4.35 THEORETICAL PROBABILITY DISTRIBUTIONS

Generally, frequency distribution are formed from the observed or experimental data. However, frequency distribution of certain populations can be deduced mathematically by fitting theoretical probability distribution under certain assumptions.

Frequency distributions can be classified under two heads:

(*i*) Observed Frequency Distributions.

(*ii*) Theoretical or Expected Frequency Distributions.

Observed frequency distributions are based on actual observation and experimentation. If certain hypothesis is assumed, it is sometimes possible to derive mathematically what the frequency distribution of certain universe should be. Such distributions are called **Theoretical Distributions.**

Theoretical probability distributions are of two types:

(*i*) Discrete probability distribution. Binomial, poisson, geometric, negative binomial, hypergeometric, multinomial, multivariate hypergeometric distributions.

(*ii*) Continuous probability distributions.

Uniform, normal Gamma, exponential, χ^2, Beta, bivariate normal, t, F-distributions.

Here, we will study three important theoretical probability distributions:

1. Binomial Distribution (or Bernoulli's Distribution)
2. Poisson's Distribution
3. Normal Distribution.

4.36 BINOMIAL PROBABILITY DISTRIBUTION

(*A.K.T.U. 2018*)

It was discovered by a Swiss Mathematician Jacob James Bernoulli in the year 1700.

This distribution is concerned with trials of a repretitive nature in which only the occurrence or non-occurrence, success or failure, acceptance or rejection, yes or no of a particular event is of interest.

For convenience, we shall call the occurrence of the event 'a success' and its non-occurrence 'a failure'.

Let there be n independent trials in an experiment. Let a random variable X denote the number of successes in these n trials. Let p be the probability of a success and q that of a failure in a single trial so that $p + q = 1$. Let the trials be independent and p be constant for every trial.

Let us find the probability of r successes in n trials.

r successes can be obtained in n trials in ${}^nC_r$ ways.

$$\therefore \quad P(X = r) = {}^nC_r P(\underbrace{S\,S\,S\ldots\ldots S}_{r \text{ times}}\quad \underbrace{F\,F\,F\ldots\ldots F}_{(n-r)\text{ times}})$$

$$= {}^nC_r \underbrace{P(S)\,P(S)\ldots\ldots P(S)}_{r \text{ factors}}\quad \underbrace{P(F)\,P(F)\ldots\ldots P(F)}_{(n-r)\text{ factors}}$$

$$= {}^nC_r \underbrace{p\,p\,p\ldots\ldots p}_{r \text{ factors}}\quad \underbrace{q\,q\,q\ldots\ldots q}_{(n-r)\text{ factors}} \quad \ldots(1)$$

$$= {}^nC_r p^r q^{n-r}$$

Hence $\boxed{P(X = r) = {}^nC_r\, p^r q^{n-r}}$ where $p + q = 1$ and $r = 0, 1, 2, ..., n$.

The distribution (1) is called the *binomial probability distribution* and X is called the *binomial variate*.

Note 1. $P(X = r)$ is usually written as $P(r)$.

Note 2. The successive probabilities $P(r)$ in (1) for $r = 0, 1, 2, ..., n$ are

$$ {}^nC_0 q^n,\ {}^nC_1 q^{n-1} p,\ {}^nC_2 q^{n-2} p^2,\ ......,\ {}^nC_n p^n $$

which are the successive terms of the binomial expansion of $(q + p)^n$. That is why this distribution is called "binomial" distribution.

Note 3. n and p occurring in the binomial distribution are called the *parameters* of the distribution.

Note 4. In a binomial distribution:

(*i*) n, the number of trials is finite.

(*ii*) each trial has only two possible outcomes usually called success and failure.

(*iii*) all the trials are independent.

(*iv*) p (and hence q) is constant for all the trials.

4.37 RECURRENCE OR RECURSION FORMULA FOR THE BINOMIAL DISTRIBUTION

In a binomial distribution,

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

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

$$ \therefore \quad \frac{P(r+1)}{P(r)} = \frac{(n-r)!}{(n-r-1)!} \times \frac{r!}{(r+1)!} \times \frac{p}{q} $$

$$ = \frac{(n-r)\times(n-r-1)!}{(n-r-1)!} \times \frac{r!}{(r+1)\times r!} \times \frac{p}{q} = \left(\frac{n-r}{r+1}\right) \cdot \frac{p}{q} $$

$$ \Rightarrow \quad \boxed{P(r+1) = \frac{n-r}{r+1} \cdot \frac{p}{q}\, P(r)} $$

which is the required recurrence formula. Applying this formula successively, we can find P(1), P(2), P(3), ..., if P(0) is known.

4.38 MEAN AND VARIANCE OF THE BINOMIAL DISTRIBUTION [A.K.T.U. 2018]

For the binomial distribution, $P(r) = {}^nC_r\, q^{n-r} p^r$

$$ \text{Mean } \mu = \sum_{r=0}^{n} r P(r) = \sum_{r=0}^{n} r \,.\, {}^nC_r q^{n-r} p^r $$

$$ = 0 + 1 \,.\, {}^nC_1\, q^{n-1} p + 2 \,.\, {}^nC_2\, q^{n-2} p^2 + 3 \,.\, {}^nC_3\, q^{n-3} p^3 + ... + n \,.\, {}^nC_n\, p^n $$

$$ = nq^{n-1} p + 2 \,.\, \frac{n(n-1)}{2\cdot 1}\, q^{n-2} p^2 + 3 \,.\, \frac{n(n-1)(n-2)}{3\cdot 2\cdot 1}\, q^{n-3} p^3 + ... + n.p^n $$

$$= nq^{n-1}p + n(n-1)\,q^{n-2}p^2 + \frac{n(n-1)(n-2)}{2.1}\,q^{n-3}p^3 + \ldots\ldots + np^n$$

$$= np[{}^{n-1}C_0 q^{n-1} + {}^{n-1}C_1 q^{n-2}p + {}^{n-1}C_2 q^{n-3}p^2 + \ldots + {}^{n-1}C_{n-1}p^{n-1}]$$

$$= np(q+p)^{n-1} = np \qquad (\because\ p+q=1)$$

Hence mean of the binomial distribution is np.

$$\text{Variance } \sigma^2 = \sum_{r=0}^{n} r^2 P(r) - \mu^2 = \sum_{r=0}^{n} [r + r(r-1)]\,P(r) - \mu^2$$

$$= \sum_{r=0}^{n} rP(r) + \sum_{r=0}^{n} r(r-1)\,P(r) - \mu^2 = \mu + \sum_{r=2}^{n} r(r-1)\,{}^nC_r q^{n-r}p^r - \mu^2$$

(since the contribution due to $r = 0$ and $r = 1$ is zero)

$$= \mu + [2\cdot 1\cdot {}^nC_2 q^{n-2}p^2 + 3\cdot 2\cdot {}^nC_3 q^{n-3}p^3 + \ldots + n(n-1)\,{}^nC_n p^n] - \mu^2$$

$$= \mu + \left[2\cdot 1\cdot \frac{n(n-1)}{2.1}\,q^{n-2}p^2 + 3\cdot 2\cdot \frac{n(n-1)(n-2)}{3.2.1}\,q^{n-3}p^3 + \ldots + n(n-1)p^n\right] - \mu^2$$

$$= \mu + [n(n-1)q^{n-2}p^2 + n(n-1)(n-2)q^{n-3}p^3 + \ldots + n(n-1)p^n] - \mu^2$$

$$= \mu + n(n-1)\,p^2[q^{n-2} + (n-2)q^{n-3}p + \ldots + p^{n-2}] - \mu^2$$

$$= \mu + n\,(n-1)p^2\,[{}^{n-2}C_0 q^{n-2} + {}^{n-2}C_1 q^{n-3}p + \ldots + {}^{n-2}C_{n-2}p^{n-2}] - \mu^2$$

$$= \mu + n(n-1)p^2(q+p)^{n-2} - \mu^2 = \mu + n(n-1)p^2 - \mu^2 \qquad [\because\ q+p=1]$$

$$= np + n(n-1)p^2 - n^2p^2 = np[1-p] = npq. \qquad [\because\ \mu = np]$$

Hence the variance of the binomial distribution is npq.

Standard deviation of the binomial distribution is $\sqrt{npq}$.

4.39 MOMENT GENERATING FUNCTION OF BINOMIAL DISTRIBUTION [A.K.T.U. 2018]

1. About origin

$$M_x(t) = E(e^{tx}) = \sum_{x=0}^{n} e^{tx}\,{}^nC_x\,p^x\,q^{n-x} = \sum_{x=0}^{n} {}^nC_x\,(pe^t)^x\,q^{n-x} = (q+pe^t)^n$$

2. About mean [G.B.T.U. 2012, U.P.T.U. 2015]

$$M_{x-np}(t) = E[e^{t(x-np)}]$$

$$= e^{-npt}E(e^{tx}) = e^{-npt}M_x(t) = e^{-npt}(q+pe^t)^n$$

$$= (qe^{-pt} + pe^{t-pt})^n = (qe^{-pt} + pe^{qt})^n \qquad |\because\ 1-p=q$$

4.40 MOMENTS ABOUT MEAN OF BINOMIAL DISTRIBUTION

$$M_{x-np}(t) = (qe^{-pt} + pe^{qt})^n$$

$$= \left[q\left(1 - pt + \frac{p^2t^2}{2!} - \frac{p^3t^3}{3!} + \ldots\right) + p\left(1 + qt + \frac{q^2t^2}{2!} + \frac{q^3t^3}{3!} + \ldots\right)\right]^n$$

$$= \left[(q+p) + \frac{t^2}{2!}\,pq\,(q+p) + \frac{t^3}{3!}\,pq\,(q^2-p^2) + \frac{t^4}{4!}\,pq\,(q^3+p^3) + \ldots\right]^n$$

$$= \left[1 + \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq(q-p) + \frac{t^4}{4!} qp(1-3pq) + \ldots\right\}\right]^n$$

$$= \left[1 + {}^nC_1 \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq(q-p) + \frac{t^4}{4!} pq(1-3pq) + \ldots\right\}\right.$$

$$\left. + {}^nC_2 \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq(q-p) + \ldots\right\}^2 + \ldots\right]$$

Now,

$$\mu_2 = \text{coefficient of } \frac{t^2}{2!} = npq$$

$$\mu_3 = \text{coefficient of } \frac{t^3}{3!} = npq(q-p)$$

$$\mu_4 = \text{coefficient of } \frac{t^4}{4!} = npq(1-3pq) + 3n(n-1)p^2q^2$$

$$= 3n^2p^2q^2 + npq(1-6pq)$$

Hence,

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq}$$

$$\therefore \quad \gamma_1 = \frac{1-2p}{\sqrt{npq}}$$

and

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1-6pq}{npq}$$

$$\therefore \quad \gamma_2 = \frac{1-6pq}{npq}$$

Note 1. $\gamma_1 = \dfrac{1-2p}{\sqrt{npq}}$ gives a **measure of skewness** of the binomial distribution. If $p < \frac{1}{2}$, skewness is positive, if $p > \frac{1}{2}$, skewness is negative and if $p = \frac{1}{2}$, it is zero.

$\beta_2 = 3 + \dfrac{1-6pq}{npq}$ gives a **measure of the kurtosis** of the binomial distribution.

Note 2. If n independent trials constitute one experiment and this experiment is repeated N times then the frequency of r successes is $N \cdot {}^nC_r p^r q^{n-r}$.

4.41 APPLICATIONS OF BINOMIAL DISTRIBUTION

1. In problems concerning no. of defectives in a sample production line.
2. In estimation of reliability of systems.

$$= \left[1 + \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq\,(q-p) + \frac{t^4}{4!} qp\,(1-3pq) + \ldots\right\}\right]^n$$

$$= \left[1 + {}^nC_1 \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq\,(q-p) + \frac{t^4}{4!} pq\,(1-3pq) + \ldots\right\}\right.$$

$$\left. + {}^nC_2 \left\{\frac{t^2}{2!} \cdot pq + \frac{t^3}{3!} pq\,(q-p) + \ldots\right\}^2 + \ldots\right]$$

Now,

$$\mu_2 = \text{coefficient of } \frac{t^2}{2!} = npq$$

$$\mu_3 = \text{coefficient of } \frac{t^3}{3!} = npq\,(q-p)$$

$$\mu_4 = \text{coefficient of } \frac{t^4}{4!} = npq\,(1-3\,pq) + 3n(n-1)\,p^2q^2$$

$$= 3n^2p^2q^2 + npq\,(1-6pq)$$

Hence,

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(q-p)^2}{npq} = \frac{(1-2p)^2}{npq}$$

$$\therefore \qquad \gamma_1 = \frac{1-2p}{\sqrt{npq}}$$

and

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1-6pq}{npq}$$

$$\therefore \qquad \gamma_2 = \frac{1-6pq}{npq}$$

Note 1. $\gamma_1 = \dfrac{1-2p}{\sqrt{npq}}$ gives a **measure of skewness** of the binomial distribution. If $p < \frac{1}{2}$, skewness is positive, if $p > \frac{1}{2}$, skewness is negative and if $p = \frac{1}{2}$, it is zero.

$\beta_2 = 3 + \dfrac{1-6pq}{npq}$ gives a **measure of the kurtosis** of the binomial distribution.

Note 2. If n independent trials constitute one experiment and this experiment is repeated N times then the frequency of r successes is $N \cdot {}^nC_r\, p^r\, q^{n-r}$.

4.41 APPLICATIONS OF BINOMIAL DISTRIBUTION

1. In problems concerning no. of defectives in a sample production line.
2. In estimation of reliability of systems.

3. No. of rounds fired from a gun hitting a target.
4. In Radar detection.

ILLUSTRATIVE EXAMPLES

Example 1. (*i*) *Comment on the following statement:*

For a Binomial distribution, mean is 6 and variance is 9.

