7.8. BINOMIAL PROBABILITY DISTRIBUTION (P.T.U., May 2009)

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

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.

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

('S' Stands for success and 'F' stands for failure)

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

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

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

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

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

7.9. CONDITIONS UNDER WHICH BINOMIAL DISTRIBUTION IS APPLICABLE

(P.T.U., May 2011, Dec. 2013)

Binomial distribution is $P(X = r) = {}^nC_r\, q^{n-r} p^r$, where $p + q = 1$ and $r = 0, 1, 2, 3, ..., n$

Conditions:

(i) n, the number of trials is an experiment should be finite and fixed.

(ii) In every trial, there should be only two mutually exclusive and exhaustive outcomes—success or failure.

(iii) The trials should be independent. The out come of one trial does not affect the other trial.

(iv) p, the probability of success from trial to trial is fixed and q, the probability of failure $= 1 - p$. This is same in all trials.

7.10. PROPERTIES OF A BINOMIAL DISTRIBUTION

(P.T.U., Dec. 2004)

Binomial Distribution (B.D) is $P(X = r) = {}^nC_r\, q^{n-r} p^r$ where $p + q = 1$ and $r = 0, 1, 2, 3, ..., n$

Properties:

(i) B.D is a discrete Probability Distribution in which the random variable takes only the discrete values like 0, 1, 2, 3, ...

(ii) B.D spells out, how a total probability of 1 is distributed over several values of random variable.

(iii) B.D has two constants n and p, n- the number of trials and p- the probability of success is a single trial. Entire B.D can be determined if n and p are known because $q = 1 - p$.

(iv) Mean and variance of B.D are np and npq respectively and variance is always less than the mean.

7.11. APPLICATIONS OF BINOMIAL PROBABILITY DISTRIBUTION

Binomial distribution is applied to problems concerning

(i) To find number of defectives in a sample from production line

(ii) To estimate the realibility of the system.

(iii) To find number of rounds fired from a gun hitting a target

(iv) For radar detection.

and as q ...

7.14. FITTING A BINOMIAL DISTRIBUTION

From the given data determine the following:

(*i*) n ; which is one less than the number of variates in the given data

(*ii*) N ; the total frequency *i.e.*, the sum of all the frequencies

(*iii*) μ ; the mean $= \dfrac{\Sigma f(x)}{\Sigma f}$

(*iv*) p ; the number of successes; p is obtained from $np = \mu$ *i.e.*, $p = \dfrac{\mu}{n}$.

Also, find $q = 1 - p$.

(*v*) To get expected frequencies, find the successive terms in the expansion of $N(q+p)^n$

(*vi*) Complete the table, showing variates 0, 1, 2, 3, ..., n in 1st column, expected or theoretical frequencies represented by $N\,P(r) = N\,{}^nC_r\, q^{n-r}\, p^r$ $(r = 0, 1, 2, 3, ..., n)$ in 2nd column and the expected frequencies in round figures in 3rd column.

Note. The sum total of expected frequencies should also be equal to N.

ILLUSTRATIVE EXAMPLES

Example 1. (*a*) *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?* **(P.T.U., Dec. 2003)**

(*b*) *With the usual notation, find p for a binomial variate X if n = 6 and 9 P(X = 4) = P(X = 2).* **(P.T.U., May 2010)**

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

Binomial Distribution is $\left(\dfrac{1}{9} + \dfrac{8}{9}\right)^6$

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

(*b*) The Binomial probability distribution for $n = 6$ and variate $X = r$ is

$$P(X = r) = {}^6C_r p^r q^{6-r};\; r = 0, 1, 2, 3, ..., 6$$

Given $\qquad 9P(X = 4) = P(X = 2)$

$$\Rightarrow \quad 9\,{}^6C_4\, p^4 q^2 = {}^6C_2\, p^2 q^4$$

$$\Rightarrow \quad 9\, p^4 q^2 = p^2 q^4 \qquad \because \quad {}^6C_4 = {}^6C_2 = 15$$

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

$$9\, p^2 = 1 + p^2 - 2p$$

or $$8\, p^2 + 2p - 1 = 0$$

or $$p = \frac{-2 \pm \sqrt{4+32}}{16} = \frac{-2 \pm 6}{16}$$

$$p = \frac{1}{4}, -\frac{1}{2} \quad \text{But } p \neq \frac{-1}{2} \quad \because \quad \text{probability cannot be negative}$$

$$\therefore \quad p = \frac{1}{4}.$$

Example 2. (*a*) *If on an average, one ship out of 10 is wrecked, find the probability that out of five ships expected to arrive the port, at least four will arrive safely.* **(P.T.U., Dec. 2004, Dec. 2005)**

(*b*) *A coin is tossed four times. What is the probability of getting more than two heads?*

(P.T.U., Dec. 2006)

Solution. (*a*) p, the probability of a ship arriving safely $= 1 - \frac{1}{10} = \frac{9}{10}$

$$q = 1 - \frac{9}{10} = \frac{1}{10}$$

Binomial Distribution is $\left(\frac{1}{10} + \frac{9}{10}\right)^5$

Probability that at least four ships out of five arrive safely

$$= P(4) + P(5)$$

$$= {}^5C_4 \left(\frac{1}{10}\right)^1 \left(\frac{9}{10}\right)^4 + {}^5C_5 \left(\frac{1}{10}\right)^0 \left(\frac{9}{10}\right)^5$$

$$= 5 \cdot \frac{1}{10}\left(\frac{9}{10}\right)^4 + \left(\frac{9}{10}\right)^5$$

$$= \left(\frac{9}{10}\right)^4 \left(\frac{1}{2} + \frac{9}{10}\right) = \left(\frac{9}{10}\right)^4 \frac{14}{10}$$

$$= \left(\frac{9}{10}\right)^4 \cdot \frac{7}{5} = 0.91854.$$

(*b*) p, the probability of getting head $= \frac{1}{2}$

q, the probability of getting tail $= \frac{1}{2}$

$$n = 4\,; \text{Binomial Distribution} = \left(\frac{1}{2} + \frac{1}{2}\right)^4$$

