Chapter 3

Theoretical Probability Distributions

3.1 Introduction

Probability distributions are either discrete or continuous, depending on whether they define probabilities for discrete or continuous variables. Here we shall confine our studies to Binomial, Poisson and Normal distributions of which Binomial and Poisson distributions are discrete distributions whereas Normal distribution is a continuous probability distribution.

3.2 Binomial Distribution

A series of independent trials which result in one of the two mutually exclusive outcomes 'success' or 'failure' such that the probability of the success (or failure) in each trials is constant, then such repeated independent trials are called as 'Bernoulli trials'. A discrete random variable which results in only one of the two possible outcomes (success or failure) is called Binomial variable.

Let there be n independent finite trials in an experiment such that

i. Each trial has only two possible outcomes success and failure
ii. Probability of success (p) and probability of failure (q) are constant for all the trials and $p + q = 1$.

Then if a random variable X denotes the number of successes in n trials, then

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

$\therefore$ Binomial distribution may be given as $(q + p)^r$

3.2.1 Mean of Binomial Distribution

Mean $= \sum_{r=0}^{n} rP(r)$

$$= \sum_{r=0}^{n} r\, {}^nC_r\, q^{n-r} p^r$$

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

$$= n\, q^{n-1}p + \frac{2n(n-1)}{2!}\, q^{n-2}p^2 + \frac{3n(n-1)(n-2)}{3!}\, q^{n-3}p^3 + \cdots + np^n$$

$$= np\, \left[q^{n-1} + (n-1)\, q^{n-2}p + \frac{(n-1)(n-2)}{2!}\, q^{n-3}p^2 + \cdots + p^{n-1}\right]$$

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

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

$$= np \qquad\qquad \because q + p = 1$$

3.2.2 Variance of Binomial Distribution

Variance $= \sum_{r=0}^{n} r^2 P(r) \ - \ (\text{mean})^2$

Now $\sum_{r=0}^{n} r^2 P(r) = \sum_{r=0}^{n} r^2\ {}^{n}C_r\ q^{n-r}p^r$

$$= \ {}^{n}C_1\ q^{n-1}p + \ 2^2\ {}^{n}C_2\ q^{n-2}p^2 + \ 3^2\ {}^{n}C_3\ q^{n-3}p^3 + \cdots + n^2\ {}^{n}C_n\ p^n$$

$$= n\ q^{n-1}p + \frac{4n(n-1)}{2!}\ q^{n-2}p^2 + \frac{9n(n-1)(n-2)}{3!}\ q^{n-3}p^3 + \cdots + n^2p^n$$

$$= np\ \left[q^{n-1} + 2(n-1)\ q^{n-2}p + \frac{3(n-1)(n-2)}{2!}\ q^{n-3}p^2 + \cdots + np^{n-1}\right]$$

$$= np\ \left[q^{n-1} + (n-1)q^{n-2}p + \frac{(n-1)(n-2)}{2!}\ q^{n-3}p^2 + \cdots + p^{n-1}\right]$$

$$+np\ [(n-1)q^{n-2}p + (n-1)(n-2)\ q^{n-3}p^2 + \cdots + (n-1)p^{n-1}]$$

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

$$+np(n-1)p\ [\ {}^{n-2}C_0\ q^{n-2} + \ {}^{n-2}C_1\ q^{n-3}p + \cdots + \ {}^{n-2}C_{n-2}p^{n-2}]$$

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

$$= np[1 + (n-1)p] \qquad\qquad \because q + p = 1$$

$\therefore$ Variance $= np[1 + np - p] - n^2p^2 = np[q + np] - n^2p^2 \quad \because 1 - p = q$

$$= npq$$

$\therefore$ For Binomial distribution

Mean $= np$

$$\mu_1 = 0$$

$$\mu_2 = \ \sigma^2 = \ npq \qquad\qquad \because \ \ \mu_2 = \ variance\ (\sigma^2)$$

Similarly $\mu_3 = npq(q - p)$

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

$$\beta_1 = \frac{u_3^2}{u_2^3} = \frac{[npq(q-p)]^2}{(npq)^3} = \frac{(q-p)^2}{npq} \;, \qquad \gamma_1 = \frac{(q-p)}{\sqrt{npq}}$$

$$\beta_2 = \frac{\mu_4}{u_2^2} = \frac{1+3pq(n-2)}{npq} = 3 + \frac{1-6pq}{npq} \quad \gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}$$

Remarks:

- Binomial Distribution is symmetrical if $\beta_1 = 0$ i.e. if $\frac{(q-p)^2}{npq} = 0$ or $p = q = \frac{1}{2}$
- Binomial Distribution is positively skewed if $\gamma_1 > 0$ i.e. if $q - p > 0$ or $1 - 2p > 0$ or $p < \frac{1}{2}$
- Binomial Distribution is negatively skewed if $p > \frac{1}{2}$
- Since $0 < q < 1$, $\therefore$ for binomial distribution $npq < np$ i.e. $\text{vaiance} < \text{mean}$.