(*ii*) *A die is tossed thrice. A success is getting 1 or 6 on a toss. Find the mean and variance of the number of success.*

Sol. (*i*) $$\mu = np = 6 \quad \text{...(1)}$$
$$\sigma^2 = npq = 9 \quad \text{...(2)}$$

Dividing (2) by (1), we get

$$q = \frac{9}{6} = 1.5$$

which is impossible as $0 \le q \le 1$

$\therefore$ The above statement is **False.**

(*ii*) Prob. of getting success (1 or 6) on a toss $= \frac{2}{6} = \frac{1}{3} = p$

$$\therefore \quad q = 1 - \frac{1}{3} = \frac{2}{3}$$

No. of tosses of a die, $n = 3$

(*i*) Mean $= np = 3\left(\frac{1}{3}\right) = 1.$ (*ii*) Variance $= npq = (3)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{3}.$

Example 2. *If 10% of the bolts produced by a machine are defective, determine the probability that out of 10 bolts choosen at random*

(*i*) *1* (*ii*) *None* (*iii*) *at most 2 bolts will be defective.*

Sol. Here, $p(\text{defective}) = \frac{10}{100} = \frac{1}{10}$ (given)

$$\therefore \quad q(\text{non-defective}) = 1 - \frac{1}{10} = \frac{9}{10}$$

Also, $n = 10$, (n is no. of bolts chosen). (given)

The probability of r defective bolts out of n bolts chosen at random is given by

$$P(r) = {}^nC_r \, p^r \, q^{n-r} \quad \text{...(1)}$$

(*i*) Here $r = 1$,

$$\therefore \quad P(1) = {}^{10}C_1 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^{10-1} \quad | \text{ Using (1)}$$

$$= 10\left(\frac{1}{10}\right)\left(\frac{9}{10}\right)^9 = (.9)^9 = 0.3874 \quad \text{...(2)}$$

(ii) Here $r = 0$

$\therefore$ $$P(0) = {}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{10-0} = \left(\frac{9}{10}\right)^{10} = 0.3486 \quad \text{...(3) | Using (1)}$$

(iii) Prob. that at most 2 bolts will be defective $= P(r \le 2) = P(0) + P(1) + P(2)$ ...(4)

Now, $$P(2) = {}^{10}C_2 \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^{10-2} \quad \text{| Using (1)}$$

$$= 45\left(\frac{1}{100}\right)(0.43046) = 0.1937$$

$\therefore$ From (4), Required Probability $= P(0) + P(1) + P(2)$

$= 0.3486 + 0.3874 + 0.1937 = 0.9297.$

Example 3. *A binomial variable X satisfies the relation 9P (X = 4) = P(X = 2) when n = 6. Find the value of the parameter p and P(X = 1).*

Sol. We know that

$$P(X = r) = {}^nC_r\, p^r q^{n-r} \quad \text{...(1)}$$

$\therefore$ $$P(X = 4) = {}^6C_4\, p^4 q^2 = 15p^4q^2$$

and $$P(X = 2) = {}^6C_2 p^2 q^4 = 15\, p^2 q^4 \quad \text{| Since } n = 6$$

The given relation is

$$9\, P(X = 4) = P(X = 2) \quad \Rightarrow \quad 9(15p^4 q^2) = 15p^2q^4$$

$$\Rightarrow \quad 9p^2 = q^2 = (1-p)^2 \quad | \because \; p + q = 1$$

$$\Rightarrow \quad 9p^2 = 1 + p^2 - 2p$$

$$\Rightarrow \quad 8p^2 + 2p - 1 = 0 \quad \Rightarrow \quad (4p-1)(2p+1) = 0$$

$\therefore$ $$\boxed{p = \frac{1}{4}} \quad | \because \; p \text{ cannot be negative}$$

Now, $$P(X = 1) = {}^6C_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^5 = .3559. \quad \left| \because \; q = 1 - p = \frac{3}{4}\right.$$

Example 4. *Fit a binomial distribution to the following frequency data:*

x :	*0*	*1*	*2*	*3*	*4*
f :	*30*	*62*	*46*	*10*	*2*

Sol. The table is as follows:

x	f	fx
0	30	0
1	62	62
2	46	92
3	10	30
4	2	8
	$\Sigma f = 150$	$\Sigma fx = 192$

Mean of observations $= \dfrac{\Sigma fx}{\Sigma f} = \dfrac{192}{150} = 1.28$

$np = 1.28$ (*n* is no. of trials)

$\Rightarrow \quad 4p = 1.28$

$\Rightarrow \quad p = 0.32$

$\Rightarrow \quad q = 1 - p = 1 - 0.32 = 0.68$

$\therefore \quad N = 150$ $\quad | \because N = 4$

Also,

Hence the binomial distribution is $= N(q+p)^n = 150\,(0.68 + 0.32)^4$.

Example 5. *A student is given a true-false examination with 8 questions. If he corrects at least 7 questions, he passes the examination. Find the probability that he will pass given that he guesses all questions.*

Sol. Here, n = no. of questions asked = 8

$$p = \frac{1}{2}, q = \frac{1}{2} \qquad |\text{ Since the question can either be true or false}$$

Probability that he will pass

$$= P(r \geq 7) = P(7) + P(8)$$

$$= {}^8C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^{8-7} + {}^8C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{8-8} = 8\left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^1 + 1.\left(\frac{1}{2}\right)^8$$

$$= \left(\frac{1}{2}\right)^8 (8+1) = \frac{9}{256} = .03516.$$

Example 6. *During war, 1 ship out of 9 was sunk on an average in making a certain voyage. What was the probability that exactly 3 out of a convoy of 6 ships would arrive safely?*

Sol. p, the probability of a ship arriving safely $= 1 - \dfrac{1}{9} = \dfrac{8}{9}$; $\quad q = \dfrac{1}{9}, n = 6$

The probability that exactly 3 ships arrive safely $= P(r = 3) = {}^6C_3 \left(\dfrac{1}{9}\right)^3 \left(\dfrac{8}{9}\right)^3 = \dfrac{10240}{9^6}$.

Example 7. *A policeman fires 6 bullets on a dacoit. The probability that the dacoit will be killed by a bullet is 0.6. What is the probability that dacoit is still alive?*

Sol. Here $\quad n$ = no. of bullets fired $= 6, p = 0.6, \quad q = 1 - p = 0.4$

Probability that dacoit is still alive (not killed)

$$= P(r = 0) = {}^nC_0\, p^0\, q^{n-0} = {}^6C_0\,(.6)^0\,(.4)^6 = (.4)^6 = .004096.$$

Example 8. *If the probability of hitting a target is 10% and 10 shots are fired independently. What is the probability that the target will be hit at least once?* (A.K.T.U. 2019)

Sol. Here, $p = \dfrac{10}{100} = \dfrac{1}{10}, q = 1 - p = 1 - \dfrac{1}{10} = \dfrac{9}{10}, n = 10$

Probability that the target will be hit at least once

$$= P(r \geq 1) = 1 - P(r = 0)$$

$$= 1 - [{}^nC_0\, p^0\, q^n] = 1 - \left[{}^{10}C_0 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^{10}\right] = 0.6513.$$

Example 9. *Out of 800 families with 4 children each, how many families would be expected to have (i) 2 boys and 2 girls (ii) at least one boy (iii) no girl (iv) atmost two girls? Assume equal probabilities for boys and girls.*

(U.P.T.U. 2014)

Sol. Since probabilities for boys and girls are equal,

$$p = \text{probability of having a boy} = \frac{1}{2};\ q = \text{probability of having a girl} = \frac{1}{2}$$

$$n = 4 \qquad N = 800$$

(*i*) The expected number of families having 2 boys and 2 girls

$$= 800\ {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = 800 \times 6 \times \frac{1}{16} = 300.$$

(*ii*) The expected number of families having *at least one boy*

$$= 800\left[{}^4C_1\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right) + {}^4C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 + {}^4C_3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 + {}^4C_4\left(\frac{1}{2}\right)^4\right]$$

$$= 800 \times \frac{1}{16}\,[4 + 6 + 4 + 1] = 750$$

(*iii*) The expected number of families having no girl *i.e.*, having 4 boys

$$= 800\ .\ {}^4C_4\left(\frac{1}{2}\right)^4 = 50.$$

(*iv*) The expected number of families having *atmost two girls i.e.*, having *at least 2 boy*

$$= 800\left[{}^4C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 + {}^4C_3\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3 + {}^4C_4\left(\frac{1}{2}\right)^4\right] = 800 \times \frac{1}{16}\,[6 + 4 + 1] = 550.$$

Example 10. *Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six?*

Sol. p = the chance of getting 5 or 6 with one die $= \frac{2}{6} = \frac{1}{3}$

$$q = 1 - \frac{1}{3} = \frac{2}{3},\ n = 6,\ N = 729$$

Since dice are in sets of 6 and there are 729 sets.

The expected number of times *at least three* dice showing five or six

$$= N\ .\ P(r \geq 3)$$
$$= 729\ [P(3) + P(4) + P(5) + P(6)]$$
$$= 729\left[{}^6C_3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3 + {}^6C_4\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^4 + {}^6C_5\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^5 + {}^6C_6\left(\frac{1}{3}\right)^6\right]$$
$$= \frac{729}{3^6}\,[160 + 60 + 12 + 1] = 233.$$

Example 11. *The probability of a man hitting a target is $\frac{1}{3}$. How many times must he fire so that the probability of his hitting the target at least once is more than 90%?*

Sol. $$p = \frac{1}{3}$$

The probability of not hitting the target in n trials is q^n.

Therefore, to find the smallest n for which the probability of hitting at least once is more than 90%, we have

$$1 - q^n > 0.9$$

$$\Rightarrow \quad 1 - \left(\frac{2}{3}\right)^n > 0.9$$

$$\Rightarrow \quad \left(\frac{2}{3}\right)^n < 0.1$$

The smallest n for which the above inequality holds true is 6 hence he must fire 6 times.

Example 12. *In a bombing action, there is 50% chance that any bomb will strike the target. Two direct hits are needed to destroy the target completely. How many bombs are required to be dropped to give a 99% chance or better of completely destroying the target?*

Sol. We have, $p = \frac{50}{100} = \frac{1}{2}$

Since the probability must be greater than 0.99, if n bombs are dropped, we have

$${}^nC_2 \left(\frac{1}{2}\right)^n + {}^nC_3 \left(\frac{1}{2}\right)^n + {}^nC_4 \left(\frac{1}{2}\right)^n + \cdots\cdots + {}^nC_n \left(\frac{1}{2}\right)^n \geq 0.99$$

$$\Rightarrow \quad \left(\frac{1}{2}\right)^n [{}^nC_2 + {}^nC_3 + {}^nC_4 + \cdots\cdots + {}^nC_n] \geq 0.99$$

$$\frac{2^n - n - 1}{2^n} \geq 0.99$$

$$\Rightarrow \quad 1 - \frac{1+n}{2^n} \geq 0.99$$

$$\Rightarrow \quad \frac{1+n}{2^n} \leq 0.01$$

$$\Rightarrow \quad 2^n \geq 100n + 100$$

By trial, $n = 11$ satisfies the inequality.

Hence 11 bombs are required to be dropped.

TEST YOUR KNOWLDGE

1. (*i*) Ten coins are tossed simultaneously. Find the probability of getting at least seven heads.
   (*ii*) A die is thrown five times. If getting an odd number is a success, find the probability of getting at least four successes. *(M.T.U. 2012)*
2. (*a*) The probability of any ship of a company being destroyed on a certain voyage is 0.02. The company owns 6 ships for the voyage. What is the probability of :
   (*i*) losing one ship (*ii*) losing atmost two ships (*iii*) losing none?
   (*b*) Assume that on the average one telephone number out of fifteen called between 2 P.M. and 3 P.M. on week-days is busy. What is the probability that if 6 randomly selected telephone numbers are called (*i*) not more than 3 (*ii*) at least 3 of them will be busy?

3. (i) The incidence of occupational disease in an industry is such that the workers have a 20% chance of suffering from it. What is the probability that out of six workers chosen at random, four or more will suffer from the disease?

(ii) The probability that a man aged 60 will live to be 70 is 0.65. What is the probability that out of 10 men, now 60, at least 7 will live to be 70?

4. (i) If the mean of a binomial distribution is 3 and the variance is $\frac{3}{2}$, find the probability of obtaining at least 4 successes.

(ii) In a binomial distribution, for $n = 5$ if $P(x = 1) = 0.4096$ and $P(x = 2) = 0.2048$, then find the value of p.

(iii) The sum and product of the mean and variance of a binomial distribution are $\frac{25}{3}$ and $\frac{50}{3}$ respectively. Find the distribution.

(iv) If the probability of a defective bolt is 0.1, find (a) The mean (b) The standard deviation for the distribution in a total of 400 bolts.

(v) If the moment generating function of a random variable X is $\left(\frac{1}{3}+\frac{2}{3}e^t\right)^5$, find $P(X = 2)$.

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

(b) Four persons in a group of 20 are graduates. If 4 persons are selected at random from 20, find the probability that

(i) all are graduates (ii) at least one is a graduate.

6. A bag contains 5 white, 7 red and 8 black balls. If four balls are drawn, one by one, with replacement, what is the probability that

(i) none is white (ii) all are white

(iii) at least one is white (iv) only 2 are white?

7. (i) In a hurdle race, a player has to cross 10 hurdles. The probability that he will clear each hurdle is $\frac{5}{6}$. What is the probability that he will knock down fewer than 2 hurdles ?

(ii) If on an average one ship in every ten is wrecked, find the probability that out of 5 ships expected to arrive, 4 at least will arrive safely.

8. The prob. that a bulb produced by a factory will fuse after a use of 150 days is 0.05. Find the prob. that out of 5 such bulbs [M.T.U. (B. Pharma) 2011]

(i) None (ii) Atmost one

(iii) More than one (iv) At least one will fuse after 150 days of use.

9. Five cards are drawn successively with replacement from a well-shuffled deck of 52 cards. What is the probability that

(i) all the five cards are spades (ii) only three are spades? (iii) none is spade?