Probability of getting more than two heads $= P(3) + P(4)$

$$= {}^4C_3 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 + {}^4C_4 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4 = \left(\frac{1}{2}\right)^4 [4+1] = \frac{5}{16}.$$

Example 3. *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 three, (ii) at least three of them will be busy?*

Solution. p, the probability of a telephone number being busy between 2 P.M and 3 P.M on week-days $= \frac{1}{15}$

$$q = 1 - \frac{1}{15} = \frac{14}{15},\ n = 6;\ \text{Binomial Distribution is} \left(\frac{14}{15} + \frac{1}{15}\right)^6$$

The probability that not more than three will be busy

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

$$= {}^6C_0\left(\frac{14}{15}\right)^6 + {}^6C_1\left(\frac{14}{15}\right)^5\left(\frac{1}{15}\right) + {}^6C_2\left(\frac{14}{15}\right)^4\left(\frac{1}{15}\right)^2 + {}^6C_3\left(\frac{14}{15}\right)^3\left(\frac{1}{15}\right)^3$$

$$= \frac{(14)^3}{(15)^6}[2744 + 1176 + 210 + 20] = \frac{2744 \times 4150}{(15)^6} = 0\cdot9997$$

The probability that at least three of them will be busy

$$= P(3) + P(4) + P(5) + P(6)$$

$$= {}^6C_3\left(\frac{14}{15}\right)^3\left(\frac{1}{15}\right)^3 + {}^6C_4\left(\frac{14}{15}\right)^2\left(\frac{1}{15}\right)^4 + {}^6C_5\left(\frac{14}{15}\right)\left(\frac{1}{15}\right)^5 + {}^6C_6\left(\frac{1}{15}\right)^6 = 0.005.$$

Example 4. *If the probability of a defective bolt is 0.1, find the standard deviation (S.D.) for the defective bolts in a total of 400.* **(P.T.U., May 2006)**

Solution. $p =$ Probability of defective bolts $= 0.1 = \frac{1}{10}$

$q =$ Probability of non-defective bolts $= 1 - \frac{1}{10} = \frac{9}{10}$

$n = 400$

$$\text{S.D.} = \sqrt{npq} = \sqrt{400 \times \frac{1}{10} \times \frac{9}{10}} = 6.$$

Example 5. *The mean and variance of Binomial variable X are 2 and 1 respectively. Find the probability that X takes a value > 1.* **(P.T.U., May 2014)**

Solution. Mean of Binomial variable $X = np = 2$

Variance of Binomial variable $X = npq = 1$

Dividing;

$$q = \frac{1}{2}$$

Also we know that $p + q = 1$

$\therefore$

$$p = \frac{1}{2}$$

$$np = 2 \Rightarrow n = 4$$

The Binomial probability distribution for $n = 4$ and variate $X = r$ is

$$P(X = r) = {}^nC_r\, q^{n-r}\, p^r = {}^4C_r\left(\frac{1}{2}\right)^{4-r}\left(\frac{1}{2}\right)^r$$

$$= {}^4C_r \frac{1}{2^4} = \frac{1}{16}\,{}^4C_r$$

Now the probability that X takes *a* value > 1 is

$$P(X = 2) + P(X = 3) + P(X = 4)$$

$$= \frac{1}{16}\left\{{}^4C_2 + {}^4C_3 + {}^4C_4\right\} = \frac{1}{16}(6 + 4 + 1) = \frac{11}{16}.$$

Example 6. *Six dice the thrown 729 times. How many times do you expect at least three dice to show a five or six?* **(P.T.U., Jan 2009)**

Solution. 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 Binomial Distribution is $N\,(q + p)^n = 729\left(\frac{2}{3} + \frac{1}{3}\right)^6$

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

$$= 729\left[{}^6C_3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3 + {}^6C_4\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^4 + {}^6C_5\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^5 + {}^6C_6\left(\frac{1}{3}\right)^6\right]$$

$$= \frac{729}{3^6}[160 + 60 + 12 + 1] = 233.$$

Example 7. *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) at most two girls? Assume equal probabilities for boys and girls.*

Solution. Since probabilities for boys and girls are equal.

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

$n = 4$, $\quad N = 800 \quad \therefore$ The Binomial Distribution is $800\left(\frac{1}{2} + \frac{1}{2}\right)^4$.

(*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 the 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 *at most two girls i.e.*, having *at least 2 boys.*

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

Example 8. *What is the probability of getting the king of hearts from a pack of cards at least once in 52 cards?* **(P.T.U., Dec. 2005)**

Solution. p = the probability of getting the king of hearts from a pack of 52 cards $= \frac{1}{52}$

$\therefore \quad q = 1 - \frac{1}{52} = \frac{51}{52}$

n = the number of trials = 52

Binomial Distribution forgetting king of hearts

$$= (q + p)^{52} = \left(\frac{51}{52} + \frac{1}{52}\right)^{52}$$

Probability Distribution for getting at least one king of hearts = 1 – probability of getting no king of hearts

$$= 1 - P(X = 0)$$

$$= 1 - {}^{52}C_0\left(\frac{51}{52}\right)^{52}\left(\frac{1}{52}\right)^0$$

$$= 1 - \left(\frac{51}{52}\right)^{52} = 1 - 0.3643 - 0.6357.$$

Example 9. *The probability that a pen manufactured by a company will be defective is $\frac{1}{10}$. If 12 such pens are manufactured, find the probability that*