3.2.3 Moment Generating Function of Binomial Distribution

Moment Generating Function (MGF) about origin is expected value of e^{tr}

$$M_0(t) = E(e^{tr})$$

$$= \sum p(r)e^{tr}$$

$$= \sum {}^nC_r\, p^r q^{n-r} e^{tr}$$

$$= \sum {}^nC_r\, (pe^t)^r\, q^{n-r}$$

$$= (q + pe^t\,)^n$$

Now $\left[\frac{d}{dt} M_0(t)\right]_{t=0} = [n(q + pe^t\,)^{n-1} pe^t]_{t=0} = n(q + p\,)^{n-1} p = np$

$$\Rightarrow \mu_1' = np$$

Again $M_a(t) = e^{-at} M_0(t)$

$\therefore M_m(t) = e^{-npt}(q + pe^t\,)^n$ where $m = np$ denotes mean of the distribution

$$= (qe^{-pt} + pe^{qt}\,)^n$$

$$= \left[(\boldsymbol{q} + \boldsymbol{p}) + (-\boldsymbol{qpt} + \boldsymbol{pqt}) + \left(\frac{\boldsymbol{qp^2t^2}}{\mathbf{2!}} + \frac{\boldsymbol{pq^2t^2}}{\mathbf{2!}}\right) + \left(\frac{\boldsymbol{-qp^3t^3}}{\mathbf{3!}} + \frac{\boldsymbol{pq^3t^3}}{\mathbf{3!}}\right) + \left(\frac{\boldsymbol{qp^4t^4}}{\mathbf{4!}} + \frac{\boldsymbol{pq^4t^4}}{\mathbf{4!}}\right) + \cdots\right]^{\boldsymbol{n}}$$

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

$$= \mathbf{1} + n\left(\boldsymbol{pq}\frac{\boldsymbol{t^2}}{\mathbf{2!}} + \boldsymbol{pq}(\boldsymbol{q^2} - \boldsymbol{p^2})\frac{\boldsymbol{t^3}}{\mathbf{3!}} + \boldsymbol{pq}(\boldsymbol{q^3} + \boldsymbol{p^3})\frac{\boldsymbol{t^4}}{\mathbf{4!}} + \cdots\right)$$

$$+ \frac{n(n-1)}{2!}\left(\boldsymbol{pq}\frac{\boldsymbol{t^2}}{\mathbf{2!}} + \boldsymbol{pq}(\boldsymbol{q^2} - \boldsymbol{p^2})\frac{\boldsymbol{t^3}}{\mathbf{3!}} + \boldsymbol{pq}(\boldsymbol{q^3} + \boldsymbol{p^3})\frac{\boldsymbol{t^4}}{\mathbf{4!}} + \cdots\right)^2 + \cdots$$

$$\therefore M_m(t) = \mathbf{1} + \boldsymbol{npq}\frac{\boldsymbol{t^2}}{\mathbf{2!}} + \boldsymbol{npq}(\boldsymbol{q} - \boldsymbol{p})\frac{\boldsymbol{t^3}}{\mathbf{3!}} +$$

$$[\boldsymbol{npq}(\boldsymbol{q^2} + \boldsymbol{p^2} - \boldsymbol{pq}) + \mathbf{3}n(n-1)\boldsymbol{p^2q^2}]\frac{\boldsymbol{t^4}}{\mathbf{4!}} + \cdots \quad \because (\boldsymbol{q} + \boldsymbol{p}) = \mathbf{1}$$