10. Manish takes a step forward with probability 0.4 and backward with probability 0.6. Find the probability that at the end of 11 steps, he is one step away from the starting point.

[**Hint.** Manish will take either 6 steps forward and 5 backward or 5 steps forward and 6 backward].

11. (a) In 800 families with 5 children each, how many families would be expected to have (i) 3 boys and 2 girls, (ii) 2 boys and 3 girls, (iii) no girl (iv) at the most two girls. (Assume probabilities for boys and girls to be equal.) (A.K.T.U. 2017)

(b) Out of 800 families with 5 children each, how many would you expect to have (i) 3 boys (ii) 5 girls (iii) either 2 or 3 boys? Assume equal probabilities for boys and girls.

(M.T.U. 2012)

(c) Out of 320 families with 5 children each, what percentage would be expected to have (i) 2 boys and 3 girls (ii) at least one boy? Assume equal probability for boys and girls.

(G.B.T.U. 2011)

12. In 100 sets of ten tosses of an unbiased coin, in how many cases do you expect to get
(i) 7 heads and 3 tails (ii) at least 7 heads?

13. The following data are the number of seeds germinating out of 10 on damp filter for 80 sets of seeds. Fit a binomial distribution to this data:

x:	0	1	2	3	4	5	6	7	8	9	10	Total
f:	6	20	28	12	8	6	0	0	0	0	0	80

[**Hint.** Here $n = 10$, $N = 80$, Mean $= \frac{\Sigma fx}{\Sigma f} = 2.175$ $\therefore$ $np = 2.175$ etc.]

14. Fit a binomial distribution for the following data and compare the theoretical frequencies with the actual ones :

x:	0	1	2	3	4	5
f:	2	14	20	34	22	8

15. Fit a binomial distribution to the data given in the following table:

(i) x:	0	1	2	3	4
f:	24	41	28	5	2

(M.T.U. 2012)

(ii) x:	0	1	3	4
f:	28	62	10	4

(U.K.T.U. 2011)

16. (i) Assuming half the population of a town consumes chocolates so that the chance of an individual being consumer is $\frac{1}{2}$ and that 100 investigators each take 10 individuals to see whether they are consumers, how many investigators would you expect to report that 3 people or less were consumers?

(ii) Assuming that 20% of the population of a city are literate, so that the chance of an individual being literate is $\frac{1}{5}$ and assuming that 100 investigators each take 10 individuals to see whether they are literate, how many investigators would you expect to report 3 or less were literate?

17. Following results were obtained when 100 batches of seeds were allowed to germinate on damp filter paper in a laboratory : $\beta_1 = \frac{1}{15}$, $\beta_2 = \frac{89}{30}$. Determine the Binomial distribution. Calculate the expected frequency for $x = 8$ assuming $p > q$.

18. A coffee connoisseur claims that he can distinguish between a cup of instant coffee and a cup of percolator coffee 75% of the time. It is agreed that his claim will be accepted if he correctly identifies at least 5 of the 6 cups. Find his chances of having the claim (i) accepted (ii) rejected when he does have the ability he claims.

19. A multiple-choice test consists of 8 questions with 3 answers to each question of which only one is correct. A student answers each question by rolling a balanced die and checking the first answer if he gets 1 or 2, the second answer if he gets 3 or 4 and the third answer if he gets 5 or 6. To get a

distinction, the student must secure at least 75% correct answers. If there is no negative marking, what is the probability that the student secures a distinction?

20. An irregular six-faced die is thrown and the expectation that in 10 throws, it will give five even numbers is twice the expectation that it will give four even numbers. How many times in 10,000 sets of 10 throws each, would you expect it to give no even number?

Answers

1. (i) $\frac{11}{64}$ (ii) $\frac{3}{16}$

2. (a) (i) 0.1085, (ii) 0.9997, (iii) 0.8858 (b) (i) 0.9997 (ii) 0.005

3. (i) $\frac{53}{3125}$ (ii) 0.5137

4. (i) $\frac{11}{32}$, (ii) $\frac{1}{5}$, (iii) $\left(\frac{2}{3}+\frac{1}{3}\right)^{15}$

(iv) (a) 40 (b) 6 (v) 0.1646

5. (a) (i) 0.246 (ii) 0.345 (b) (i) 0.0016 (ii) 0.5904

6. (i) $\frac{81}{256}$, (ii) $\frac{1}{256}$, (iii) $\frac{175}{256}$, (iv) $\frac{27}{128}$ **7.** (i) $\frac{5}{2}\left(\frac{5}{6}\right)^9$, (ii) 0.91854

8. (i) $\left(\frac{19}{20}\right)^5$ (ii) $\frac{6}{5}\left(\frac{19}{20}\right)^4$ (iii) $1-\frac{6}{5}\left(\frac{19}{20}\right)^4$ (iv) $1-\left(\frac{19}{20}\right)^5$

9. (i) $\left(\frac{1}{4}\right)^5$ (ii) $90\left(\frac{1}{4}\right)^5$ (iii) $\left(\frac{3}{4}\right)^5$ **10.** 0.36787

11. (a) (i) 250 (ii) 250 (iii) 25 (iv) 400

(b) (i) 250 (ii) 25 (iii) 500

(c) (i) 31.25% (ii) 96.875%

12. (i) 12 nearly (ii) 17 nearly **13.** $80\,(0.7825 + 0.2175)^{10}$

14. $100\,(0.432 + 0.568)^5$ **15.** (i) $100\,(0.7 + 0.3)^4$, (ii) $104\,(0.7404 + 0.2596)^4$

16. (i) 17 (ii) 88 **17.** $100\left(\frac{1}{4}+\frac{3}{4}\right)^{20}$, 0.075168752

18. (i) 0.534 (ii) 0.466

19. 0.0197 **20.** 1.

POISSON DISTRIBUTION

4.42 POISSON DISTRIBUTION AS A LIMITING CASE OF BINOMIAL DISTRIBUTION

[G.B.T.U. 2013 ; M.T.U. 2013, 2014]

Poisson distribution was discovered by S.D. Poisson in the year 1837.

If the parameters n and p of a binomial distribution are known, we can find the distribution. But in situations where n is very large and p is very small, application of binomial distribution is very labourious. However, if we assume that as $n \to \infty$ and $p \to 0$ such that np always remains finite, say λ, we get the Poisson approximation to the binomial distribution.

Now, for a binomial distribution

$$P(X = r) = {}^nC_r q^{n-r} p^r$$

$$= \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!} \times (1-p)^{n-r} \times p^r$$

$$= \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!} \times \left(1-\frac{\lambda}{n}\right)^{n-r} \times \left(\frac{\lambda}{n}\right)^r \qquad \Big| \text{ Since } np = \lambda \quad \therefore \quad p = \frac{\lambda}{n}$$

$$= \frac{\lambda^r}{r!} \times \frac{n(n-1)(n-2)\ldots(n-r+1)}{n^r} \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\ldots\ldots\left(\frac{n-r+1}{n}\right) \times \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$= \frac{\lambda^r}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\ldots\left(1-\frac{r-1}{n}\right) \times \frac{\left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}}{\left(1-\frac{\lambda}{n}\right)^r}$$

As $n \to \infty$, each of the $(r-1)$ factors

$$\left(1-\frac{1}{n}\right), \left(1-\frac{2}{n}\right), \ldots\ldots, \left(1-\frac{r-1}{n}\right) \quad \text{tends to 1. Also } \left(1-\frac{\lambda}{n}\right)^r \text{ tends to 1.}$$

Since $\underset{x \to \pm\infty}{\text{Lt}} \left(1+\frac{1}{x}\right)^x = e$, the Naperian base. $\quad \therefore \quad \left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda} \longrightarrow e^{-\lambda}$ as $n \to \infty$

Hence in the limiting case when $n \to \infty$, we have

$$\boxed{P(X = r) = \frac{e^{-\lambda}.\lambda^r}{r!} \quad (r = 0, 1, 2, 3, \ldots)} \qquad \ldots(1)$$

where λ is a finite number $= np$.

(1) represents a probability distribution which is called the *Poisson probability distribution.*

Note 1. λ is called the parameter of the distribution.

Note 2. The sum of the probabilities $P(r)$ for $r = 0, 1, 2, 3, \ldots\ldots$ is 1, since

$$P(0) + P(1) + P(2) + P(3) + \ldots\ldots = e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!} + \frac{\lambda^3 e^{-\lambda}}{3!} + \ldots\ldots$$

$$= e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots\ldots\right) = e^{-\lambda}.e^{\lambda} = 1.$$

4.43 RECURRENCE FORMULA FOR THE POISSON DISTRIBUTION

For Poisson distribution, $\quad P(r) = \dfrac{e^{-\lambda}\,\lambda^r}{r\,!}$ and $P(r+1) = \dfrac{e^{-\lambda}\,\lambda^{r+1}}{(r+1)\,!}$

$$\therefore \quad \frac{P(r+1)}{P(r)} = \frac{\lambda r\,!}{(r+1)\,!} = \frac{\lambda}{r+1}$$

or $$P(r+1) = \frac{\lambda}{r+1} P(r),\ r = 0, 1, 2, 3, \ldots\ldots$$

This is called the *recurrence* or *recursion formula* for the Poisson distribution.

4.44 MEAN AND VARIANCE OF THE POISSON DISTRIBUTION (*A.K.T.U. 2018*)

For the Poisson distribution, $\quad P(r) = \dfrac{e^{-\lambda}\,\lambda^r}{r\,!}$

$$\text{Mean } \mu = \sum_{r=0}^{\infty} rP(r) = \sum_{r=0}^{\infty} r\,.\frac{e^{-\lambda}\,\lambda^r}{r\,!}$$

$$= e^{-\lambda} \sum_{r=1}^{\infty} \frac{\lambda^r}{(r-1)\,!} = e^{-\lambda}\left(\lambda + \frac{\lambda^2}{1\,!} + \frac{\lambda^3}{2\,!} + \ldots\ldots\right)$$

$$= \lambda e^{-\lambda}\left(1 + \frac{\lambda}{1\,!} + \frac{\lambda^2}{2\,!} + \ldots\ldots\right) = \lambda e^{-\lambda}\,.\,e^{\lambda} = \lambda$$

Thus, the mean of the Poisson distribution is equal to the parameter λ.

$$\text{Variance } \sigma^2 = \sum_{r=0}^{\infty} r^2 P(r) - \mu^2 = \sum_{r=0}^{\infty} r^2\,.\frac{\lambda^r e^{-\lambda}}{r\,!} - \lambda^2 = e^{-\lambda}\sum_{r=1}^{\infty}\frac{r^2\lambda^r}{r\,!} - \lambda^2$$

$$= e^{-\lambda}\left[\frac{1^2\,.\,\lambda}{1\,!} + \frac{2^2\,.\,\lambda^2}{2\,!} + \frac{3^2\,.\,\lambda^3}{3\,!} + \frac{4^2\,.\,\lambda^4}{4\,!} + \ldots\ldots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[1 + \frac{2\lambda^2}{1\,!} + \frac{3\lambda^2}{2\,!} + \frac{4\lambda^3}{3\,!} + \ldots\ldots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[1 + \frac{(1+1)\lambda}{1\,!} + \frac{(1+2)\lambda^2}{2\,!} + \frac{(1+3)\lambda^3}{3\,!} + \ldots\ldots\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[\left(1 + \frac{\lambda}{1\,!} + \frac{\lambda^2}{2\,!} + \frac{\lambda^3}{3\,!} + \ldots\ldots\right) + \left(\frac{\lambda}{1\,!} + \frac{2\lambda^2}{2\,!} + \frac{3\lambda^3}{3\,!} + \ldots\ldots\right)\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\left[e^{\lambda} + \lambda\left(1 + \frac{\lambda}{1\,!} + \frac{\lambda^2}{2\,!} + \ldots\ldots\right)\right] - \lambda^2$$

$$= \lambda e^{-\lambda}\,[e^{\lambda} + \lambda e^{\lambda}] - \lambda^2 = \lambda e^{-\lambda}\,.\,e^{\lambda}(1+\lambda) - \lambda^2 = \lambda(1+\lambda) - \lambda^2 = \lambda.$$

Hence, the variance of the Poisson distribution is also λ.

Thus, the mean and the variance of the Poisson distribution are each equal to the parameter λ.

Note 1. The mean and the variance of the Poisson distribution can also be derived from those of the binomial distribution in the limiting case when $n \to \infty$, $p \to 0$ and $np = \lambda$.

Mean of Binomial distribution is np.

$\therefore$ Mean of Poisson distribution $= \underset{n \to \infty}{\text{Lt}}\ np = \underset{n \to \infty}{\text{Lt}}\ \lambda = \lambda$

Variance of Binomial distribution is $npq = np\,(1-p)$

$\therefore$ Variance of Poisson distribution $= \underset{n \to \infty}{\text{Lt}}\ np\,(1-p) = \underset{n \to \infty}{\text{Lt}}\ \lambda\left(1 - \dfrac{\lambda}{n}\right) = \lambda.$

Note 2. For Poisson distribution, $\mu_3 = \lambda$ and $\mu_4 = 3\lambda^2 + \lambda$.

Coefficients of skewness and kurtosis are given by

$$\beta_1 = \frac{1}{\lambda} \text{ and } \gamma_1 = \frac{1}{\sqrt{\lambda}}\text{. Also, } \beta_2 = 3 + \frac{1}{\lambda} \text{ and } \gamma_2 = \frac{1}{\lambda}$$

Hence Poisson distribution is always a skewed distribution.

Remark. While fitting the Poisson distribution to a given data, we round the figures to the nearest integer but it should be kept in mind that the total of the observed and the expected frequencies should be same.

