}} \sigma^2 \int_0^{\infty} \sqrt{2t}\, e^{-t}\, dt = \frac{2}{\sqrt{\pi}} \sigma^2 \int_0^{\infty} e^{-t} \cdot t^{\frac{3}{2} - 1}\, dt$$

Std. Deviation $= \sqrt{\text{variance}}$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \left\lceil \tfrac{3}{2} \right. = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \left\lceil \tfrac{1}{2} \right. = \sigma^2. \qquad \left(\text{as } \left\lceil \tfrac{1}{2} \right. = \sqrt{\pi}\right).$$

Therefore, standard deviation of the normal distribution $= \sigma$.

6. The first quartile Q_1 and the third quartile Q_3 are **equidistant from the mean.**
7. **The points of Inflexion of the normal curve.**

On putting $\frac{d^2 y}{dx^2} = 0$ such that $\frac{d^3 y}{dx^3} \neq 0$, we get the points of inflexion at $x = \mu \pm \sigma$, which are equidistant from the mean.

8. Moments about the mean of the normal distribution.

$$\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n+1} \sigma^{2n+1} e^{-z^2/2} dz \quad \left(\text{on putting } \frac{x-\mu}{\sigma} = z\right)$$

$= 0$ since the integrand is an odd function of z.

Here all the odd order moments about the mean are zero.

And $$\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{\sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n} e^{-z^2/2} dz \quad \left(\text{again putting } \frac{x-\mu}{\sigma} = z\right)$$

$$= \sqrt{\frac{2}{\pi}} \sigma^{2n} \int_0^{\infty} z^{2n} e^{-z^2/2} dz = \sqrt{\frac{2}{\pi}} \sigma^{2n} \cdot 2^{n-1/2} \int_0^{\infty} e^{-t} \cdot t^{n-1/2} dt \quad (\text{now putting } z^2 = 2t)$$

$$= \frac{2^n \sigma^{2n}}{\sqrt{\pi}} \int_0^{\infty} e^{-t} t^{n+\frac{1}{2}-1} dt = \frac{2^n \sigma^{2n}}{\sqrt{\pi}} \Gamma\left(n+\frac{1}{2}\right)$$

$$= \frac{2^n \sigma^{2n}}{\sqrt{\pi}} \left(n-\frac{1}{2}\right) \Gamma\left(n-\frac{1}{2}\right) = 2^{n-1} \sigma^2 \cdot \sigma^{2n-2} (2n-1) \Gamma\left(n-\frac{1}{2}\right)$$

$$= (2n-1)\, \sigma^2 \mu_{2n-2}$$

$$= (2n-1)(2n-3)\, \sigma^4 \mu_{2n-4} = (2n-1)(2n-3)(2n-5)\, \sigma^6 \mu_{2n-6}$$

$$= (2n-1)(2n-3)(2n-5) \ldots 3.1\ \sigma^{2n-2} \mu_2$$

$$= (2n-1)(2n-3)(2n-5) \ldots 5.3.1\ \sigma^{2n} \quad (\text{as } \mu_2 = \sigma^2).$$

9. Skewness and Kurtosis of the Normal Distribution

Here $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0$ as $\mu_3 = 0$ and $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3 \cdot \sigma^4}{(\sigma^2)^2} = 3.$

Therefore the coefficient of skewness is zero for the normal curve, that is, it is not skewed anywhere. Also since the kurtosis $= \beta_2 = 3$, the curve is **mesokurtic.**

10. Shape of the Normal Curve.

Here mean = median = mode = μ, the curve is uninodal and perfectly symmetrical about the mean and ordinates are decreasing fast on both sides of mean. It is better called **bell-shaped** and X-axis is an asymptote to the normal probability curve.