Now $M_m(t) = \mathbf{1} + \boldsymbol{\mu_1 t} + \boldsymbol{\mu_2}\frac{\boldsymbol{t^2}}{\mathbf{2!}} + \boldsymbol{\mu_3}\frac{\boldsymbol{t^3}}{\mathbf{3!}} + \boldsymbol{\mu_4}\frac{\boldsymbol{t^4}}{\mathbf{4!}}$

$$\therefore \mathbf{1} + \boldsymbol{\mu_1 t} + \boldsymbol{\mu_2}\frac{\boldsymbol{t^2}}{\mathbf{2!}} + \boldsymbol{\mu_3}\frac{\boldsymbol{t^3}}{\mathbf{3!}} + \boldsymbol{\mu_4}\frac{\boldsymbol{t^4}}{\mathbf{4!}} = \mathbf{1} + \boldsymbol{npq}\frac{\boldsymbol{t^2}}{\mathbf{2!}} + \boldsymbol{npq}(\boldsymbol{q} - \boldsymbol{p})\frac{\boldsymbol{t^3}}{\mathbf{3!}} +$$

$$[\boldsymbol{npq}(\boldsymbol{q^2} + \boldsymbol{p^2} - \boldsymbol{pq}) + \mathbf{3}n(n-1)\boldsymbol{p^2q^2}]\frac{\boldsymbol{t^4}}{\mathbf{4!}} + \cdots$$

Comparing coefficients of different powers of t on both sides, we get

$$\mu_1 = 0$$

$$\mu_2 = npq$$

$$\mu_3 = npq(q-p)$$

$$\mu_4 = \boldsymbol{npq}(\boldsymbol{q^2} + \boldsymbol{p^2} - \boldsymbol{pq}) + \mathbf{3}n(n-1)\boldsymbol{p^2q^2}$$

$$= \boldsymbol{npq}(\mathbf{1} - \mathbf{3}\boldsymbol{pq}) + \mathbf{3}n(n-1)\boldsymbol{p^2q^2} \qquad \because \boldsymbol{q^2} + \boldsymbol{p^2} = (\boldsymbol{q} + \boldsymbol{p})^{\mathbf{2}} - \mathbf{2}\boldsymbol{pq}$$

$$= \boldsymbol{npq}(\mathbf{1} - \mathbf{3}\boldsymbol{pq} + \mathbf{3}\boldsymbol{npq} - \mathbf{3}\boldsymbol{pq})$$

$$= npq[1 + 3pq(n-2)]$$

Example1 Show that Variance of a binomial distribution is less than or equal to $\frac{n}{4}$.

Solution: Variance (σ^2) of a binomial distribution is npq

$$\therefore \sigma^2 = npq = np(1-p) = np - np^2 = f(p) \text{ say}$$

For $f(p)$ to be maximum

$$f'(p) = 0 \text{ and } f''(p) < 0$$

Now $f'(p) = n - 2np = 0 \Rightarrow p = \frac{1}{2}$

Also $f''(p) = -2n < 0$

$\therefore$ $f(p)$ is maximum at $p = \frac{1}{2}$

Thus maximum variance is at $p = \frac{1}{2}$, $q = \frac{1}{2}$

i.e. maximum variance $= npq = n.\frac{1}{2}.\frac{1}{2} = \frac{n}{4}$ $\therefore$ Variance $\leq \frac{n}{4}$

Example2 6 dice are thrown 729 times. How many times would you expect at least three dice to show 1 or 2?

Solution: Here the Binomial Distribution (B.D.) is given by $N(q+p)^n$

Where $p = \frac{2}{6} = \frac{1}{3}, = \frac{2}{3}, n = 6, N = 729$

$\therefore$ B.D. is given by $729\left(\frac{2}{3} + \frac{1}{3}\right)^6$ and if X is random variable showing number of successes, then

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

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

Example 3 A die is tossed 3 times. Find mean and variance of number of successes if getting 5 or 6 is considered as success.

Solution: Here $p = \frac{2}{6} = \frac{1}{3}$, $q = \frac{2}{3}, n = 3$

$\therefore$ Mean $= np = \frac{3}{3} = 1$, Variance $= npq = 3 \times \frac{1}{3} \times \frac{2}{3} = \frac{2}{3}$

Example4 If the sum of mean and variance of Binomial Distribution is 4.8 for 5 trials. Find the distribution.

Solution: Given $np + npq = 4.8, n = 5$

$\Rightarrow np(1+q) = 4.8$

$\Rightarrow 5(1-q)(1+q) = 4.8 \quad \because n = 5$

$\Rightarrow 1 - q^2 = 9.6 \;\; \Rightarrow q^2 = \frac{1}{25} \;\; \Rightarrow q = \frac{1}{5}$

$\therefore q = \frac{1}{5}$, $p = \frac{4}{5}$ and distribution is given by $\left(\frac{1}{5} + \frac{4}{5}\right)^5$

Example5 For a binomial distribution; mean is 4 and standard deviation is $\sqrt{2}$. Find the distribution.