4.45 MODE OF POISSON DISTRIBUTION

Let $$P(x = r) = e^{-\lambda}\frac{\lambda^r}{r!},\ r = 0, 1, 2, \ldots\ldots, \infty$$

The value of r which has a greater probability than any other value is the mode of the Poisson distribution. Thus r is mode if

$$P(X = r) \geq P(X = r + 1) \text{ and } P(X = r) \geq P(X = r - 1)$$

$$\Rightarrow \quad \frac{e^{-\lambda}.\lambda^r}{r!} \geq \frac{e^{-\lambda}.\lambda^{r+1}}{(r+1)!} \quad \text{and} \quad \frac{e^{-\lambda}.\lambda^r}{r!} \geq \frac{e^{-\lambda}.\lambda^{r-1}}{(r-1)!}$$

$$\Rightarrow \quad 1 \geq \frac{\lambda}{r+1} \quad \text{and} \quad \frac{\lambda}{r} \geq 1$$

$$\Rightarrow \quad r \geq \lambda - 1 \quad \text{and} \quad r \leq \lambda \quad i.e., \quad \lambda - 1 \leq r \leq \lambda$$

Case I. If λ is a positive integer, there are two modes $\lambda - 1$ and λ.

Case II. If λ is not a positive integer, there is one mode and is the integral value between $\lambda - 1$ and λ.

4.46 APPLICATIONS OF POISSON DISTRIBUTION

This distribution is applied to problems concerning :

(*i*) Arrival pattern of defective vehicles in a workshop.

(*ii*) Patients in a hospitals.

(*iii*) Telephone calls.

(*iv*) Demand pattern for certain spare parts.

(*v*) Number of fragments from a shell hitting a target.

(*vi*) Emission of radioactive (α) particles.

ILLUSTRATIVE EXAMPLES

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

(M.T.U. 2013)

Sol. λ, the parameter of Poisson distribution = Variance = 2

Recurrence relation for the Poisson distribution is

$$P(r+1) = \frac{\lambda}{r+1} P(r) = \frac{2}{r+1} P(r) \qquad \ldots(1)$$

Now $\quad P(r) = \dfrac{e^{-\lambda} \lambda^r}{r!} \quad \Rightarrow \quad P(0) = \dfrac{e^{-2}(2)^0}{0!} = e^{-2} = 0.1353$

Putting $r = 0, 1, 2, 3$ in (1), we get

$$P(1) = 2P(0) = 2 \times 0.1353 = 0.2706\,; \qquad P(2) = \frac{2}{2} P(1) = 0.2706$$

$$P(3) = \frac{2}{3} P(2) = \frac{2}{3} \times 0.2706 = 0.1804\,; \quad P(4) = \frac{2}{4} P(3) = \frac{1}{2} \times 0.1804 = 0.0902.$$

Now, $P(r \geq 4) = 1 - [P(0) + P(1) + P(2) + P(3)]$

$= 1 - [0.1353 + 0.2706 + 0.2706 + 0.1804] = 0.1431.$

Example 2. *Using Poisson distribution, find the probability that the ace of spades will be drawn from a pack of well-shuffled cards at least once in 104 consecutive trials.* (U.P.T.U. 2015)

Sol. $\qquad p = \dfrac{1}{52},\ n = 104$

$\therefore \qquad \lambda = np = 104 \times \dfrac{1}{52} = 2$

Prob. (at least once) $= P(r \geq 1) = 1 - P(0)$

$$= 1 - \frac{e^{-\lambda}.\lambda^0}{0!} = 1 - e^{-2} = 1 - 0.135335 \approx 0.8647.$$

Example 3. (*i*) *Fit a Poisson distribution to the following data and calculate theoretical frequencies.*

Deaths:	*0*	*1*	*2*	*3*	*4*
Frequencies:	*122*	*60*	*15*	*2*	*1*

[U.P.T.U. 2014 ; U.K.T.U. 2010]

(*ii*) *The frequency of accidents per shift in a factory is shown in the following table:*

Accident per shift	*Frequency*
0	*192*
1	*100*
2	*24*
3	*3*
4	*1*
Total	320

Calculate the mean number of accidents per shift. Fit a Poisson distribution and calculate theoretical frequencies.

Sol. (*i*) Mean of given distribution $= \dfrac{\Sigma fx}{\Sigma f}$

$$\Rightarrow \qquad \lambda = \frac{60 + 30 + 6 + 4}{200} = 0.5$$

Required Poisson distribution $= \text{N} \,.\, \dfrac{e^{-\lambda}\,.\,\lambda^r}{r\,!} = 200\,.\,\dfrac{e^{-.5}\,(.5)^r}{r\,!} = (121.306)\dfrac{(0.5)^r}{r\,!}$

r	*N. P(r)*	*Theoretical frequency*
0	$121.306 \times \dfrac{(0.5)^0}{0\,!} = 121.306$	121
1	$121.306 \times \dfrac{(0.5)^1}{1\,!} = 60.653$	61
2	$121.306 \times \dfrac{(0.5)^2}{2\,!} = 15.163$	15
3	$121.306 \times \dfrac{(0.5)^3}{3\,!} = 2.527$	3
4	$121.306 \times \dfrac{(0.5)^4}{4\,!} = 0.3159$	0
		Total = 200

(*ii*) Mean number of accidents per shift

$$\lambda = \frac{\Sigma fx}{\Sigma f} = \frac{100 + 48 + 9 + 4}{320} = 0.5031$$

$\therefore$ Required Poisson distribution

$$= \text{N}\,.\,\frac{e^{-\lambda}\,.\,\lambda^r}{r\,!} = 320\,.\,\frac{e^{-.5031}\,(.5031)^r}{r\,!} = \frac{(193.48)(.5031)^r}{r\,!}$$

r	*N. P(r)*	*Theoretical frequency*
0	193.48	194
1	97.34	97
2	24.38	24
3	4.10	4
4	0.51	1
		Total = 320

Example 4. (*i*) *Suppose that a book of 600 pages contains 40 printing mistakes. Assume that these errors are randomly distributed throughout the book and r, the number of errors per page has a Poisson distribution. What is the probability that 10 pages selected at random will be free from errors?*

(*ii*) *Wireless sets are manufactured with 25 solders joints each, on the average 1 joint in 500 is defective. How many sets can be expected to be free from defective joints in a consignment of 10000 sets?*

Sol. (*i*)
$$p = \frac{40}{600} = \frac{1}{15}, \quad n = 10$$

$$\therefore \quad \lambda = np = 10\left(\frac{1}{15}\right) = \frac{2}{3}$$

$$P(r) = \frac{e^{-\lambda}\lambda^r}{r!} = \frac{e^{-2/3}(2/3)^r}{r!}$$

$$\therefore \quad P(0) = \frac{e^{-2/3}(2/3)^0}{0!} = e^{-2/3} = 0.51.$$

(*ii*)
$$p = \frac{1}{500}, \; n = 25$$

$$\therefore \quad \lambda = np = 25 \times \frac{1}{500} = \frac{1}{20} = 0.05$$

No. of sets in a consignment, N= 10000

Probability of having no defective joint = $P(r = 0) = \dfrac{e^{-.05}(0.05)^0}{0!} = 0.9512.$

$\therefore$ The expected no. of sets free from defective joints = 0.9512 × 10000 = 9512.

Example 5. *A manufacturer knows that the condensors he makes contain on an average 1% of defectives. He packes them in boxes of 100. What is the probability that a box picked at random will contain 4 or more faulty condensors?*

Sol.
$$p = 0.01, \; n = 100$$

$$\therefore \quad \lambda = np = 1$$

$$P(r) = \frac{e^{-\lambda}\lambda^r}{r!} = \frac{e^{-1}}{r!}$$

P(4 or more faulty condensors) = P(4) + P(5) + ... + P(100)

$$= 1 - [P(0) + P(1) + P(2) + P(3)]$$

$$= 1 - \left[\frac{e^{-1}}{0!} + \frac{e^{-1}}{1!} + \frac{e^{-1}}{2!} + \frac{e^{-1}}{3!}\right]$$

$$= 1 - e^{-1}\left[1 + 1 + \frac{1}{2} + \frac{1}{6}\right] = 1 - \frac{8}{3e} = 1 - 0.981 = 0.019.$$

Example 6. (*i*) *If the probabilities of a bad reaction from a certain injection is 0.0002, determine the chance that out of 1000 individuals more than two will get a bad reaction.*

(*ii*) *The probability that a man aged 50 years will die within a year is 0.01125. What is the probability that of 12 such men, at least 11 will reach their 51st birthday?*

(*Given*: $e^{-.135} = 0.87366$)

Sol. (*i*) Here, $p = 0.0002, \quad n = 1000$

$\therefore \quad \lambda = np = 1000 \times 0.0002 = 0.2.$

Since the prob. of bad reaction is very small and no. of trials is very high, we use Poisson distribution here.

The prob. that out of 100 individuals, more than 2 will get a bad reaction is

$$= P(r > 2) = 1 - P(r \leq 2) = 1 - [P(0) + P(1) + P(2)] \quad ...(1)$$

Now, $$P(0) = \frac{e^{-0.2}(0.2)^0}{0} = 0.8187 \quad \text{(Here } r = 0\text{)}$$

$$P(1) = \frac{e^{-0.2}(0.2)^1}{1} = 0.1637 \quad \text{(Here } r = 1\text{)}$$

and $$P(2) = \frac{e^{-0.2}(0.2)^2}{2} = 0.0164. \quad \text{(Here } r = 2\text{)}$$

$\therefore$ From (1), Reqd. probability $= 1 - [0.8187 + 0.1637 + 0.0164] = 0.0012.$

(*ii*) $p = 0.01125, n = 12$

$\therefore \quad \lambda = np = 12 \times 0.01125 = 0.135$

P (at least 11 survive) = P(atmost 1 dies)

$$= P(0) + P(1) = \frac{e^{-\lambda}.\lambda^0}{0!} + \frac{e^{-\lambda}.\lambda^1}{1!}$$

$$= e^{-.135}(1 + 0.135) = 1.135 \times 0.87366 = 0.9916.$$

Example 7. *A car-hire firm has two cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with mean 1.5. Calculate the proportion of days on which neither car is used and the proportion of days on which some demand is refused ($e^{-1.5} = 0.2231$).*

Sol. Since the number of demands for a car is distributed as a Poisson distribution with mean $\lambda = 1.5$.

$\therefore$ Proportion of days on which neither car is used

= Probability of there being no demand for the car

$$= \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-1.5} = 0.2231$$

Proportion of days on which some demand is refused

= probability for the number of demands to be more than two

$$= 1 - P(x \leq 2) = 1 - \left(e^{-\lambda} + \frac{\lambda e^{-\lambda}}{1!} + \frac{\lambda^2 e^{-\lambda}}{2!}\right)$$

$$= 1 - e^{-1.5}\left(1 + 1.5 + \frac{(1.5)^2}{2}\right) = 0.1912625.$$

Example 8. *Suppose the number of telephone calls on an operator received from 9 : 00 to 9 : 05 follow a Poisson distribution with a mean 3. Find the probability that*

(i) The operator will receive no calls in that time interval tomorrow.

(ii) In the next three days, the operator will receive a total of 1 call in that time interval.

(Given: $e^{-3} = 0.04978$)

Sol. Here, $\lambda = 3$

(i) $$P(0) = \frac{e^{-\lambda} . \lambda^0}{0!} = e^{-3} = 0.04978$$

(ii) Required probability = P(0) P(0) P(1) + P(0) P(1) P(0) + P(1) P(0) P(0)

$$= 3\left\{\frac{e^{-\lambda} . \lambda^0}{0!}\right\}^2 \frac{e^{-\lambda} . \lambda^1}{1!} = 9(e^{-3})^3 = 0.00111.$$

Example 9. *The no. of arrivals of customers during any day follows Poisson distribution with a mean of 5. What is the probability that the total no. of customers on two days selected at random is less than 2? (Given: $e^{-10} = 4.54 \times 10^{-5}$)*

Sol. $\lambda = 5$

Arrival of Customers

I day	*II day*	*Total*
0	0	0
0	1	1
1	0	1

Required probability $= P(0)\,P(0) + P(0)\,P(1) + P(1)\,P(0)$

$$= \frac{e^{-5}\, 5^0}{0!} . \frac{e^{-5}\, 5^0}{0!} + \frac{e^{-5}\, 5^0}{0!} . \frac{e^{-5}\, 5^1}{1!} + \frac{e^{-5}\, 5^1}{1!} . \frac{e^{-5}\, 5^0}{0!}$$

$$= e^{-10} + 2 . 5 . e^{-10} = 11\, e^{-10} = 11 \times 4.54 \times 10^{-5}$$

$$= 4.994 \times 10^{-4}.$$

Example 10. *An insurance company finds that 0.005% of the population dies from a certain kind of accident each year. What is the probability that the company must pay off no more than 3 of 10,000 insured risks against such incident in a given year?*

Sol. $$p = \frac{0.005}{100} = 0.00005, \quad n = 10000$$

$\therefore$ $$\lambda = np = 10000 \times 0.00005 = 0.5$$

Required Probability $= 1 - P(r \le 3) = 1 - [P(0) + P(1) + P(2) + P(3)]$

$$= 1 - \left[\frac{e^{-0.5}(0.5)^0}{0!} + \frac{e^{-0.5}(0.5)^1}{1!} + \frac{e^{-0.5}(0.5)^2}{2!} + \frac{e^{-0.5}(0.5)^3}{3!}\right]$$

$$= 1 - e^{-5}\,[1 + 0.5 + 0.125 + 0.021] = 0.0016.$$

Example 11. *In a certain factory turning out razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Calculate the approximate number of packets containing no defective, one defective and two defective blades in a consignment of 10,000 packets. (Given: $e^{-0.02} = 0.9802$)*

Sol. p(defective) = 0.002

$\therefore$ $n = 10$ (no. of blades in a packet)

$$\lambda = np = 10 \times 0.002 = 0.02$$

No. of packets in the consignment, N = 10,000.