(i) *exactly two will be defective*

(ii) *at least two will be defective*

(iii) *none will be defective.*

Solution. p = Probability of defective pens $= \frac{1}{10}$

q = Probability of non-defective pens $= 1 - \frac{1}{10} = \frac{9}{10}$

$n = 12.$

Binomial Distribution for defective pens $= \left(\frac{9}{10} + \frac{1}{10}\right)^{12}$

(*i*) Probability that exactly two will be defective

$$= {}^{12}C_2 \left(\frac{9}{10}\right)^{10} \left(\frac{1}{10}\right)^2 = \frac{12 \times 11}{2} \frac{9^{10}}{10^{12}} = 0.2301.$$

(*ii*) Probability that at least two will be defective

$$= 1 - \text{Probability that at the most one is defective}$$

$$= 1 - [P(0) + P(1)] = 1 - \left[{}^{12}C_0 \left(\frac{9}{10}\right)^{12} + {}^{12}C_1 \left(\frac{9}{10}\right)^{11} \left(\frac{1}{10}\right)\right]$$

$$= 1 - \left(\frac{9}{10}\right)^{11} \left(\frac{9}{10} + \frac{12}{10}\right) = 1 - \frac{(9)^{11} \cdot 21}{(10)^{12}} = 0.3412.$$

(*iii*) Probability that non will be defective

$$= {}^{12}C_0 \left(\frac{9}{10}\right)^{12} \left(\frac{1}{10}\right)^0 = \left(\frac{9}{10}\right)^2 = (0.9)^{12} = 0.2833.$$

Example 10. *Fit a Binomial distribution to the following data and compare the theoretical frequencies with the actual ones*

x	*0*	*1*	*2*	*3*	*4*
f	*28*	*62*	*46*	*10*	*4*

Solution. For B.D. we have to find p which is not given

$\therefore$ We first find $\mu = \frac{\Sigma f(x)}{\Sigma f}$

$$\therefore \quad \mu = \frac{0.28 + 1.62 + 2.46 + 3.10 + 4.4}{28 + 62 + 46 + 10 + 4} = \frac{200}{150} = \frac{4}{3}$$

$\mu = np \therefore p = \frac{4}{3n}$, where $n = 4$ (one less than number of variates in the given table)

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

N = Total frequency = 150

Binomial Distribution $= 150\left(\frac{2}{3} + \frac{1}{3}\right)^4$

x	$NP(r)$	Theoretical frequency	Actual frequency
0	$150\times{}^4C_0\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^0 = 150\times\frac{16}{81} = 29.6$	30	28
1	$150\times{}^4C_1\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^1 = 150\times4\times\frac{8}{81} = 59.2$	59	62
2	$150\times{}^4C_2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^2 = 150\times6\times\frac{4}{81} = 44.4$	44	46
3	$150\times{}^4C_3\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right)^3 = 150\times4\times\frac{2}{81} = 14.8$	15	10
4	$150\times{}^4C_4\left(\frac{2}{3}\right)^0\left(\frac{1}{3}\right)^4 = 150\times\frac{1}{81} = 1.9$	2	4
		Total = 150	150

Example 11. *The Probability of a man hitting a target is $\frac{1}{4}$. If he fires 7 times, then what is the probability of his hitting the target at least twice? Also find that how many times must he fire so that the probability of his hitting the target at least once is greater than $\frac{2}{3}$?* **(P.T.U., May 2010)**

Solution. $p =$ the probability of hitting the target $= \frac{1}{4}$

$$q = 1 - p = \frac{3}{4}$$

$$n = 7$$

Binomial Distribution for hitting the target

$$= (q+p)^n = \left(\frac{3}{4}+\frac{1}{4}\right)^7$$

Probability Distribution for hitting the target at least twice

$$= 1 - P(X=0) - P(X=1)$$

$$= 1 - {}^7C_0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^7 - {}^7C_1\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^6$$

$$= 1 - \frac{3^7}{4^7} - 7\frac{3^6}{4^7} = 1 - \frac{3^6}{4^7}(10) = \frac{4547}{8192}$$

Now let he fires the target n times so that the probability of his hitting at least once is $> \frac{2}{3}$

i.e.,

$$1 - P(X=0) > \frac{2}{3}$$

or $\quad 1 - {}^nC_0\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^n > \frac{2}{3}$ or $1-\left(\frac{3}{4}\right)^n > \frac{2}{3}$ or $\frac{1}{3} > \left(\frac{3}{4}\right)^n$

or $\quad 4^n > 3^{n+1}$ which holds good for $n = 4$

$\because \quad$ for $n = 1, 2, 3, 4^n < 3^{n+1} \quad \therefore \quad n = 4$

TEST YOUR KNOWLEDGE

1. Ten coins are tossed simultaneously. Find the probability of getting at least seven heads.

   [**Hint:** $p = \frac{1}{2}, q = \frac{1}{2}, n = 10$; Required Probability = (P(7) + P(8) + P(9) + P (10))]

2. 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 at most two ships (*iii*) losing none?

   [**Hint:** $p = 0.02$, $q = 0.98$, $n = 6$; Required probability (*i*) P(1), (*ii*) P(0) + P(1) + P(2) (*iii*) P(0)]

3. The probability that a man aged 60 will live to be 70 is 0·65. What is the probability that out of ten men now 60, at least 7 would live to be 70?