Solution: Given $np = 4$, $npq = 2$

$\Rightarrow \frac{npq}{np} = \frac{2}{4} \;\; \Rightarrow q = \frac{2}{4} = \frac{1}{2} \therefore p = \frac{1}{2}$

Also $np = 4$ or $\frac{n}{2} = 4 \Rightarrow n = 8$

$\therefore$ The distribution is given by $\left(\frac{1}{2} + \frac{1}{2}\right)^8$

Example6 Find the expected number of the defective bulbs in a lot of 100 bulbs; if one out of 5 bulbs is defective. Also find the standard deviation, coefficient of skewness γ_1 and determine whether the distribution curve is leptokurtic, mesokurtic or platykurtic.

Solution: We have $p = \frac{1}{5} = 0.2$, $q = 1 - 0.2 = 0.8, n = 1000$

Expected number of defective bulbs $= np = 100 \times 0.2 = 20$

Also standard deviation $= \sqrt{npq} = \sqrt{100 \times 0.2 \times 0.8} = 4$

$$\gamma_1 = \frac{(q-p)}{\sqrt{npq}} = \frac{(0.8-0.2)}{4} = 0.15$$

$$\beta_2 = 3 + \frac{1-6pq}{npq} = 3 + \frac{1-6(0.2)(0.8)}{16} = 3.0025$$

$\therefore$ The curve is a bit leptokurtic.

Example7 If the probability of success of an event is $\frac{1}{20}$; how many trials are required in order that the probability of getting at least one success, is just greater than $\frac{1}{2}$?

Solution: Here $p = \frac{1}{20}, q = \frac{19}{20}$

Let n be the required number of trials such that the probability of getting at least one success, is just greater than $\frac{1}{2}$

i.e. $P\ (X \geq 1) > \frac{1}{2}$

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

$\Rightarrow\ P(X = 0) < 1 - \frac{1}{2}$

$\Rightarrow\ {}^{n}C_0 \left(\frac{19}{20}\right)^{n} \left(\frac{1}{20}\right)^{0} < \frac{1}{2}$

$\Rightarrow\ \left(\frac{19}{20}\right)^{n} < \frac{1}{2}$

$\Rightarrow n \log_{10} \frac{19}{20} < \log_{10} \frac{1}{2}$

$\Rightarrow n > \frac{\log_{10}\frac{1}{2}}{\log_{10}\frac{19}{20}} \qquad \because \log_{10} \frac{19}{20} < 0$

$\Rightarrow n > \frac{-0.3010}{-0.02228} = 13.5099 \quad \therefore n = 14$

Example 8 The probability of a man hitting a target is 1/3. How many times must he take the shot so that the probability of hitting the target at least once is less than 90%?

Solution: Here $p = \frac{1}{3}, q = \frac{2}{3}$

Let n be the number of shots so that the probability of hitting the target, at least once, is less than 90%

i.e. $P\ (X \geq 1) < \frac{9}{10}$

$\Rightarrow 1 - P(X = 0) < \frac{9}{10}$

$\Rightarrow\ P(X = 0) > 1 - \frac{9}{10}$

$\Rightarrow\ {}^{n}C_0 \left(\frac{2}{3}\right)^{n} \left(\frac{1}{3}\right)^{0} > \frac{1}{10}$

$\Rightarrow\ \left(\frac{2}{3}\right)^{n} > \frac{1}{10}$

$\Rightarrow n \log_{10}\frac{2}{3} > \log_{10}\frac{1}{10}$

$\Rightarrow n < \frac{\log_{10}\frac{1}{10}}{\log_{10}\frac{2}{3}} \qquad \because \log_{10}\frac{2}{3} < 0$

$\Rightarrow n < \frac{-1}{-0.1761} = 5.6786 \qquad \therefore n = 5$

Example 9 Assuming that half the population are consumers of chocolates, so that the chances of an individual being consumer is $\frac{1}{2}$ and assuming that each of the 25 surveyors takes 10 individuals to see whether they are consumers. How many surveyors would you expect to report that three or less people were consumers?