(*i*) Probability of having no defective = $P(0) = \frac{e^{-0.02}(0.02)^0}{0!} = 0.9802$ | Here $r = 0$

Approximate no. of packets having zero defective in the consignment = 0.9802×10000
$= 9802$

(*ii*) Probability of having one defective = $P(1) = \frac{e^{-0.02}(0.02)^1}{1!} = 0.9802 \times 0.02 = 0.019604$

Approximate no. of packets having one defective in the consignment
$= 0.019604 \times 10000 \approx 196.$

(*iii*) Probability of having two defective blades

$$P(2) = \frac{e^{-0.02}(0.02)^2}{2!} = \frac{(0.980198) \times (0.0004)}{2} = 0.000196.$$

$\therefore$ Approximate no. of packet having two defectives in the consignment
$= 0.000196 \times 10000 = 1.96 \approx 2.$

Example 12. (*i*) *Six coins are tossed 6400 times. Using the Poisson distribution, determine the approximate probability of getting six heads x times.*

(*ii*) *A Poisson distribution has a double mode at x = 3 and x = 4. What is the probability that x will have one or the other of these two values?*

Sol. (*i*) Probability of getting one head with one coin $= \frac{1}{2}$.

$\therefore$ The probability of getting six heads with six coins $= \left(\frac{1}{2}\right)^6 = \frac{1}{64}$

$\therefore$ Average number of six heads with six coins in 6400 throws $= np = 6400 \times \frac{1}{64} = 100$

$\therefore$ The mean of the Poisson distribution = 100.

Approximate probability of getting six heads x times when the distribution is Poisson

$$= \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(100)^x \cdot e^{-100}}{x!}.$$

(*ii*) Since 2 modes are given when λ is an integer, modes are $\lambda - 1$ and λ.

$\therefore \qquad \lambda - 1 = 3 \Rightarrow \lambda = 4$

Probability (when $r = 3$) $= \frac{e^{-4}(4)^3}{3!}$

Probability (when $r = 4$) $= \frac{e^{-4}(4)^4}{4!}$

$\therefore$ Required Probability $= P(r = 3 \text{ or } 4) = P(r = 3) + P(r = 4)$

$$= \frac{e^{-4}(4)^3}{3!} + \frac{e^{-4}(4)^4}{4!} = \frac{64}{3} e^{-4} = 0.39073.$$

Example 13. *For a Poisson distribution with mean m, show that*

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

Sol. $$\mu_r = \sum_{x=0}^{\infty}(x-m)^r\,.\frac{e^{-m}.m^x}{x!}$$

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

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

$$\Rightarrow \quad \mu_{r+1} = m\frac{d\mu_r}{dm} + mr\,\mu_{r-1}.$$

Example 14. *Show that in a Poisson distribution with unit mean, mean deviation about mean is* $\left(\frac{2}{e}\right)$ *times the standard deviation.* (G.B.T.U. 2012)

Sol. Here, $\lambda = 1$

$$\therefore \quad P(X = x) = \frac{e^{-1}.(1)^x}{x!} = \frac{e^{-1}}{x!}\ ;\ x = 0, 1, 2, \ldots$$

Mean deviation about mean 1 is

$$= \sum_{x=0}^{\infty}|x-1|\,p(x) = e^{-1}\sum_{x=0}^{\infty}\frac{|x-1|}{x!} = e^{-1}\left[1 + \frac{1}{2!} + \frac{2}{3!} + \frac{3}{4!} + \ldots\right] \qquad \ldots(1)$$

we have, $$\frac{n}{(n+1)!} = \frac{\overline{n+1}-1}{(n+1)!} = \frac{1}{n!} - \frac{1}{(n+1)!}$$

$\therefore$ From (1), Mean deviation about mean $= e^{-1}\left[1 + \left(1 - \frac{1}{2!}\right) + \left(\frac{1}{2!} - \frac{1}{3!}\right) + \ldots\right]$

$$= e^{-1}(1+1) = \frac{2}{3}\times 1 = \frac{2}{e}\times \text{S.D.} \qquad |\text{ Since, variance = mean = 1}$$

TEST YOUR KNOWLEDGE

1. If X is a Poisson variate such that P(X = 2) = 9P(X = 4) + 90P(X = 6), find the standard deviation.
2. If a random variable has a Poisson distribution such that P(1) = P(2), find
   (*i*) mean of the distribution (*ii*) P(4).
3. Suppose that X has a Poisson distribution. If $P(X = 2) = \frac{2}{3}P(X = 1)$ find, (*i*) P(X = 0) (*ii*) P(X = 3).

4. A certain screw making machine produces on average 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws.

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

(*ii*) The experience shows that 4 industrial accidents occur in a plant on an average per month. Calculate the probabilities of less than 3 accidents in a certain month. Use Poisson distribution. (Given : $e^{-4} = 0.01832$). [*M.T.U.* (*MBA*) *2011*]

6. (*i*) Suppose a book of 585 pages contains 43 typographical errors. If these errors are randomly distributed throughout the book, what is the probability that 10 pages, selected at random, will be free from errors ?

(*ii*) Assume that the probability of an individual coalminer being killed in a mine accident during a year is $\frac{1}{2400}$. Use Poisson's distribution to calculate the probability that in a mine employing 200 miners there will be at least one fatal accident in a year.

7. (*i*) A manufacturer of cotter pins knows that 5% of his product is defective. If he sells cotter pins in boxes of 100 and guarantee that not more than 10 pins will be defective, what is the approximate prob. that a box will fail to meet the guaranteed quality?

(*ii*) An insurance company insures 4000 people against loss of both eyes in a car accident. Based on previous data it was assumed 10 persons out of 1,00,000 will have such type of injury in car accident. What is probability that more than 2 of the insured will collect on their policy in a given year? (*M.T.U. 2013*)

8. Records show that the probability is 0.00002 that a car will have a flat tyre while driving over a certain bridge. Use Poisson distribution to find the probability that among 20,000 cars driven over the bridge, not more than one will have a flat tyre.

9. Between the hours of 2 and 4 P.M., the average no. of phone calls per minute coming into the switch board of a company is 2.5. Find the probability that during a particular minute, there will be no phone call at all. [Given : $e^{-2} = 0.13534$ and $e^{-0.5} = 0.60650$.]

10. (*i*) Fit a Poisson distribution to the following data given the number of yeast cells per square for 400 squares :

No. of cells per sq.	:	0	1	2	3	4	5	6	7	8	9	10
No. of squares	:	103	143	98	42	8	4	2	0	0	0	0

It is given that $e^{-1.3225} = 0.2665$.

(*ii*) Data was collected over a period of 10 years, showing number of deaths from horse kicks in each of the 200 army corps. The distribution of deaths was as follows:

No. of deaths	:	0	1	2	3	4	Total
Frequency	:	109	65	22	3	1	200

Fit a Poisson distribution to the data and calculate the theoretical frequencies. [*M.T.U.* (*MBA*) *2011*]

(*iii*) The following table gives the no. of days in a 50 day period during which automobile accidents occured in a city.

No. of accidents	:	0	1	2	3	4
No. of days	:	21	18	7	3	1

Fit a Poisson distribution to the data. (*G.B.T.U. 2011*)

11. The number of accidents in a year involving taxi drivers in a city follows a Poisson distribution with mean equal to 3. Out of 1000 taxi drivers, find approximately the number of drivers with

(*i*) no accidents (*ii*) more than 3 accidents in a year.

12. 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) at least 2 accidents

(iii) atmost 3 accidents (iv) between 2 and 5 accidents.

13. It is given that 2% of the electric bulbs manufactured by a company are defective. Using Poisson distribution, find the probability that a sample of 200 bulbs will contain (i) no defective bulb (ii) two defective bulbs (iii) at the most three defective bulbs. (G.B.T.U. 2011)

14. A manager accepts the work submitted by his typist only when there is no mistake in the work. The typist has to type on an average 20 letters per day of about 200 words each. Find the chance of her making a mistake if

(i) less than 1% of the letters submitted by her are rejected

(ii) on 90% days all the letters submitted by her are accepted.

As the probability of making a mistake is small, you may use Poisson distribution. (Take $e = 2.72$)

[**Hint:** (i) $e^{-200p} \geq 0.99$ (ii) $e^{-4000p} = 0.90$]

15. In a certain factory manufacturing razor blades, there is a small chance of 0.002 for any blade to be defective. The blades are supplied in packets of 10. Use suitable distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 50,000 packets. (A.K.T.U. 2018)

Answers

1. 1 **2.** (i) 2 (ii) $\frac{2}{3e^2}$ **3.** (i) $e^{-(4/3)}$ (ii) $\frac{32}{81}e^{-(4/3)}$ **4.** 0.035

5. (i) 0.0008 (ii) 0.23816

6. (i) 0.4795 (ii) 0.08 **7.** (i) 0.0136875 (ii) 0.007926

8. 0.938 **9.** 0.08208

10. (i) Theoretical frequencies are 107, 141, 93, 41, 13, 4, 1, 0, 0, 0, 0.

(ii) $108.67 \times \frac{(0.61)^r}{r!}$; 109, 66, 20, 4 and 1 (iii) $(50)\,\frac{e^{-0.9}\,(0.9)^r}{r!}$; 20, 18, 8, 3 and 1.

11. (i) 50 (ii) 353

12. (i) 2 (ii) 91 (iii) 43 (iv) 39

13. (i) 0.018315 (ii) 0.146525 (iii) 0.43347

14. (i) $p \leq 0.0000506$ (ii) $p = 0.0000263$

15. 49010, 980, 10

4.47 NORMAL DISTRIBUTION

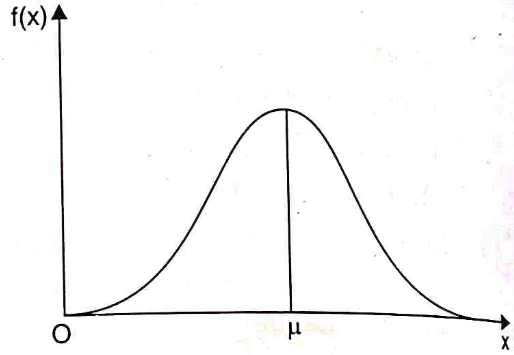
The normal distribution is a continuous distribution. It can be derived from the binomial distribution in the limiting case when n, the number of trials is very large and p, the probability of a success, is close to $\frac{1}{2}$. The general equation of the normal distribution is given by

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

where the variable x can assume all values from $-\infty$ to $+\infty$. μ and σ, called the parameters of the distribution, are respectively the mean and the standard deviation of the distribution and $-\infty < \mu < \infty$, $\sigma > 0$. x is called the normal variate and $f(x)$ is called probability density function of the normal distribution.

If a variable x has the normal distribution with mean μ and standard deviation σ, we briefly write $x : N(\mu, \sigma^2)$.

The graph of the normal distribution is called the *normal curve.* It is bell-shapped and symmetrical about the mean μ. The two tails of the curve extend to $+\infty$ and $-\infty$ towards the positive and negative directions of the x-axis respectively and gradually approach the x-axis without ever metting it. The curve is unimodal and the mode of the normal distribution coincides with its mean μ. The line $x = \mu$ divides the area under the normal curve above x-axis into two equal parts. Thus, the median of the distribution also coincides with its mean and mode. The area under the normal curve between any two given ordinates $x = x_1$ and $x = x_2$ represents the probability of values falling into the given interval. The total area under the normal curve above the x-axis is 1.



4.48 BASIC PROPERTIES OF THE NORMAL DISTRIBUTION

The probability density function of the normal distribution is given by

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

(*i*) $f(x) \geq 0$ (*ii*) $\int_{-\infty}^{\infty} f(x)\, dx = 1,$

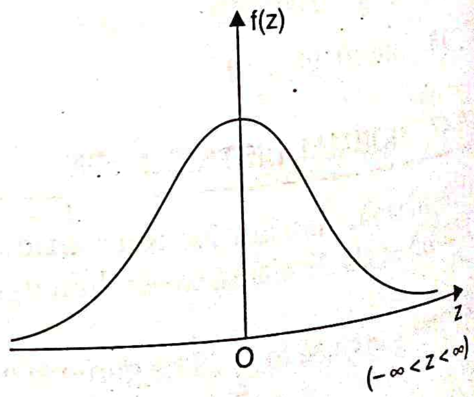
i.e., the total area under the normal curve above the x-axis is 1.

(*iii*) The normal distribution is symmetrical about its mean.

(*iv*) It is a unimodal distribution. The mean, mode and median of this distribution coincide.

4.49 STANDARD FORM OF THE NORMAL DISTRIBUTION

If X is a normal random variable with mean μ and standard deviation σ, then the random variable $Z = \frac{X - \mu}{\sigma}$ has the normal distribution with mean 0 and standard deviation 1. The random variable Z is called the *standardized* (or *standard*) normal random variable.



The probability density function for the normal distribution in standard form is given by

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$$

It is free from any parameter. This helps us to compute areas under the normal probability curve by making use of standard tables.

Note 1. If $f(z)$ is the probability density function for the normal distribution, then

$$P(z_1 \le Z \le z_2) = \int_{z_1}^{z_2} f(z)\,dz = F(z_2) - F(z_1), \quad \text{where} \quad F(z) = \int_{-\infty}^{z} f(z)\,dz = P(Z \le z)$$

The function $F(z)$ defined above is called the *distribution function* for the normal distribution.

Note 2. The probabilities $P(z_1 \le Z \le z_2)$, $P(z_1 < Z \le z_2)$, $P(z_1 \le Z < z_2)$ and $P(z_1 < Z < z_2)$ are all regarded to be the same.

Note 3. $F(-z_1) = 1 - F(z_1)$.