   [**Hint:** $p = 0.65$, $q = 0.35$, $n = 10$, Required probability = P(7) + P(8) + P(9) + P(10)]

4. The incidence of occupational decrease 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? [**Hint:** p = Probability of suffering from disease = $\frac{20}{100} = \frac{1}{5}, q = \frac{4}{5}, n = 6$ **(P.T.U., Dec. 2003)**

   Required Probability = P(4) + P(5) + P(6)]

5. A pair of dice is thrown 200 times. If getting a sum of 9 is considered a success, find the mean and variance of number of successes. [**Hint:** p = Probability of number of successes = $\frac{4}{36} = \frac{1}{9}, q = \frac{8}{9}, n = 200$

   mean = $np = 200 \times \frac{1}{9}$, variance = $npq = \frac{200}{9} = \frac{8}{9}$]

6. If the chance that one of ten telephone lines is busy at an instant is 0.2.

   (*i*) What is the chance that 5 of the lines are busy?

   (*ii*) What is the probability that all the lines are busy?

7. If on an average 1 vessel in every 10 is wrecked, find the probability that out of 5 vessels expected to arrive, at least 4 will arrive safely.

8. Determine the Binomial distribution whose mean is 9 and standard deviation is 3/2.

9. 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 ?

   [**Hint:** p = Prob. of white balls = $\frac{5}{20}, = \frac{1}{4}$, q = Prob. of non-white balls $1 - \frac{1}{4} = \frac{3}{4}$.

   Balls to be drawn = 4 *i.e.*, $n = 4$.

POISSON DISTRIBUTION

7.15. POISSON DISTRIBUTION

Poisson distribution is related to the probability of events which are extremely rare but which have a large number of opportunities for occurrance. Poisson distribution deals with cases like: The number of persons killed in a railway accident during a particular year, the number of children born blind per year in a large city.

7.16. POISSON DISTRIBUTION AS A LIMITING CASE OF BINOMIAL DISTRIBUTION

(P.T.U., May 2005, 2007, Jan. 2009, May 2011)

Poisson distribution is a limiting case of Binomial Distribution. In B.D when parameters n and p are reasonably known, we can easily find distribution, but in situations where n is very large and p is very small then application of B.D becomes very laborious, then Poisson distribution is applied. So we apply Poisson Distribution under following conditions:

(*i*) n, the number of trials is indefinitely large *i.e.*, $n \to \infty$

(*ii*) p, the constant probability of success for each trial is indefinitely small *i.e.*, $p \to 0$

(*iii*) $np = \lambda$ (say) is finite.

7.17. RECURRENCE FORMULA FOR THE POISSON DISTRIBUTION

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

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

This is called the recurrence formula for the Poisson Distribution.

7.19. PROPERTIES OF POISSON DISTRIBUTION (P.D.)

Poisson Distribution (P.D.) is $P(r) = \dfrac{\lambda^r e^{-\lambda}}{r!}$

Properties:

(*i*) It is a discrete Probability Distribution;

(*ii*) n; the number of trials is indefinitely large *i.e.*, $n \to \infty$

(*iii*) p; the probability of success for each trial is very small *i.e.*, $p \to 0$

(*iv*) $np = \lambda$(say), the parameter is finite

(*v*) Mean and variance of P.D are equal and each $= \lambda$, the parameter

(*vi*) Sum of the probabilities $P(r)$ for $r = 1, 2, 3, \ldots \infty$ is always equal to 1.

7.20. FITTING A POISSON DISTRIBUTION

For fitting a Poisson's Distribution determine the following from the given data

(*i*) value of λ = mean of the given distribution

$$= \frac{\Sigma f\,(x)}{\Sigma f}$$

(*ii*) N = Total frequency = sum of all the frequencies.

(*iii*) Find the value of $e^{-\lambda}$ *i.e.*, calculate the frequency of zero *i.e.*, $P(X = 0)$ where $P(X = r) = \dfrac{\lambda^r e^{-\lambda}}{r!}$

(*iv*) Compute the succession N P (0), N P (1), N P (2) etc.,

(*v*) Prepare the table showing variant $r = 0, 1, 2 \ldots, n$ in the 1st column, expected or theoretical frequencies represented by NP (r) in the 2nd column and the expected frequencies in round figures in the 3rd column.

Note. The sum or total of expected frequencies should also be equal to N.

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.*

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

Recurrence relation for the Poisson distribution is

$$P(r+1) = \frac{\lambda}{r+1}P(r) = \frac{2}{r+1}P(r) \quad ...(1)$$

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

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

$$P(1) = 2P(0) = 2 \times 0.1353 = 0.2706; \qquad P(2) = \frac{2}{2}P(1) = 1.2706$$

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

Example 2. *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.*

Solution. Here $p = \frac{1}{2400}$, $n = 200$; $\quad \therefore \lambda = np = \frac{200}{2400} = \frac{1}{12} = \cdot 083$

$$\therefore \quad P(r) = \frac{\lambda^r e^{-\lambda}}{r!} = \frac{(\cdot 083)^r e^{-\cdot 083}}{r!}$$

P(at least one fatal accident) = 1 – P(no fatal accident)

$$= 1 - P(0) = 1 - \frac{(\cdot 083)^0 e^{-\cdot 083}}{0!} = 1 \cdot 92 = \cdot 08.$$

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

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

(given $e^{-0.5} = 0.61$) **(P.T.U., May 2007, 2011)**

Solution. Mean of given distribution, $\lambda = \frac{\Sigma fx}{\Sigma f}$

$$\therefore \quad \lambda = \frac{0(122) + 1(60) + 2(15) + 3(2) + 4(1)}{122 + 60 + 15 + 2 + 1} = \frac{100}{200} = \frac{1}{2} = 0.5$$