Solution: The probability of an individual to be consumer $(p) = \frac{1}{2}, \therefore q = \frac{1}{2}$

Also $n = 10$, $N = 25$

$\therefore$ B.D. is given by $25\left(\frac{1}{2}+\frac{1}{2}\right)^{10}$ and if X is random variable showing number of successes, then

$$P(X \le 3) = 25\left[{}^{10}C_0 \left(\frac{1}{2}\right)^{10} + {}^{10}C_1 \left(\frac{1}{2}\right)^{9}\left(\frac{1}{2}\right)^{1} + {}^{10}C_2 \left(\frac{1}{2}\right)^{8}\left(\frac{1}{2}\right)^{2} + {}^{10}C_3 \left(\frac{1}{2}\right)^{7}\left(\frac{1}{2}\right)^{3}\right]$$

$$= 25\left(\frac{1}{2}\right)^{10}\left[1 + 10 + \frac{10\times 9}{2!} + \frac{10\times 9\times 8}{3!}\right] = 25(0.171875) = 4.3$$

So we can expect four surveyors to report that three or less people were consumers.

Example10: Fit a binomial distribution to the following data and compare theoretical frequencies with actual ones

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

Solution: Mean of the given distribution $= \frac{\sum fx}{\sum f}$, $\sum f = 80$

$$= \frac{0+20+56+36+32+30+0}{80} = \frac{87}{40} = 2.175$$

Let mean of binomial distribution to be fitted $= np = 2.175$

Also $n = 10 \therefore p = 0.2175 \quad q = 1 - 0.2175 = 0.7825$

$\therefore$ B.D. is given by $80(0.7825 + 0.2175)^{10}$

Theoretical frequencies using binomial distribution are given in the table below:

x	P(r) = ${}^nC_r\, q^{n-r}p^r$	Theoretical frequencies (f) = 80× $P(r)$
0	${}^{10}C_0\,(0.7825)^{10}(0.2175)^0 = 0.086$	$6.9 = 7$(say)
1	${}^{10}C_1\,(0.7825)^9(0.2175)^1 = 0.239$	$19.1 = 19$(say)
2	${}^{10}C_2\,(0.7825)^8(0.2175)^2 = 0.299$	$23.9 = 24$(say)
3	${}^{10}C_3\,(0.7825)^7(0.2175)^3 = 0.22$	$17.8 = 18$(say)
4	${}^{10}C_4\,(0.7825)^6(0.2175)^4 = 0.11$	$8.6 = 9$ (say)
5	${}^{10}C_5\,(0.7825)^5(0.2175)^5 = 0.04$	$2.9 = 3$ (say)
6	${}^{10}C_6\,(0.7825)^4(0.2175)^6 = 0.008$	$0.66 = 0$ (say)
7	${}^{10}C_7\,(0.7825)^3(0.2175)^7 = 0.001$	$0.11 = 0$ (say)
8	${}^{10}C_8\,(0.7825)^2(0.2175)^8 = 0$	0
9	${}^{10}C_9\,(0.7825)^1(0.2175)^9 = 0$	0

3.3 Poisson Distribution

Result: Poisson distribution with $P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$ is a limiting case of Binomial distribution, under the conditions $\boldsymbol{i.}\ n \to \infty$ $\boldsymbol{ii.}\ p \to 0$ $\boldsymbol{iii.}\ np = \lambda$ is finite

Proof: In a Binomial distribution

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

$$= {}^nC_r\,(1-p)^{n-r}p^r$$

$$= {}^nC_r\left(1-\frac{\lambda}{n}\right)^{n-r}\left(\frac{\lambda}{n}\right)^r \qquad \because np = \lambda$$

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

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

Taking limit as $n \to \infty$

$$P(r) = \frac{\lambda^r}{r!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{\lambda^r}{r!} \lim_{n\to\infty} \left[\left(1 - \frac{\lambda}{n}\right)^{\frac{-n}{\lambda}}\right]^{-\lambda}$$

$\therefore P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$ $\quad\quad\quad \because \lim_{n\to\infty} \left(1 + \frac{1}{x}\right)^{x} = e$

3.3.1 Mean of Poisson Distribution

$$\text{Mean} = \sum_{r=0}^{\infty} rP(r)$$

$$= \sum_{r=0}^{\infty} \frac{re^{-\lambda}\lambda^r}{r!}$$

$$= \frac{e^{-\lambda}\lambda^1}{1!} + \frac{2e^{-\lambda}\lambda^2}{2!} + \frac{3e^{-\lambda}\lambda^3}{3!} + \frac{4e^{-\lambda}\lambda^4}{4!} + \cdots$$