4.50 NORMAL DISTRIBUTION AS A LIMITING FORM OF BINOMIAL DISTRIBUTION (when $p = q$)

Let $N(q+p)^n$ be the binomial distribution. If $p = q$ then $p = q = \frac{1}{2}$ (since $p + q = 1$) and consequently the binomial distribution is symmetrical. Let n be an even integer say $2k$, k being an integer. Since $n \to \infty$, the frequencies of r and $r+1$ successes can be written in following forms:

$$f(r) = N.\,{}^{2k}C_r \left(\frac{1}{2}\right)^{2k}$$

$$f(r+1) = N.\,{}^{2k}C_{r+1} \left(\frac{1}{2}\right)^{2k}$$

$$\therefore \quad \frac{f(r+1)}{f(r)} = \frac{{}^{2k}C_{r+1}}{{}^{2k}C_r} = \frac{2k-r}{r+1}$$

The frequency of r successes will be greater than the frequency of $(r+1)$ successes if

$$f(r) > f(r+1)$$

$$\Rightarrow \quad \frac{f(r+1)}{f(r)} < 1$$

$$\Rightarrow \quad 2k - r < r + 1$$

$$\Rightarrow \quad r > k - \frac{1}{2} \qquad \ldots(1)$$

In a similar way, the frequency of r successes will be greater than the frequencies of $(r-1)$ successes if $r < k + \frac{1}{2}$ ...(2)

In view of (1) and (2), we observe that if $k - \frac{1}{2} < r < k + \frac{1}{2}$ the frequency corresponding to r successes will be the greatest. Clearly, $r = k$ is the value of the success corresponding to which the frequency is maximum. Suppose it is y_0. Then, we have

$$y_0 = N.\,{}^{2k}C_k \left(\frac{1}{2}\right)^{2k} = N.\frac{2k!}{k!\,k!}\left(\frac{1}{2}\right)^{2k}$$

Let y_x be the frequency of $k + x$ successes then, we have

$$y_x = N \,.\, {}^{2k}C_{k+x}\left(\frac{1}{2}\right)^{2k} = N \,.\left(\frac{1}{2}\right)^{2k} \cdot \frac{2k\,!}{(k+x)\,!\,(k-x)\,!}$$

Now,

$$\frac{y_x}{y_0} = \frac{k\,!\,k!}{(k+x)\,!\,(k-x)\,!} = \frac{k\,(k-1)\,(k-2)\ldots(k-x+1)}{(k+x)\,(k+x-1)\ldots(k+1)}$$

$$= \frac{\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\ldots\left\{1-\frac{x-1}{k}\right\}}{\left(1+\frac{1}{k}\right)\left(1+\frac{2}{k}\right)\ldots\left(1+\frac{x}{k}\right)}$$

Taking log on both sides,

$$\log\frac{y_x}{y_0} = \left[\log\left(1-\frac{1}{k}\right) + \log\left(1-\frac{2}{k}\right) + \cdots + \log\left(1-\frac{x-1}{k}\right)\right]$$

$$-\left[\log\left(1+\frac{1}{k}\right) + \log\left(1+\frac{2}{k}\right) + \cdots + \log\left(1+\frac{x}{k}\right)\right] \quad \ldots(3)$$

Now, writing expression for each term and neglecting higher powers of $\frac{x}{k}$ (very small quantity), we get from (3),

$$\log\frac{y_x}{y_0} = -\frac{1}{k}\{1+2+3+\cdots+(x-1)\} - \frac{1}{k}\{1+2+3+\ldots+(x-1)+x\}$$

$$= -\frac{2}{k}\{1+2+3+\cdots+(x-1)\} - \frac{x}{k}$$

$$= -\frac{2}{k}\,\frac{(x-1)\,x}{2} - \frac{x}{k} = -\frac{x^2}{k}$$

$\therefore$

$$y_x = y_0\, e^{-x^2/k}$$

$\Rightarrow$

$$\boxed{y_x = y_0\, e^{-x^2/2\sigma^2}} \qquad \Big|\, \because \ \sigma^2 = npq = \frac{n}{4} = \frac{k}{2}$$

which is **normal distribution.**

4.51 MEAN AND VARIANCE OF NORMAL DISTRIBUTION (A.K.T.U. 2015, 2018)

1. The A.M. of a continuous distribution $f(x)$ is given by

$$\text{A.M. } (\bar{x}) = \frac{\int_{-\infty}^{\infty} x\, f(x)\, dx}{\int_{-\infty}^{\infty} f(x)\, dx} \qquad \Big|\text{ By definition}$$

Consider the normal distribution with μ, σ as the parameters then

$$\bar{x} = \int_{-\infty}^{\infty} x.\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Since $\int_{-\infty}^{\infty} f(x)\, dx$ = area under normal curve = 1

Put $\frac{x-\mu}{\sigma} = z$ so that $x = \mu + \sigma z$ $\therefore$ $dx = \sigma dz$

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

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

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

$$= \mu + \frac{\sigma}{\sqrt{2\pi}} \left(\frac{e^{-z^2/2}}{-1}\right)_{-\infty}^{\infty}$$

$$\boxed{\bar{x} = \mu}$$

2. By definition,

$$\text{Variance} = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\, dx + \bar{x}^2 \int_{-\infty}^{\infty} f(x)\, dx - 2\bar{x} \int_{-\infty}^{\infty} x f(x)\, dx$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\, dx + \bar{x}^2 - 2\bar{x}\,\bar{x} \qquad \because \int_{-\infty}^{\infty} f(x)\, dx = 1 \text{ and } \int_{-\infty}^{\infty} x f(x)\, dx = \bar{x}$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\, dx - \bar{x}^2 \qquad \text{...(1)}$$

Now, Let $I = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2} dx$

Put $\frac{x-\bar{x}}{\sigma} = z$ so that $x = \bar{x} + \sigma z$ $\therefore$ $dx = \sigma\, dz$

Hence,

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

$$= \frac{1}{\sqrt{2\pi}}\left[\sigma^2 \int_{-\infty}^{\infty} z^2 e^{-z^2/2}\, dz + \bar{x}^2 \int_{-\infty}^{\infty} e^{-z^2/2}\, dz + 2\sigma\bar{x} \int_{-\infty}^{\infty} z e^{-z^2/2}\, dz\right]$$

$$= \frac{-\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z\, d(e^{-z^2/2}) + \bar{x}^2 .1 + 2\sigma\bar{x} .0$$

$$= -\frac{\sigma^2}{\sqrt{2\pi}}\left(z e^{-z^2/2}\right)_{-\infty}^{\infty} + \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z^2/2}\, dz + \bar{x}^2$$

$$= 0 + \sigma^2 \cdot 1 + \bar{x}^2 = \sigma^2 + \bar{x}^2$$

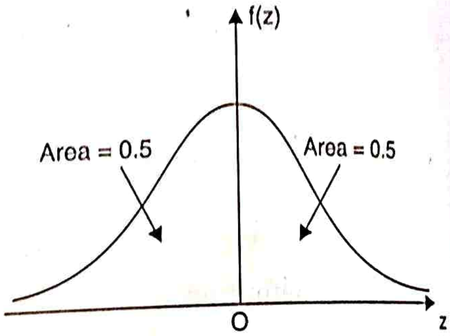
$\therefore$ From (1), $\boxed{\text{Variance} = \sigma^2 + \bar{x}^2 - \bar{x}^2 = \sigma^2}$

$\therefore$ The standard deviation of the normal distribution is σ.

4.52 AREA UNDER THE NORMAL CURVE

By taking $z = \dfrac{x-\mu}{\sigma}$, standard normal curve is formed. The total area under this curve is 1.

The area under the curve is divided into two equal parts by $z = 0$. The area between the ordinate $z = 0$ and any other ordinate can be noted from the supplied table. It should be noted that mean for the normal distribution is 0.



4.53 APPLICATIONS OF NORMAL DISTRIBUTION

De Moivre made the discovery of this distribution in 1733.

This distribution has an important application in the theory of errors made by chance in experimental measurements. Its more applications are in computation of hit probability of a shot and in statistical inference in almost every branch of science.

ILLUSTRATIVE EXAMPLES

Example 1. *A sample of 100 dry battery cells tested to find the length of life produced the following results:*

$$\bar{x} = 12 \text{ hours}, \ \sigma = 3 \text{ hours}.$$

Assuming the data to be normally distributed, what percentage of battery cells are expected to have life

(*i*) *more than 15 hours* (*ii*) *less than 6 hours* (*iii*) *between 10 and 14 hours?*

(A.K.T.U. 2018)

Sol. Here x denotes the length of life of dry battery cells.

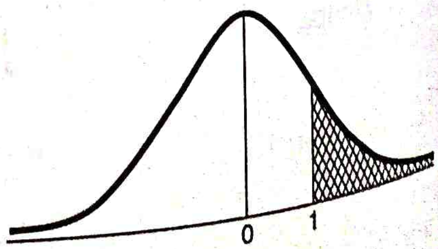
Also, $z = \dfrac{x-\bar{x}}{\sigma} = \dfrac{x-12}{3}$.

(*i*) When $x = 15$, $z = 1$

$\therefore$ $P(x > 15) = P(z > 1)$

$= P(0 < z < \infty) - P(0 < z < 1)$

$= .5 - 0.3413 = 0.1587 = 15.87\%$.

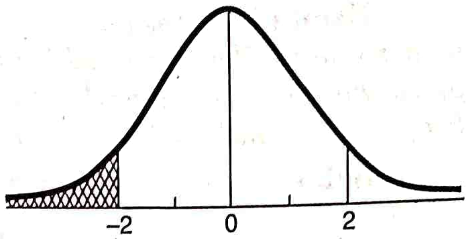


(ii) When $x = 6,\ z = -2$

$\therefore$ $P(x < 6) = P(z < -2)$

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

$= 0.5 - 0.4772 = 0.0228 = 2.28\%.$



(iii) When $x = 10, z = -\dfrac{2}{3} = -0.67$

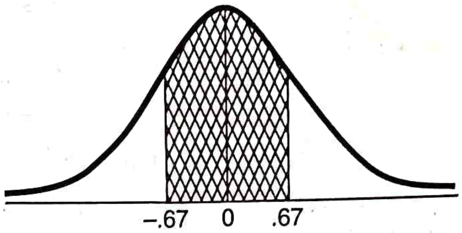
When $x = 14, z = \dfrac{2}{3} = 0.67$

$P(10 < x < 14)$

$= P(-0.67 < z < 0.67)$

$= 2P(0 < z < 0.67) = 2 \times 0.2485$

$= 0.4970 = 49.70\%.$



Example 2. *In a sample of 1000 cases, the mean of a certain test is 14 and S.D. is 2.5. Assuming the distribution to be normal, find*

(i) *how many students score between 12 and 15?*

(ii) *how many score above 18?*

(iii) *how many score below 8?*

(iv) *how many score 16?*

Sol. (i) $$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{12 - 14}{2.5} = -0.8$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{15 - 14}{2.5} = 0.4$$

Area lying between – 0.8 and 0.4

= Area between 0 to 0.8 + Area between 0 to 0.4

$= 0.2881 + 0.1554 = 0.4435$

Reqd. no. of students $= 1000 \times 0.4435 = 444$ (app.)

(ii) $$z = \frac{18 - 14}{2.5} = 1.6$$

Area right to 1.6 = 0.5 – (Area between 0 and 1.6) = 0.5 – 0.4452 = 0.0548

Reqd. no. of students $= 1000 \times 0.0548 = 54.8 \approx 55$ (app.)

(iii) $$z = \frac{8 - 14}{2.5} = -2.4$$

Area left to – 2.4 = 0.5 – (Area between 0 and 2.4) = 0.5 – 0.4918 = 0.0082

$\therefore$ Reqd. no. of students $= 1000 \times 0.0082 = 8.2 \approx 8$ (app.).

(iv) $$z_1 = \frac{15.5 - 14}{2.5} = 0.6$$

$$z_2 = \frac{16.5 - 14}{2.5} = 1$$

Area between 0.6 and 1 $= 0.3413 - 0.2257 = 0.1156$

$\therefore$ Reqd. no. of students $= 1000 \times 0.1156 = 115.6 \approx 116$ (app.).

Example 3. *Assume mean height of soldiers to be 68.22 inches with a variance of 10.8 inches square. How many soldiers in a regiment of 1,000 would you expect to be over 6 feet tall, given that the area under the standard normal curve between z = 0 and z = 0.35 is 0.1368 and between z = 0 and z = 1.15 is 0.3746.* [G.B.T.U. (C.O.) 2011]

Sol. x = 6 feets = 72 inches

$$\therefore \quad z = \frac{x-\mu}{\sigma} = \frac{72-68.22}{\sqrt{10.8}} = 1.15$$

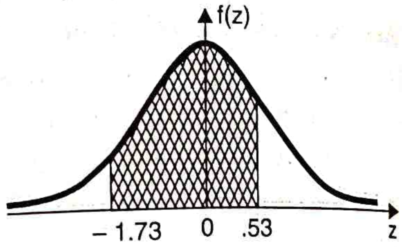
$$P(x > 72) = P(z > 1.15) = 0.5 - P(0 \le z \le 1.15)$$
$$= 0.5 - 0.3746 = 0.1254$$

$\therefore$ Expected no. of soldiers $= 1000 \times 0.1254 = 125.4 \approx 125$ (app.).

Example 4. *A large number of measurement is normally distributed with a mean 65.5″ and S.D. of 6.2″. Find the percentage of measurements that fall between 54.8″ and 68.8″.*

Sol. Mean μ = 65.5 inches, S.D. σ = 6.2 inches

x_1 = 54.8 inches, x_2 = 68.8 inches



$$\therefore \quad z_1 = \frac{x_1-\mu}{\sigma} = \frac{54.8-65.5}{6.2} = -1.73$$

and $$z_2 = \frac{x_2-\mu}{\sigma} = \frac{68.8-65.5}{6.2} = 0.53$$

Now, $$P(-1.73 \le z \le 0.53) = P(-1.73 \le z \le 0) + P(0 \le z \le 0.53)$$
$$= P(0 \le z \le 1.73) + P(0 \le z \le 0.53)$$
$$= 0.4582 + 0.2019 = 0.6601 \quad | \text{ By table}$$

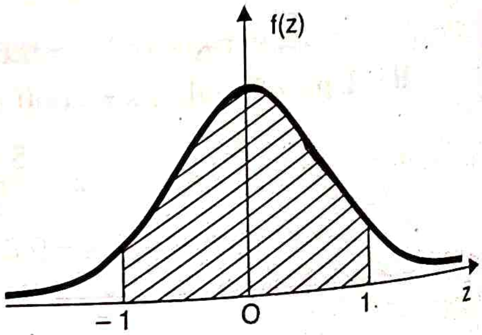
$\therefore$ Reqd. percentage of measurements = 66.01%.