11. Moment Generating Function (M.G.F.) of the Normal Distribution

Since $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $-\infty < x < \infty$, the moment generating function about the origin is given by

$$M_0(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx = \int_{-\infty}^{\infty} \frac{e^{tx}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad \left(\text{putting } \frac{x-\mu}{\sigma} = z\right)$$

$$M_0(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} e^{t(\mu+\sigma z)} \cdot \sigma\, dz = \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2t\sigma z)} \cdot \sigma\, dz$$

$$= \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t\sigma)^2} \cdot e^{t^2\sigma^2/2} dz = \frac{e^{(t\mu + t^2\sigma^2/2)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t\sigma)^2} dz$$

$$= \frac{e^{(t\mu+t^2\sigma^2/2)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{\frac{2}{\pi}} e^{(t\mu+t^2\sigma^2/2)} \int_0^{\infty} e^{-y^2/2} dy \quad (\text{on putting } z - t\sigma = y)$$

$$= \sqrt{\frac{2}{\pi}} e^{(t\mu+t^2\sigma^2/2)} \cdot \sqrt{\frac{\pi}{2}} \quad (\text{as done earlier too})$$

Thus, $M_0(t) = e^{(t\mu + t^2\sigma^2/2)}$.

12. Standard Normal Probability Curve.

The normal curve is $y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$.

If we make the area under this curve equal to one we arrive at a normal distribution which is independent of N, μ and σ. The normal distribution in this form is called **unit normal distribution or standard normal distribution** and the new normal curve is

$$y = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad \left(z = \frac{x-\mu}{\sigma}\right)$$

which is called *unit normal probability curve* or **standard normal probability curve.**

13. Probability Integral.

The probability of x lying between x_1 and x_2 is equal to the area under the normal curve from x_1 to x_2 and the X-axis.

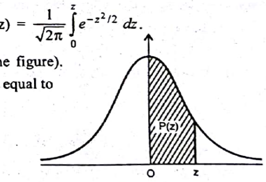
$$P(x_1 < x < x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-(x-\mu)^2/2\sigma^2} dx = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-z^2/2} dz \quad \left(\text{where } z = \frac{x-\mu}{\sigma}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{z_2} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{z_1} e^{-z^2/2} dz$$

where $z_1 = \frac{x_1 - \mu}{\sigma}$ and $z_2 = \frac{x_2 - \mu}{\sigma}$.

$\therefore$ $P(x_1 < x < x_2) = P(z_2) - P(z_1)$ where $P(z) = \frac{1}{\sqrt{2\pi}} \int_0^{z} e^{-z^2/2} dz$.

This P (z) is called the **probability integral** (see the figure). Therefore, the total area under the normal curve is equal to



$$P(-\infty < z < \infty) = 2P(0 < z < \infty)$$

$$= 2P(\infty) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-z^2/2} dz$$

$= 2(0.5) = 1$ as seen from the standard normal distribution table.

4. Area property of the normal distribution, is

$$P(\mu - \sigma < x < \mu + \sigma) = P(-1 < z < 1) = 0.6826$$

$$P(\mu - 2\sigma < x < \mu + 2\sigma) = P(-2 < z < 2) = 0.9544$$

$$P(\mu - 3\sigma < x < \mu + 3\sigma) = P(-3 < z < 3) = 0.9973.$$

EXAMPLE 31.22. A normal curve has mean $\bar{x} = 20$ and the variance $= 100$. Find the area between $x = 26$, $x = 38$ and also between $x = 15$, $x = 40$.

μ=

SOLUTION: mean $= \bar{x} = 20$, $\sigma = 10$. The normal variate $z = \dfrac{x - \mu}{\sigma}$

$\therefore$ For $x = 26$, $z = \dfrac{26 - 20}{10} = 0.6$ and for $x = 38$, $z = \dfrac{38 - 20}{10} = 1.8$.

From the Table No. 1 (Appendix) P (0.6) = Prob. $(0 \le z \le 0.6) = 0.2257$

and P (1.8) = Prob. $(0 \le z \le 1.8) = 0.4641$.

$\therefore$ Area under the curve from $x = 26$ to $x = 38$ is equal to the difference between two areas

$= 0.4641 - 0.2257 = 0.2384$.

Next, from the Table No. 1 (Appendix) for $x = 15$, $z = \dfrac{15 - 20}{10} = -0.50$

and for $x = 40$, $z = \dfrac{40 - 20}{10} = +2.00$

Since the two z's are on opposite sides of the mean the area between them has to be taken as sum of their tabular values.