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

$$= e^{-\lambda}\,\lambda\, e^{\lambda} = \lambda$$

3.3.2 Variance of Poisson Distribution

$$\text{Variance} = \sum_{r=0}^{n} r^2 P(r) \; - \; (\text{mean})^2$$

$$\text{Now } \sum_{r=0}^{n} r^2 P(r) \; = \sum_{r=0}^{\infty} \frac{r^2 e^{-\lambda}\lambda^r}{r!}$$

$$= \frac{e^{-\lambda}\lambda^1}{1!} + \frac{2^2 e^{-\lambda}\lambda^2}{2!} + \frac{3^2 e^{-\lambda}\lambda^3}{3!} + \frac{4^2 e^{-\lambda}\lambda^4}{4!} + \cdots$$

$$= e^{-\lambda}\,\lambda\left(1 + 2\lambda + \frac{3\lambda^2}{2!} + \frac{4\lambda^3}{3!} + \cdots\right)$$

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

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

$$= e^{-\lambda}\,\lambda\left[e^{\lambda} + \lambda e^{\lambda}\right]$$

$$\Rightarrow \sum_{r=0}^{n} r^2 P(r) = e^{-\lambda}\,\lambda e^{\lambda}[1 + \lambda] = \lambda + \lambda^2$$

$$\therefore \text{Variance} = \lambda + \lambda^2 - \lambda^2 = \lambda$$

3.3.3 Moments of Poisson Distribution

Since Poisson distribution is a limiting case of binomial distribution, therefore mean and moments may be obtained from Binomial distribution by taking $np = \lambda$, $p \to 0$ and $q \to 1$ as limit $n \to \infty$

Mean $= \lim_{n\to\infty} np = \lambda$,

$$\mu_1 = 0$$

$$\mu_2 = \sigma^2 = \lim_{n\to\infty} np\, q = \lambda \qquad \because \; \mu_2 = variance\ (\sigma^2)$$

Similarly $\mu_3 = \lim_{n\to\infty} np\, q(q-p) = \lambda.1(1-0) = \lambda$

$$\mu_4 = \lim_{n\to\infty} np\, q[1 + 3pq(n-2)]$$

$$= \lim_{n\to\infty} np\, q[1 + 3npq - 6pq]$$

$$= \lambda.1[1 + 3\lambda.1 - 6.0.1]$$

$$= \lambda + 3\lambda^2$$

$$\beta_1 = \frac{u_3^2}{u_2^3} = \frac{\lambda^{\ 2}}{\lambda^3} = \frac{1}{\lambda}$$

$$\beta_2 = \frac{\mu_4}{u_2^2} = \frac{\lambda + 3\lambda^2}{\lambda^2} = \frac{1}{\lambda} + 3$$

Example11: If the standard deviation of a Poisson variate X is $\sqrt{3}$, then find the probability that X is strictly positive.

Solution: Here variance (λ)= 3

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

The probability that X is strictly positive is:

$$P(X > 0) = 1 - \mathrm{P}(X = 0)$$

$$= 1 - \frac{e^{-3}3^0}{0!} = 1 - e^{-3}$$

Example12: $i.$ If the probability of a bad reaction from a certain injection is 0.001, determine the chance that out of 1000 individuals more than two will have bad reaction.

$ii.$ A manufacturer who produces medicine bottles finds that 0.1% of the bottles are defective. The bottles are packed in boxes containing 500 bottles. Find the probability that in 100 such boxes how many are expected to contain (a) no defective (b) at least two defective bottles.

Solution: $i.$ Here $p = 0.001$, $n = 1000$ $\lambda = np = 1000 \times 0.001 = 1$

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

Probability that more than two individuals will have bad reaction is given by:

$$P(X > 2) = 1 - P(X \leq 2) = \{1 - [P(X = 0) + P(X = 1) + P(X = 2)]\}$$

$$= \left\{1 - e^{-1}\left[\frac{(1)^0}{0!} + \frac{(1)^1}{1!} + \frac{(1)^2}{2!}\right]\right\} = 1 - e^{-1}\left[\frac{5}{2}\right] = 1 - 0.9197 = 0.0803$$

$ii.$ Here $p = 0.1\% = \frac{0.1}{100} = 0.001$, $n = 500$ $\lambda = np = 500 \times 0.001 = 0.5$