Example 5. *A manufacturer knows from experience that the resistance of resistors he produces is normal with mean μ = 100 ohms and standard deviation σ = 2 ohms. What percentage of resistors will have resistance between 98 ohms and 102 ohms?*

Sol. $\mu = 100\ \Omega, \quad \sigma = 2\ \Omega, \quad x_1 = 98\ \Omega, \quad x_2 = 102\ \Omega$

$$\therefore \quad z_1 = \frac{x_1-\mu}{\sigma} = \frac{98-100}{2} = -1$$

and $$z_2 = \frac{x_2-\mu}{\sigma} = \frac{102-100}{2} = 1.$$



Now, $$P(98 < x < 102) = P(-1 < z < 1)$$
$$= P(-1 \le z \le 0) + P(0 \le z \le 1)$$
$$= P(0 \le z \le 1) + P(0 \le z \le 1)$$
$$= 0.3413 + 0.3413 = 0.6826.$$

$\therefore$ Percentage of resistors having resistance between 98 Ω and 102 Ω = 68.26%.

Example 6. *In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution. It is given that if* $f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{1}{2}x^2}\,dx$ *then f(0.5) = 0.19 and f(1.4) = 0.42.* (A.K.T.U. 2019, 2018)

Sol. Let μ and σ be the mean and S.D. respectively.

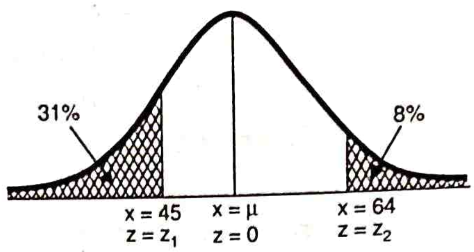
31% of the items are under 45.

$\Rightarrow$ Area to the left of the ordinate $x = 45$ is 0.31

When $x = 45$, let $z = z_1$

$P(z_1 < z < 0) = 0.5 - 0.31 = 0.19$

From the tables, the value of z corresponding to this area is 0.5



$$\therefore \quad z_1 = -0.5 [z_1 < 0]$$

When $x = 64$, let $z = z_2$

$$P(0 < z < z_2) = 0.5 - 0.08 = 0.42$$

From the tables, the value of z corresponding to this area is 1.4.

$$\therefore \quad z_2 = 1.4$$

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

$$\therefore \quad -0.5 = \frac{45 - \mu}{\sigma} \quad \text{and} \quad 1.4 = \frac{64 - \mu}{\sigma}$$

$$\Rightarrow \quad 45 - \mu = -0.5\sigma \qquad ...(1)$$

$$\text{and} \quad 64 - \mu = 1.4\sigma \qquad ...(2)$$

Subtracting $\quad -19 = -1.9\sigma \quad \therefore \quad \sigma = 10$

From (1), $\quad 45 - \mu = -0.5 \times 10 = -5 \quad \therefore \quad \mu = 50.$

Example 7. *The life of army shoes is normally distributed with mean 8 months and standard deviation 2 months. If 5000 pairs are insured, how many pairs would be expected to need replacement after 12 months?* $\left[\textit{Given that } P(z \geq 2) = 0.0228 \textit{ and } z = \frac{x - \mu}{\sigma}\right].$

(A.K.T.U. 2018)

Sol. Mean $(\mu) = 8$, Standard Deviation $(\sigma) = 2$

Number of pairs of shoes = 5000, Total months $(x) = 12$

when $x = 12$, $\quad z = \frac{x - \mu}{\sigma} = \frac{12 - 8}{2} = 2$

Area $(z \geq 2) = 0.0228$

Number of pairs whose life is more than 12 months = $5000 \times 0.0228 = 114$

Pair of shoes needing replacement after 12 months = $5000 - 114 = 4886$.

Example 8. *The mean inside diameter of a sample of 200 washers produced by a machine is 0.502 cm and the standard deviation is 0.005 cm. The purpose for which these washers are intended allows a minimum tolerance in the diameter of 0.496 to 0.508 cm, otherwise the washers are considered defective. Determine the percentage of defective washers produced by the machine. Assume the diameters are normally distributed.*

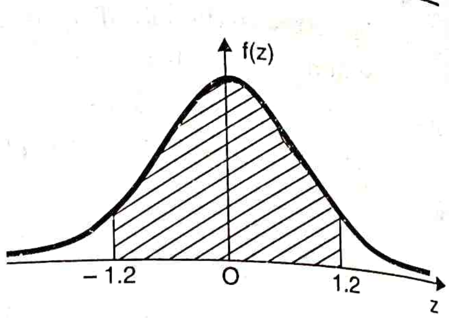
Sol. Given: Mean $\mu = 0.502$ cm, S.D. $\sigma = 0.005$ cm, $x_1 = 0.496$ cm, $x_2 = 0.508$ cm.

Now, $\quad z_1 = \frac{x_1 - \mu}{\sigma} = \frac{0.496 - 0.502}{0.005} = -1.2$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{0.508 - 0.502}{0.005} = 1.2$$

Area for non-defective washers

$$= P(-1.2 \le z \le 1.2)$$
$$= P(-1.2 \le z \le 0) + P(0 \le z \le 1.2)$$
$$= P(0 \le z \le 1.2) + P(0 \le z \le 1.2)$$
$$= 0.3849 + 0.3849 = 0.7698$$
$$= 76.98\%.$$



$\therefore$ Percentage of defective washers $= 100 - 76.98$
$$= 23.02\%.$$

Example 9. *Assuming that the diameters of 1000 brass plugs taken consecutively from a machine, form a normal distribution with mean 0.7515 cm and standard deviation 0.002 cm, how many of the plugs are likely to be rejected if the approved diameter is 0.752 ± 0.004 cm.*

Sol. Tolerance limits of the diameter of non-defective plugs are

$$0.752 - 0.004 = 0.748 \text{ cm.} \quad \text{and} \quad 0.752 + 0.004 = 0.756 \text{ cm.}$$

Standard normal variable, $z = \dfrac{x - \mu}{\sigma}$

If $x_1 = 0.748$, $z_1 = \dfrac{0.748 - 0.7515}{0.002} = -1.75$

If $x_2 = 0.756$, $z_2 = \dfrac{0.756 - 0.7515}{0.02} = 2.25$

Area from $(z_1 = -1.75)$ to $(z_2 = 2.25)$

$$= P(-1.75 \le z \le 2.25) = P(-1.75 \le z \le 0) + P(0 \le z \le 2.25)$$
$$= P(0 \le z \le 1.75) + P(0 \le z \le 2.25) = 0.4599 + 0.4878 = 0.9477$$

Number of plugs which are likely to be rejected $= 1000 \times (1 - 0.9477) = 1000 \times .0523 = 52.3$

Hence approximately 52 plugs are likely to be rejected.

Example 10. *If the heights of 300 students are normally distributed with mean 64.5 inches and standard deviation 3.3 inches, find the height below which 99% of the students lie.*

Sol. Mean $\mu = 64.5$ inches, S.D. $\sigma = 3.3$ inches

Area between 0 and $\dfrac{x - 64.5}{3.3} = 0.99 - 0.5 = 0.49$

From the table, for the area 0.49, $z = 2.327$

The corresponding value of x is given by

$$\frac{x - 64.5}{3.3} = 2.327$$

$\Rightarrow$ $x - 64.5 = 7.68$

$\Rightarrow$ $x = 7.68 + 64.5 = 72.18$ inches.

Hence 99% students are of height less than 6 ft. 0.18 inches.

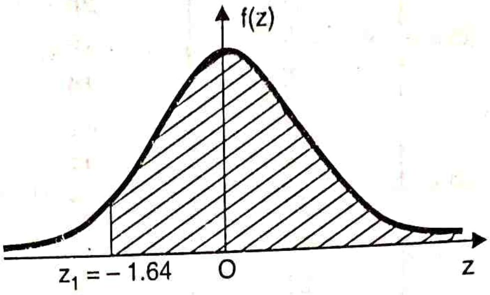
Example 11. *The income of a group of 10,000 persons was found to be normally distributed with mean ₹ 750 p.m. and standard deviation of ₹ 50. Show that, of this group, about 95% had income exceeding ₹ 668 and only 5% had income exceeding ₹ 832. Also find the lowest income among the richest 100.*

Sol. Given: $\mu = 750, \sigma = 50$

Standard normal variable, $z = \dfrac{x-\mu}{\sigma}$

(i) If $x_1 = 668$, $z_1 = \dfrac{x_1-\mu}{\sigma} = \dfrac{668-750}{50} = -1.64$

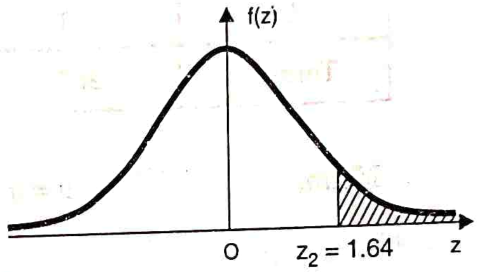
$$\begin{aligned} P(x_1 > 668) &= P(z_1 > -1.64) \\ &= 0.5 + P(-1.64 \le z \le 0) \\ &= 0.5 + P(0 \le z \le 1.64) \\ &= 0.5 + 0.4495 \\ &= 0.9495 \end{aligned}$$



$\therefore$ Required percentage of persons having income exceeding ₹ 668 = 94.95% ≈ 95% (approx.)

(ii) If $x_2 = 832$, $z_2 = \dfrac{x_2-\mu}{\sigma} = \dfrac{832-750}{50} = 1.64$

$$\begin{aligned} P(x_2 > 832) &= P(z_2 > 1.64) \\ &= 0.5 - P(0 \le z \le 1.64) \\ &= 0.5 - 0.4495 = 0.0505 \end{aligned}$$



$\therefore$ Required percentage of persons having income exceeding ₹ 832 = 5.05% ≈ 5% (approx.)

(iii) Let x be the lowest income among the richest 100 persons *i.e.,* 1% of 10,000.

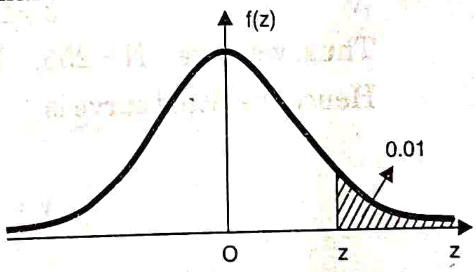
Thus, area between O and $z = 0.49$ (see figure) by Normal distribution table,

$$z = 2.33$$

Thus, $$\frac{x-\mu}{\sigma} = 2.33$$

$\Rightarrow$ $$\frac{x-750}{50} = 2.33$$

$\Rightarrow$ $$x = 866.5$$



Hence ₹ 866.5 is the minimum income among the richest 100 persons.

Example 12. *255 metal rods were cut roughly 6 inches over size. Finally the lengths of the over size amount, were measured exactly and grouped with 1 inch intervals, there being in all 12 groups* $\frac{1}{2}'' - 1\frac{1}{2}''$, $1\frac{1}{2}'' - 2\frac{1}{2}''$, ..., $11\frac{1}{2}'' - 12\frac{1}{2}''$.

The frequency distribution for the 255 lengths was as follows:

Length (inches) Central value	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Frequency	*2*	*10*	*19*	*25*	*40*	*44*	*41*	*28*	*25*	*15*	*5*	*1*

Fit a normal curve to this data.

Sol. The equation of the normal curve for N observations is

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

x	f	$u = x - 6$	fu	fu^2
1	2	– 5	– 10	50
2	10	– 4	– 40	160
3	19	– 3	– 57	171
4	25	– 2	– 50	100
5	40	– 1	– 40	40
6	44	0	0	0
7	41	1	41	41
8	28	2	56	112
9	25	3	75	225
10	15	4	60	240
11	5	5	25	125
12	1	6	6	36
Total	255		66	1300

Mean, $$\mu = a + \frac{\Sigma fu}{\Sigma f} = 6 + \frac{66}{255} = 6.259$$

Variance, $$\sigma^2 = \frac{\Sigma fu^2}{\Sigma f} - \left(\frac{\Sigma fu}{\Sigma f}\right)^2 = \frac{1300}{225} - \left(\frac{66}{255}\right)^2 = 5.031$$

$\therefore$ $$\sigma = 2.243$$

Thus, we have $N = 255$, Mean, $\mu = 6.259''$, S.D. $\sigma = 2.243''$

Hence the fitted curve is

$$y = \frac{255}{2.243\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-6.259}{2.243}\right)^2} \qquad | \text{ From (1)}$$

$$= \frac{113.68}{\sqrt{2\pi}} e^{-0.099(x-6.259)^2}.$$

Example 13. *Show that the area under the normal curve is unity.*

Sol. Area under the normal curve is given by

$$A = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Put $\frac{x-\mu}{\sigma} = z$ so $dx = \sigma\, dz$

$\therefore$ $$A = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-z^2/2} (\sigma dz) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-z^2/2} dz$$

Now, $$A \cdot A = A^2 = \left(\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-x^2/2} dx\right)\left(\sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-y^2/2} dy\right)$$

$$= \frac{2}{\pi} \int_0^{\infty}\int_0^{\infty} e^{-\left(\frac{x^2+y^2}{2}\right)} dx\, dy \qquad | \text{ where } x \text{ and } y \text{ are dummy variables}$$

Put $x = r\cos\theta$, $y = r\sin\theta$ so that $J = r$ changing to polar coordinates,

$$A^2 = \frac{2}{\pi}\int_0^{\pi/2}\int_0^{\infty} e^{-r^2/2}\, r\, dr\, d\theta = \int_0^{\infty} e^{-r^2/2}\, d\left(\frac{r^2}{2}\right) = 1$$

$\therefore$ A = Area under the normal curve = 1

Example 14. *Prove that for normal distribution, the mean deviation from the mean equals to $\frac{4}{5}$ of the standard deviation approximately.*

Sol. Let μ and σ be the mean and standard deviation of the normal distribution. Then by definition,

Mean deviation from the mean

$$= \int_{-\infty}^{\infty} |x-\mu|\, f(x)\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} |x-\mu|\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \sigma|z|\, e^{-\frac{1}{2}z^2}\, \sigma\, dz \qquad \left|\begin{array}{l} \text{when } \dfrac{x-\mu}{\sigma} = z \\ \Rightarrow \quad dx = \sigma\, dz \end{array}\right.$$

$$= \sigma\sqrt{\frac{2}{\pi}}\int_0^{\infty} z e^{-z^2/2}\, dz$$

$$= \sigma\sqrt{\frac{2}{\pi}}\left[-e^{-z^2/2}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}\,\sigma = 0.7979\,\sigma \approx 0.8\sigma \approx \frac{4}{5}\sigma$$

TEST YOUR KNOWLEDGE

1. In a test on 2000 electric bulbs, it was found that the life of a particular make, was normally distributed with an average life of 2040 hours and S.D. of 60 hours, estimate the number of bulbs likely to burn for

   (*i*) more than 2150 hours (*ii*) less than 1950 hours

   (*iii*) more than 1920 hours but less than 2160 hours.