$\therefore$ Required Poisson distribution $= N\frac{\lambda^r e^{-\lambda}}{r!}$, where $N = \Sigma f = 200$

$$\therefore \quad NP(r) = 200\frac{(0.5)^r e^{-0.5}}{r!} = \frac{200(0.61)(0.5)^r}{r!} = \frac{122(0.5)^r}{r!} \quad (\text{Given } e^{-0.5} = 0.61)$$

$\therefore$ Poisson distribution is :

r	$NP(r)$	*Theoretical frequencies*
0	$P(0) = 122$	122
1	$P(1) = \dfrac{122(0.5)}{1!} = 61$	61
2	$P(2) = \dfrac{122(0.5)^2}{2!} = 15.25$	15
3	$P(3) = \dfrac{122(0.5)^3}{3!} = 2.54$	2
4	$P(4) = \dfrac{122(0.5)^4}{4!} = 0.32$	0
		Total = 200

Example 4. *If the probability that an individual suffers a bad reaction from a certain injection is 0.001. Find the probability that out of 2000 individuals*

(i) exactly 3 individuals will suffer a bad reaction

(ii) none will suffer a bad reaction **(P.T.U., May 2005)**

(iii) more than one individual will suffer **(P.T.U., Jan. 2009)**

(iv) more than two individual will suffer. **(P.T.U., Dec. 2011)**

Solution. Here $p = 0.001, n = 2000$

$\lambda = np = 2000 \times 0.001 = 2$

$$\therefore \quad P(r) = \frac{\lambda^r e^{-\lambda}}{r!} = \frac{2^r e^{-2}}{r!} = \frac{1}{e^2} \cdot \frac{2^r}{r!}$$

(i) P (exactly 3 individual will suffer a bad reaction)

$$\Rightarrow \quad P(3) = \frac{1}{e^2} \frac{2^3}{3!} = \frac{8}{e^2 . 6} = \frac{4}{3e^2} = \frac{4}{3(2.718)^2} = 0.18$$

(ii) $P(\text{none will suffer}) = P(0) = \dfrac{1}{e^2} \dfrac{2^0}{0!} = \dfrac{1}{e^2} = \dfrac{1}{(2.718)^2} = 0.135$

(iii) $P(\text{more than one}) = P(2) + P(3) + P(4) + \ldots\ldots + P(2000)$

$$= 1 - [P(0) + P(1)] = 1 - \left[\frac{2^0}{e^2 0!} + \frac{1}{e^2} \frac{2^1}{1!}\right] = 1 - \frac{1}{e^2}[1+2] = 1 - 3(0.135) = 0.595.$$

(iv) $P(\text{more than two}) = 1 - [P(0) + P(1) + P(2)] = 1 - \left[\dfrac{2^0}{e^2 0!} + \dfrac{2^1}{e^2 1!} + \dfrac{2^2}{e^2 2!}\right]$

$$= 1 - \frac{1}{e^2}[1 + 2 + 2]$$

$$= 1 - 5(0.135)$$

$$= 0.325$$

Example 5. *Six coins are tossed 1600 times. Using the Poisson distribution, determine the approximate probability of getting six heads x times.*

Solution. 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 $= np = 1600 \times \frac{1}{64} = 25$

$\therefore$ The mean of the Poisson distribution $= \lambda = 25$.

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

$$= \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(25)^x e^{-25}}{x!}.$$

Example 6. *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 fifty first birthday?*

Solution. $p =$ the probability that a man aged 50 years will die within a year $= 0.01125$

$n = 12$

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

Let X denote the number of men aged fifty years who will die within a year.

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

P (at least 11 will reach their 51^{st} birthday)

= P (at the most one man dies within a year)

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

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

$$= e^{-0.135}(1 + 0.135) = e^{-0.135}(1.135)$$

$$= (0.8731)(1.135) = 0.9916..$$

Example 7. *A product is 0.5% defective and is packed in cartoons of 100. What percentage contains not more than 3 defective?* **(P.T.U., May 2004)**

Solution. $p =$ The probability that a product is defective $= 0.5\%$

$$\therefore \quad p = \frac{0.5}{100} = 0.005$$

$n = 100$

$\lambda = np = 100\,(0.005) = 0.5$

Let X denotes the number of defective products

$$\therefore \quad P(X = r) = \frac{\lambda^r e^{-\lambda}}{r!} = \frac{(0.5)^r e^{-0.5}}{r!}$$

P(not more than 3 defective) $= P(X \leq 3) = P(0) + P(1) + P(2) + P(3)$

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

$$= e^{-0.5}\left[1 + (0.5) + \frac{0.25}{2} + \frac{0.125}{6}\right]$$

TEST YOUR KNOWLEDGE

1. Fit a Poisson distribution to the following:

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

2. If the probability of a bad reaction from a certain injection is 0.001, determine the chance that out of 2000 individuals more than two will get a bad reaction. **(P.T.U., Jan. 2009)**

3. If X is a Poisson variate such that P(X = 2) = 9P(X = 4) + 90P(X = 6), find the standard deviation.

4. If a random variable has a Poisson distribution such that P(1) = P(2), find

   (*i*) mean of the distribution (*ii*) P(4).

5. 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).

6. 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.

   [Hint: $p=\frac{2}{100}, n=500 \therefore \lambda=np=10, e^{-10}=.000045$ Required probability P(15)]

7. 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?

   [Hint: $p=\frac{10}{100}=\frac{1}{10}, n=7, \lambda=\frac{7}{10}; P(r)=\frac{e^{-\lambda}\lambda^r}{r!}$ Required probability = P(5) + P(6) + P(7)]

8. 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
Frequencies:	109	65	22	3	1	200

   Fit a Poisson distribution to the data and calculate the theoretical frequencies. **(P.T.U., 2005)**

9. 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. $(e^{-1.5}=0.2231)$

10. 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

11. The probability that a man aged 35 years will die before reaching the age of 40 years may be taken as 0.018. Out of a group of 400 men, now aged 35 years, what is probability that 2 men will die within the next 5 years?