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

(a) Probability of zero defective bottles in a box of 500 bottles is given by:

$$P(X = 0) = \frac{e^{-0.5}(0.5)^0}{0!} = e^{-0.5} = 0.6065$$

$\therefore$ Number of boxes having no defective bottle out of 100 boxes

$= 100 \times 0.6065 = 60.65$ approx

(b) Probability of at least 2 defective bottles in a box of 500 is given by:

$$P(X \geq 2) = 1 - P(X < 2) = \{1 - [P(X = 0) + P(X = 1)]\}$$

$$= \left\{1 - e^{-0.5}\left[\frac{(0.5)^0}{0!} + \frac{(0.5)^1}{1!}\right]\right\} = 1 - e^{-0.5}[1.5] = 0.0902$$

$\therefore$ Number of boxes having at least 2 defective bottles out of 100 boxes

$= 100 \times 0.0902 = 9.02$ approx

Example13: In a certain factory producing tyres, there is a small chance of 1 in 500 tyres to be defective. The tyres are supplied in lots of 10. Using Poisson distribution, calculate the approximate number of lots containing $i.$ no defective

$ii.$ at least one defective tyre in a consignment of 10,000 lots.

Solution: Here $p = \frac{1}{500}$, $n = 10$ $\lambda = np = \frac{10}{500} = 0.02$

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!} = \frac{e^{-0.02}(0.02)^r}{r!}$$

$i.$ Probability of no defective tyre in a lot is given by:

$$P(X = 0) = \frac{e^{-0.02}(0.02)^0}{0!} = e^{-0.02} = 0.9802$$

∴ Number of lots containing no defective tyre = $10000 \times 0.9802 = 9802$ approx

$ii.$ Probability of at least one defective tyre in a lot is given by:

$$1 - P(X = 0) = 1 - 0.9802 = 0.0198$$

∴ Number of lots containing at least one defective tyre

$= 10000 \times 0.0198 = 198$ Approx

Example14: A skilled typist kept a record of his mistakes made per day during 300 working days. Fit a Poisson distribution to compare theoretical frequencies with actual ones

Mistakes per day	0	1	2	3	4	5	6
Number of days	143	90	42	12	9	3	1

Solution: Mean of the given distribution= $\frac{\sum fx}{\sum f}$, $\sum f = 300$

$$= \frac{0+90+84+36+36+15+6}{300} = \frac{89}{100} = 0.89 = \lambda$$

Mistakes per day	$P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$	Theoretical frequency $300\times P(r)$
0	$\frac{e^{-(0.89)}(0.89)^0}{0!} = 0.411$	123.3=123 (say)
1	$\frac{e^{-(0.89)}(0.89)^1}{1!} = 0.365$	109.5=110 (say)
2	$\frac{e^{-(0.89)}(0.89)^2}{2!} = 0.163$	48.9=49 (say)
3	$\frac{e^{-(0.89)}(0.89)^3}{3!} = 0.048$	14.4=14 (say)
4	$\frac{e^{-(0.89)}(0.89)^4}{4!} = 0.011$	3.3=3 (say)

5	$\frac{e^{-(0.89)}(0.89)^5}{5!} = 0.002$	0.6=1 (say)
6	$\frac{e^{-(0.89)}(0.89)^6}{6!} = 0.0003$	0.09=0 (say)

Example15: The distribution of number of road accidents per day in a city is Poisson with mean 5. Find the number of days out a year when there will be

$i.$ at most 2 accidents $ii.$ between 3 and 5 accidents

Solution: Here $\lambda = np = 5$

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!} = \frac{e^{-5}(5)^r}{r!}$$

$i.$ Probability of at most 2 accidents per day is given by:

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

$$= e^{-5}\left[\frac{(5)^0}{0!} + \frac{(5)^1}{1!} + \frac{(5)^2}{2!}\right] = 0.1247$$

∴ Number of days in a year having at most 2 accidents per day

$$= 365 \times 0.1247 = 45.5 \text{ approx}$$

$ii.$ Probability of 3 to 5 accidents per day is given by:

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

$$= e^{-5}\left[\frac{(5)^3}{3!} + \frac{(5)^4}{4!} + \frac{(5)^5}{5!}\right] = 0.4913$$