Areas corresponding to $z = 0.50$ and $z = 2.00$ are 0.1915 and 0.4772, respectively, therefore the required area $= 0.1915 + 0.4772 = 0.6687$. **Ans.**

EXAMPLE 31.23. In a normal distribution, 7% of the items are under 35 and 89% are under 63. Determine the mean and variance of the distribution. Given that $P(0 \le z \le 0.18) = 0.07$ and $P(0 \le z \le 1.48) = 0.43$, $P(0 \le z \le 1.23) = 0.39$.

SOLUTION: Since 7% of items are under 35, hence 43% of the items lie between mean $\bar{X}$ and 35. From the Table No. 1 (Appendix) the value of z for the area 0.43 is 1.48. For it we should take

$$z = -1.48 \text{ hence } \frac{35 - \bar{X}}{\sigma} = -1.48 \qquad \text{...(1)}$$

Next, since 89% items are under 63 it implies that 39% items lie between $x = \bar{X}$ and $x = 63$ for which $z = +1.23$. $\therefore$ $\dfrac{63 - \bar{X}}{\sigma} = 1.23$ ...(2)

From (1) and (2) we have $\bar{X} - 1.48\,\sigma = 35$, and $\bar{X} + 1.23\,\sigma = 63$

which gives $\sigma = 10.33$ and $\bar{X} = 50.3$. **Ans.**

Therefore the mean and standard deviation of the distribution are 50. 3 and 10.33.

EXAMPLE 31.24. The income of a group of 10,000 persons was found to be normally distributed with mean Rs. 750 p.m. and standard deviation Rs. 50. show that, in this group about 95% had income exceeding Rs. 668 and only 5% had income exceeding Rs 832. Also find the lowest income among the richest 100.

SOLUTION: $\bar{X} = 750$, $\sigma = 50$, $X = 668$ hence $z = \dfrac{668 - 750}{50} = -1.64$.

Area to the right of the ordinate at -1.64 is $0.4495 + 0.50 = 0.9495$.

Therefore, the expected number of persons getting above Rs. 668 = 95% approximately of 10,000 = 9500.

Next, the standard normal variate corresponding to 832 is

$$z = \frac{832 - 750}{50} = \frac{82}{50} = 1.64.$$

The area on the right of the ordinate at $z = 1.64$ is $0.50000 - 0.4495 = 0.0505 = 5\%$ approx.

Also the number of persons getting Rs. 832 and above, is $= 10{,}000 \times 0.0505 = 505$.

Next, let x be the lowest income among the 100 richest persons. Richest 100 = 1% of 10,000 the total number of persons. Here $z = \dfrac{x - \mu}{\sigma} = \dfrac{x - 750}{50}$. Area on the right of x is 0.01.

Area between $z = 0$ and z(to be obtained) $= 0.5 - 0.01 = 0.49$. Corresponding to this area the value of z from the Table No. I (Appendix) = 2.33.

$$2.33 = \frac{x - 750}{50} \quad \text{or} \quad x = 750 + 50\,(2.33) = 866.5.$$

$\therefore$ Lowest income among the 100 richest persons = Rs. 866.5. **Ans.**

EXAMPLE 31.25. In a normal distribution of marks, 31% are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution. It is given that if

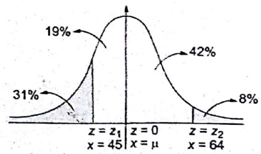
$$f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-x^2/2}\, dx \quad \text{then } f(0.5) = 0.19 \text{ and } f(1.4) = 0.42.$$

SOLUTION: Let μ and σ denote the mean and standard deviation of the distribution respectively. Since 31% of marks are under 45, the area to the left of the ordinate at $x = 45$ is 0.31. Also since 8% of marks are over 64, the area to the right of the ordinate at $x = 64$ is 0.08.

Let $z = \dfrac{x - \mu}{\sigma}$ we are to find values of z when $x = 45$ and $x = 64$.

Since $f(0.5) = 0.19$ we have $\dfrac{45 - \mu}{\sigma} = -0.5$

and since $f(1.4) = 0.42$ we have $\dfrac{64 - \mu}{\sigma} = 1.4$



Solving the above equations for μ and σ, we get

$\mu = 50$ and $\sigma = 10$. **Ans.**