   [Hint: $p=0.018, n=400, \lambda=np=7.2, e^{-\lambda}=e^{-7.2}=0.0007466$]

7.22. NORMAL DISTRIBUTION

(P.T.U., May 2005, May 2007)

The normal distribution is a continuous distribution. Any quantity whose variation depends on random causes is distributed according to the normal law. Its importance lies in the fact that a large number of distributions approximate to the normal 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,

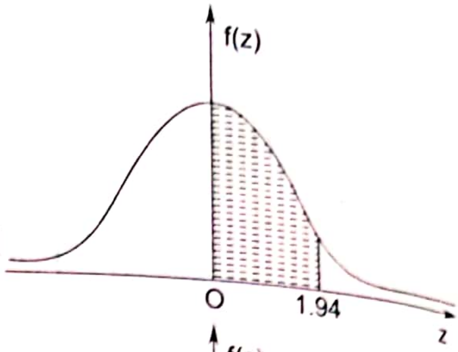
Example 1. *Find the area under the normal curve in the following cases :*

(*i*) $z = 0$ and $z = 1.94$

(*ii*) $z = -0.46$ and $z = 0$

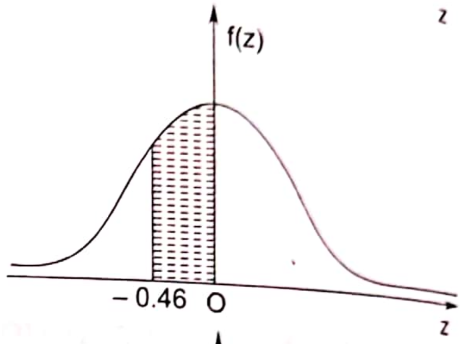
(*iii*) $z = -0.68$ and $z = 2.21$

(*iv*) *To the left of* $z = -0.6$.



Solution. (*i*) Required area $= P(0 \le z \le 1.94)$

$= 0.4738$ (From the table of Normal curve given at the end of the book)

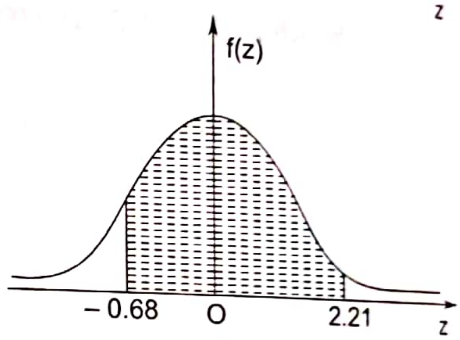


(*ii*) Required area

$$= P(-0.46 \le z \le 0)$$
$$= P(0 \le z \le 0.46)$$
$$= 0.1772.$$

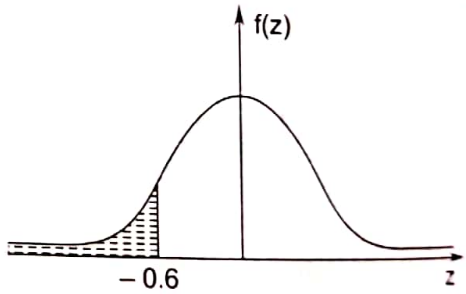
$$= 2\int_0^{\infty} \frac{(x-\mu)^2}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

($\because$ Integrand is even)



(*iii*) Required area

$$= P(-0.68 \le z \le 2.21)$$
$$= P(-0.68 \le z \le 0) + P(0 \le z \le 2.21)$$
$$= P(0 \le z \le 0.68) + P(0 \le z \le 2.21)$$
$$= 0.2518 + 0.4865$$
$$= 0.7383.$$



(*iv*) Required area

$$= P(-\infty \le z \le -0.6)$$
$$= 0.5 - P(-0.6 \le z \le 0)$$
$$= 0.5 - P(0 \le z \le 0.6)$$
$$= 0.5 - 0.2257$$
$$= 0.2743.$$

Example 2. *Find the value of z in each of the following cases:*

(*i*) *Area between 0 and z is 0.3621.* (*ii*) *Area to the left of z is 0.7642.*

Solution. (*i*) As 0.3621 is < 0.5

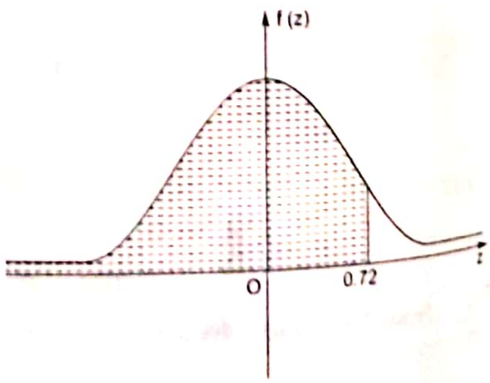
$\therefore$ Area is either on the right or on the left hand side of $z = 0$.