2. An aptitude test for selecting officers in a bank is conducted on 1000 candidates. The average score is 42 and the standard deviation of score is 24. Assuming normal distribution for the scores, find

   (*i*) the number of candidates whose scores exceed 60.

   (*ii*) the number of candidates whose scores lie between 30 and 60.

3. (*i*) In a normal distribution exactly normal, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution?

   (*ii*) In a normal distribution, 0.0107 of the items lie below 42 and 0.0446 of the items lie above 82. What is the mean and standard deviation of the normal distribution?

4. If Z is a standard normal variable, find the following probabilities:

   (*i*) P(Z < 1.2) (*ii*) P(Z > – 1.2) (*iii*) P(– 1.2 < Z < 1.3).

5. An aptitude test was conducted on 900 employees of the Metro Tyres Limited in which the mean score was found to be 50 units and standard deviation was 20. On the basis of this information, you are required to answer the following questions:

(i) What was the number of employees whose mean score was less than 30?

(ii) What was the number of employees whose mean score exceeded 70?

(iii) What was the number of employees whose mean score were between 30 and 70?

$\frac{x-\mu}{\sigma}$	0.25	0.50	0.70	1.00	1.25	1.50
Area	0.0987	0.1915	0.2734	0.3413	0.3944	0.4332

[U.P.T.U. (MBA) 2009]

6. (a) Students of a class were given a mechanical aptitude test. Their marks were found to be normally distributed with mean 60 and standard deviation 5. What percent of students scored?

(i) more than 60 marks? (ii) less than 56 marks? (iii) between 45 and 65 marks?

(b) 2000 students appeared in an examination. Distribution of marks is assumed to be normal with mean $\mu = 30$ and $\sigma = 6.25$. How many students are expected to get marks?

(i) between 20 and 40 (ii) less than 35 and (iii) above 50.

[U.P.T.U. (MBA) 2012]

(c) Suppose the weight W of 600 male students are normally distributed with mean $\mu = 70$ kg and standard deviation $\sigma = 5$ kg. Find number of students with weight

(i) between 69 and 74 kg (ii) more than 76 kg. (G.B.T.U. 2013)

7. (a) In an intelligence test administered to 1000 students, the average score was 42 and standard deviation 24. Find:

(i) the expected number of students scoring more than 50.

(ii) the number of students scoring between 30 and 54.

(iii) the value of score exceeded by top 100 students. [G.B.T.U. (MBA) 2010]

(b) The average monthly sales of 5000 firms are normally distributed. Its mean and standard deviation are ₹ 36000 and ₹ 10000 respectively. Find:

(i) the no. of firms having sales over ₹ 40000.

(ii) the no. of firms having sales between ₹ 30000 and ₹ 40000.

[Given area under normal curve from 0 to z for Z (0.4) = 0.1554 and Z (0.6) = 0.2257]

[G.B.T.U. (MBA) 2010]

(c) The daily wages of 1000 workers are distributed around a mean of ₹ 140 and with a standard deviation of ₹ 10. Estimate the number of workers whose daily wages will be

(i) between ₹ 140 and ₹ 144 (ii) less than ₹ 126

(iii) more than ₹ 160. (G.B.T.U. 2012)

8. (a) Records kept by the goods inwards department of a large factory show that the average no. of lorries arriving each week is 248. It is known that the distribution approximates to be normal with a standard deviation of 26.

If this pattern of arrival continues, what percentage of weeks can be expected to have number of arrivals of:

(i) less than 229 per week? (ii) more than 280 per week?

(b) Pipes for tobacco are being packed in fancy plastic boxes. The length of the pipe is normally distributed with $\mu = 5''$ and $\sigma = 0.1''$. The internal length of the boxes is 5.2″. What is the probability that the box would be small for the pipe?

[Given that : $\phi(1.8) = 0.9641$, $\phi(2) = 0.9772$, $\phi(2.5) = 0.9938$]

(c) A manufacturer of envelopes knows that the weight of the envelopes is normally distributed with mean 1.9 gm and variance 0.01 square gm. Find how many envelopes weighing

(i) 2 gm or more

(ii) 2.1 gm or more, can be expected in a given packet of 1000 envelopes? (U.P.T.U. 2015)

[Given: if t is the normal variable, then $\phi(0 \le t \le 1) = 0.3413$ and $\phi(0 \le t \le 2) = 0.4772$]

9. (a) East-East Airlines has the policy of employing only Indian woman whose height is between 62 inches and 69 inches. If the height of Indian women is approximately normally distributed with a mean of 64 inches and a standard deviation of 3 inches, Out of the 1000 applications received find the number of applicants that would be

(i) too tall (ii) too short (iii) of acceptable height.

(b) The mean height of 500 students is 151 cm and the standard deviation is 15 cm. Assuming that the heights are normally distributed, find how many students have heights between 120 and 155 cm?

(c) The monthly mess bill of a student who is staying in the hostel follows a normal distribution with a mean of ₹ 2000 and a standard deviation of ₹ 185. What is the probability that in the next month, his bill will go above ₹ 2400? [*U.P.T.U. (MBA) 2009*]

10. In an examination taken by 500 candidates, the average and the standard deviation of marks obtained (normally distributed) are 40% and 10%. Find approximately

(i) how many will pass, if 50% is fixed as a minimum?

(ii) what should be the minimum if 350 candidates are to pass?

(iii) how many have scored marks above 60%?

11. The income distribution of workers in a certain factory was found to be normal with mean ₹ 500 and standard deviation equal to ₹ 50. There were 228 persons getting above ₹ 600. How many workers were there in all?

(Area under the standard normal curve between height at 0 and 2 is 0.4772).

12. In a certain examination, the percentage of passes and distinction were 46 and 9 respectively. Estimate the average marks obtained by the candidates, the minimum pass and distinction marks being 40 and 75 respectively. Assume the distribution of marks to be normal.

$$\left[\textbf{Hint. } \frac{40-\mu}{\sigma} = 0.1 \quad \text{and} \quad \frac{75-\mu}{\sigma} = 1.34.\right]$$

13. The marks obtained by a no. of students for a certain subject are assumed to be approximately normally distributed with mean value 65 and S.D. of 5. If 3 students are taken at random from this set, what is the probability that exactly 2 of them will have marks over 70?

14. In an examination, it is laid down that a student passes if he secures 30% or more marks. He is placed in the first, second or third division according as he secures 60% or more marks, between 45% and 60% marks and marks between 30% and 45% respectively. He gets distinction incase he secures 80% or more marks. It is noticed from the result that 10% of the students failed in the examination, whereas 5% of them obtained distinction.

Calculate the percentage of students placed in the second division. (Assume normal distribution of marks).

[**Hint.** $P(X < 30) = 0.10$, $P(X \geq 80) = 0.05$]

15. Determine the minimum marks a student must get in order to receive an A grade if the top 10% of the students are awarded A grades in an examination where the mean mark is 72 and standard deviation is 9.

16. (i) When the mean of marks was 50% and S.D. 5% then 60% of the students failed in an examination. Determine the 'grace' marks to be awarded in order to show that 70% of the students passed. Assume that the marks are normally distributed.

$$\left[\textbf{Hint. } \frac{x_1 - 0.5}{0.05} = 0.25, \quad \frac{x_2 - 0.5}{0.05} = -0.52\right]$$

(ii) In a university examination of a particular year, 60% of the students failed when mean of the marks was 50% and S.D. 5%. University decided to relax the conditions of passing by lowering the pass marks to show its result 70%. Find the minimum marks for a student to pass supposing the marks to be normally distributed and no change in the performance of students takes place.

17. How does a normal distribution differ from a binomial distribution? What are the important properties of normal distribution? [M.T.U. (MBA) 2012]

18. If the skulls are classified as A, B and C according as the length-breadth index is under 75, between 75 and 80 or over 80, find approximately (assuming that the distribution is normal) the mean and standard deviation of a series in which A are 58%, B are 38% and C are 4%, being given that if $f(t) = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-(x^2/2)}\, dx$ then $f(0.20) = 0.08$ and $f(1.75) = 0.46$.

[**Hint:** $P(X < 75) = 0.58$, $P(X > 80) = 0.04$]

19. The following table gives frequencies of occurrence of a variable X between certain limits:

Variable X	*Frequency*
Less than 40	30
40 or more but less than 50	33
50 and more	37

The distribution is exactly normal. Find the distribution and also obtain the fequency between X = 50 and X = 60.

20. The marks X obtained in Mathematics by 1000 students are normally distributed with mean 78% and standard deviation 11%.

Determine:

(*i*) how many students got marks above 90%?

(*ii*) What was the highest marks obtained by the lowest 10% of students?

(*iii*) Within what limits did the middle 90% of the students lie?

Answers

1. (*i*) 67 (*ii*) 134 (*iii*) 1909 **2.** (*i*) 227 (*ii*) 465
3. (*i*) $\bar{x} = 50.3, \sigma = 10.33$ (*ii*) $\mu = 65, \sigma = 10$
4. (*i*) 0.8849 (*ii*) 0.8849 (*iii*) 0.7881
5. (*i*) 143 (*ii*) 143 (*iii*) 614
6. (*a*) (*i*) 50%, (*ii*) 21.2%, (*iii*) 84% (*b*) (*i*) 1781, (*ii*) 1576, (*iii*) 1 (*c*) (*i*) 220, (*ii*) 69
7. (*a*) (*i*) 371, (*ii*) 383, (*iii*) 72.72 (*b*) (*i*) 1723 (*ii*) 1906 (*c*) (*i*) 155, (*ii*) 81, (*iii*) 23
8. (*a*) (*i*) 23% (*ii*) 11% (*b*) 0.0228 (*c*) (*i*) 159 (*ii*) 23
9. (*a*) (*i*) 48 (*ii*) 251 (*iii*) 701 (*b*) 294 (*c*) 0.0154
10. (*i*) 79 (*ii*) 35% (*iii*) 11 **11.** 10,000 **12.** 37.2
13. 0.06357 **14.** 34% **15.** 84 marks
16. (*i*) 3.85 (*ii*) 47.4. **18.** $\mu = 74.35, \sigma = 3.23$ **19.** $\mu = 46.12, \sigma = 11.76$, 25
20. (*i*) 138 (*ii*) 63.92% (*iii*) between 60 and 96.

ASSIGNMENT-IV

(For Section–A)

1. Define a random variable. Define expectation.

2. Two events A and B have probabilities 0.25 and 0.50 respectively. The probability that both events A and B occurs in 0.14. Find the probability that neither A nor B occurs. (M.T.U. 2013)

$\left[\textbf{Hint.}\ P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)\right]$

3. Define probability density function and cumulative distribution function.

4. What is the total probability theorem? (M.T.U. 2012)
5. Define marginal and conditional distribution.
6. Find the moment generating function of Poisson distribution. (M.T.U. 2013)
7. Find the parameters p and q of the Binomial distribution whose mean is 9 and variance is $\frac{9}{4}$.
(M.T.U. 2012)
8. If the sum of the mean and variance of a Binomial distribution of 5 trials is $\frac{9}{5}$, find $P(X \geq 1)$.
(M.T.U. 2013)
9. 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? (M.T.U. 2013)
10. If $P(X = 0) = P(X = 1) = k$ in a Poisson distribution then what is k?
11. For a Poisson variate X if $P(X = 1) = P(X = 2)$ then find $P(X = 4)$.
12. Find the total area under the curve of *p.d.f.* of a normal curve.
13. If for a Poisson distribution, $P(2) = P(3)$ then what is its probability function?
14. Find the parameters of a binomial distribution with mean = 8 and variance = 4.
15. If X is a normal variate with mean 30 and S.D. 5, find the probabilities that
(*i*) $26 \leq X \leq 40$ (*ii*) $X \geq 45$ and (*iii*) $|X - 30| > 5$
16. Define :
(*i*) Binomial distribution (*ii*) Poisson distribution (*iii*) Normal distribution
(A.K.T.U. 2018)
17. Find mean, variance and third moment about mean for the binomial distribution.
(A.K.T.U. 2018)
18. Find mean and variance of poisson distribution. (A.K.T.U. 2018)

Answers

2. 0.39	6. $M_x(t) = e^{\lambda(e^t - 1)}$	7. $q = \frac{1}{4}, p = \frac{3}{4}$	8. 0.67232
9. 0.9862	10. $\frac{1}{e}$	11. $\frac{2}{3e^2}$	12. 1
13. $\frac{e^{-3}(3)^x}{x!}$	14. $n = 16, p = \frac{1}{2}$		
15. (*i*) 0.7653	(*ii*) 0.00135	(*iii*) 0.3174.	18. $\lambda = np$