From table, we see that this area 0.3621 corresponds to

$z = 1.0 + 0.09$

$\therefore$ $z = \pm(1.0 + 0.09) = \pm 1.09$

$\therefore$ $-1.09 \le z \le 0$ or $0 \le z \le 1.09$



(*ii*) As $0.7642 > 0.5$

$\therefore$ Area between 0 and z is

$$= 0.7642 - 0.5 = 0.2642$$

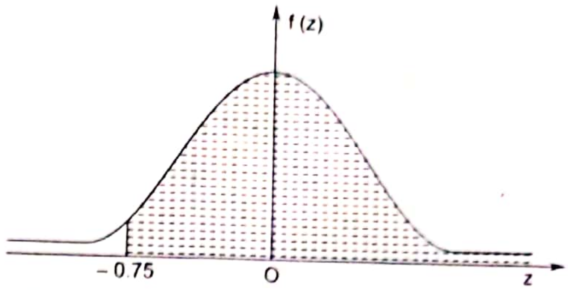
From the table 0.2642 corresponds to $z = 0.7 + 0.02 = 0.72$

$\therefore$ z lies between $-\infty$ and 0.72

i.e., $-\infty < z \le 0.72$ or simply $z \le 0.72$

Example 3. *Let x be normal random variable with mean 10 and S.D. 4. Determine the probability $P(x \geq 7)$.*

Solution. When $x = 7$, $z = \dfrac{x-\mu}{\sigma} = \dfrac{7-10}{4} = -0.75$



$$\therefore \quad P(x \geq 7) = P(z \geq -0.75) = P(-0.75 \leq z \leq 0) + P(0 \leq z < \infty)$$
$$= P(0 \leq z \leq 0.75) + 0.5$$
$$= 0.2734 + 0.5$$
$$= 0.7734.$$

Example 4. *A sample of 100 dry battery cells tested to find the length of life produced the following results:*

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

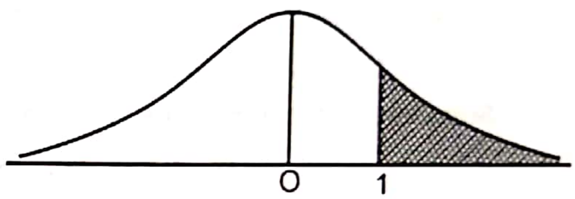
Assuming the data to be normally distributed, what percentage of battery cells are expected to have life.

(*i*) *more then 15 hours* (*ii*) *less than 6 hours*

(*iii*) *between 10 and 14 hours?*

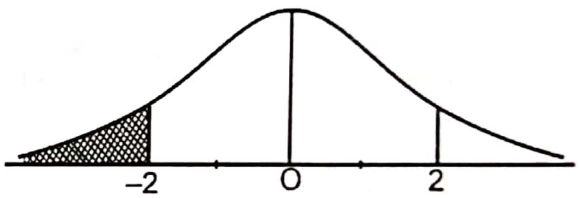
Solution. Here, x denoted the length of life of dry battery cells.

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



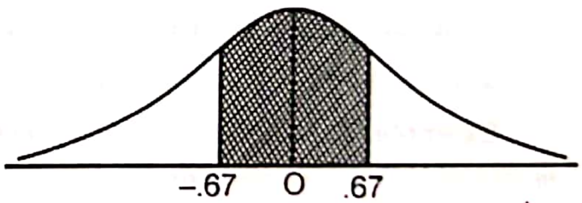
(*i*) When $x = 15$, $z = 1$

$$\therefore \quad P(x > 15) = P(z > 1)$$
$$= P(0 < z < \infty) - P(0 < z < 1)$$
$$= .5 - 0.3413 = 0.1587 = 15.87\%.$$



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

$$\therefore \quad P(x < 6) = P(z < -2)$$
$$= P(z > 2) = P(0 < z < \infty) - P(0 < z < 2)$$
$$= 5 - 0.4772 = 0.0228 = 2.28\%.$$



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

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

$$P(10 < x < 14) = P(-0.67 < z < 0.67)$$
$$= 2P(0 < z < 0.67) = 2 \times 0.2487 = 0.4974 = 49.74\%.$$

Example 5. *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.* **(P.T.U., May 2006, Dec. 2013)**

Solution. Let $\bar{x}$ 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 $z = z_1$

$$P(z_1 < z < 0) = .5 - .31 = .19$$

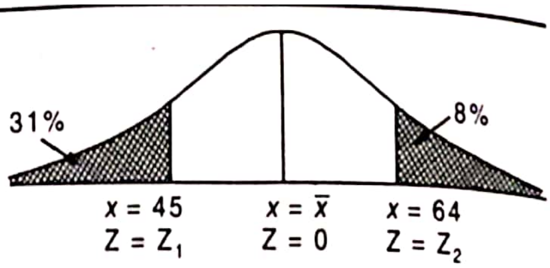
From the tables, the value of z corresponding to this area is 0.5

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

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

$$P(0 < z < z_2) = .5 - .08 = .42$$

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



$$z_2 = 1.4$$

Since, $z = \dfrac{x - \bar{x}}{\sigma}$

$$-0.5 = \frac{45 - \bar{x}}{\sigma} \quad \text{and} \quad 1.4 = \frac{64 - \bar{x}}{\sigma}$$

$\Rightarrow$ $\quad 45 - \bar{x} = -0.5\sigma$ ...(1)

and $\quad 64 - \bar{x} = 1.4\sigma$ ...(2)

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

From (1), $\quad 45 - \bar{x} = -0.5 \times 10 = -5 \qquad \therefore \bar{x} = 50.$

Example 6. *On a statistics examination the mean score was 78 and S.D was 10.*

(i) Determine standard score of two boys whose score was 93 and 62 respectively.

(ii) Determine the score of two students whose standard score was – 0.6 and 1.4 respectively.

Solution. Mean score $= \mu = 78$ and S.D $= \sigma = 10$

$\therefore \qquad z = \dfrac{x - 78}{10}$ ...(1) where x stands for score and z stands for standard score.

(*i*) When $x = 93$, then $z = \dfrac{93 - 78}{10} = 1.5$

when $x = 62$, $\qquad z = \dfrac{62 - 78}{10} = -1.6$

(*ii*) When $z = -0.6$, $x = 10z + 78 = -6 + 78 = 72$

when $z = 1.4$, $\qquad x = (1.4)\,10 + 78 = 14 + 78 = 92.$

Example 7. *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?

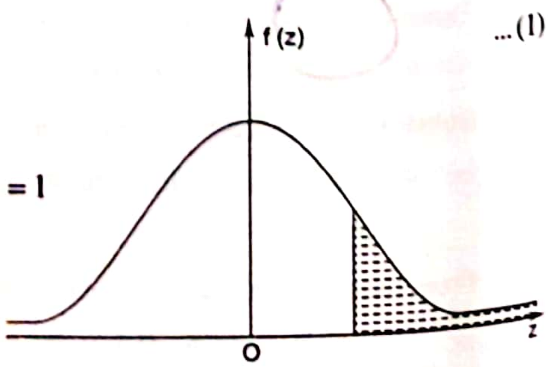
(ii) what should be the minimum if 350 candidates are to pass?

(iii) how many have scored marks above 60%? **(P.T.U., May 2004, Dec. 2011)**

Solution. Average deviation *i.e.*, mean $\mu = 40\% = 0.4$

Standard deviation $\quad$ *i.e.*, $\sigma = 10\% = 0.1$

$\therefore \qquad z = \dfrac{x - \mu}{\sigma} = \dfrac{x - 0.4}{0.1}$ ...(1)



(*i*) When 50% is fixed as minimum marks

then $\qquad x \geq 50\% = 0.5 \quad \therefore z \geq \dfrac{0.5 - 0.4}{0.1} = \dfrac{0.1}{0.1} = 1$

Probability of pass students $= P(x \geq 0.5) = P(z \geq 1)$

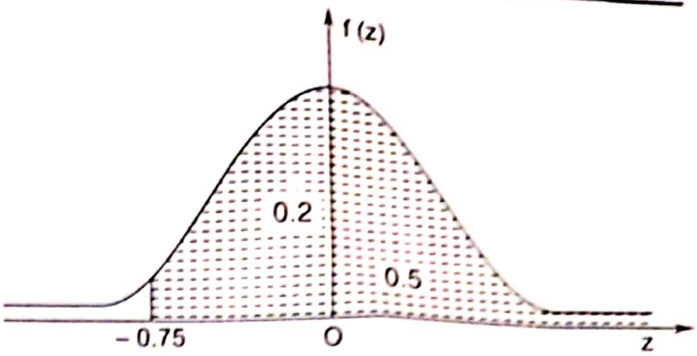
$$= 0.5 - P(0 \leq z \leq 1)$$

$$= 0.5 - 0.3413 = 0.1587$$

Total number of pass students

$$= 500 \times (0.1587) = 79.35 = 79 \text{ students.}$$

(ii) Number of pass students 350



Probability of pass students $= \dfrac{350}{500} = 0.7.$

As $\quad 0.7 > 0.5$

$\therefore$ Area 0.7 = area 0.2 + area 0.5

Area 0.2 will be on the left hand side of O

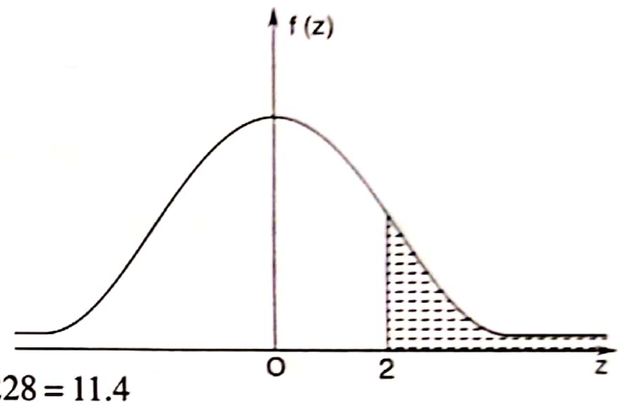
$\therefore$ From tables we see that 0.2 on the LHS of 'O' corresponds to the value $z = -0.53$

From (1), $\quad -0.53 = \dfrac{x - 0.4}{0.1} \quad$ or $\quad x = 0.347$

$$= 34.7\%.$$

$\therefore$ If 350 candidates are to pass then minimum % of marks required is 34.7 *i.e.*, 35%.

(iii) When $x \geq 60\% = 0.6$



then $\quad z = \dfrac{0.6 - 0.4}{0.1} = 2$

$\therefore$ Probability of students getting more than 60%

$$= P(x \geq 0.6) = P(z \geq 2)$$
$$= 0.5 - P(0 \leq z \leq 2)$$
$$= 0.5 - 0.4772 = 0.228.$$

Total number of students getting more than 60% $= 500 \times 0.0228 = 11.4$

i.e., required number of students = 11.

TEST YOUR KNOWLEDGE

1. The mean height of 500 male students in a certain college is151 cm and the standard deviation is 15 cm. Assuming the heights are normally distributed, find how many students have heights between 120 and 155 cm?
2. An aptitude test from 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. **[Hint:** Consult example 4]
3. (a) In a normal distribution, 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviation of the distribution? **[Hint:** Consult solved example 5]
   (P.T.U., May 2014)
   (b) In a normal distribution 12% of the items are under 30 and 55% are under 60. Find the mean and S.D of the distribution. **[Hint:** $z_1 = -1.175$; $z_2 = 1.0365$] **(P.T.U., Jan. 2008)**
4. Let X denote the number of scores on a test. If X is normally distributed with mean 100 and standard deviation 15, find the probability that X does not exceed 130.
5. It is known ... community