∴ Number of days in a year having 3 to 5 accidents per day

$$= 365 \times 0.4913 = 179.3 \text{ approx}$$

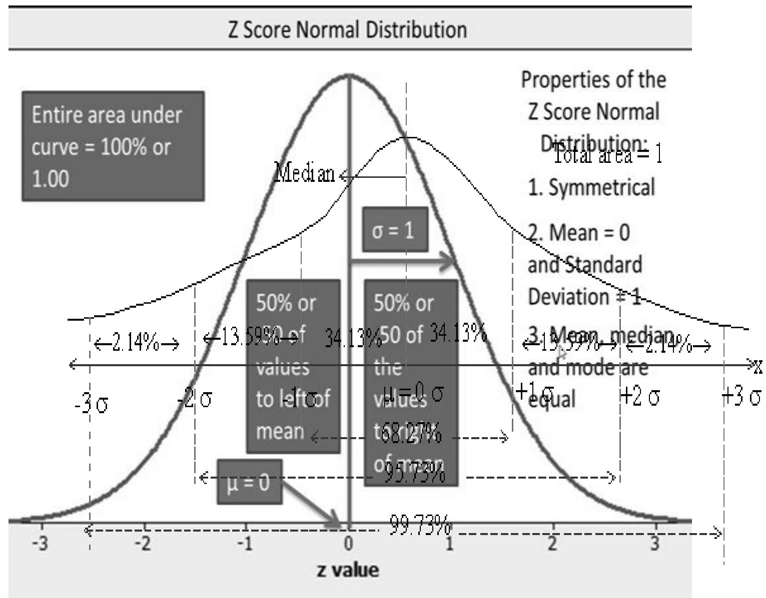
3.4 Normal Distribution

The normal distribution developed by Gauss is a continuous distribution and is very useful in practical applications. It can be considered as the limiting form of the Binomial Distribution when the number of trials (n), is very large and neither p nor q is very small. The probability curve of a normal variate x with mean μ and standard deviation σ is given by:

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

Any normal variate $\boldsymbol{x}$ with mean $\boldsymbol{\mu}$ and standard deviation $\boldsymbol{\sigma}$ is changed to a standard normal variate $z = \frac{x-\mu}{\sigma}$, and hence the probability density function of z is given by: $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, -\infty < z < \infty$

The normal distribution with mean μ and variance σ^2 is denoted by $N(\mu\,, \sigma^2)$. Adjoining figure shows a normal distribution curve for standard normal variate z.



Properties of Normal Curve:

- The graph of $p(x)$ or $\phi(z)$ is a bell shaped curve.
- Since the distribution is symmetrical, mean, mode and median coincide at $x = \mu$ or $z = 0$. Also $\beta_1 = 0 \Rightarrow \gamma_1 = 0$ and $\beta_2 = 3 \Rightarrow \gamma_2 = 0$
- The ordinate at $x = \mu$ or $z = 0$, divides the whole area into two equal parts. Also since the total area under the probability curve is 1, area to the right of the ordinate as well as to the left of the ordinate at $x = \mu$ or $z = 0$ is 0.5.
- Area under the curve between the ordinates $a < z < b$ gives the probability of variate z taking values between a and b. Area is concentrated more towards the middle and goes on decreasing on the either sides of the curve, i.e. tails, but never becomes zero. The curve never intersects x-axis at any finite point. i.e. x-axis is its asymptote.
- Since the curve is symmetrical about mean. The first quartile Q_1 and the third quartile Q_3 lie at the same distance on the two sides of the mean μ. The distance of any quartile from μ is $0.6745\,\sigma$ units.

  Thus $Q_1 = \mu - 0.6745\sigma$ or $\mu - \frac{2}{3}\sigma$

  $Q_3 = \mu + 0.6745\sigma$ or $\mu + \frac{2}{3}\sigma$
- Area under the normal curve is distributed as follows:

  Area between $x = \mu - \sigma$ and $x = \mu + \sigma$ is 18.27%

  Area between $x = \mu - 2\sigma$ and $x = \mu + 2\sigma$ is 95.45%

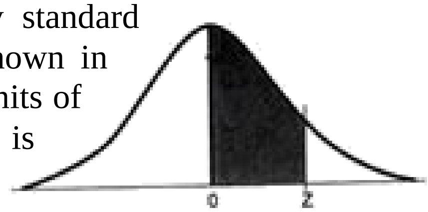
  Area between $x = \mu - 3\sigma$ and $x = \mu + 3\sigma$ is 99.73%
- Points of inflection are $\mu \pm \sigma$

- Quartile deviation (QD) is
  $\frac{1}{2}\left(Q_3 - Q_1\right) \frac{2}{3}\sigma$
- Mean deviation (MD) is $\frac{4}{5}\sigma$
  $\therefore QD\!:MD\!:SD \equiv 10\!:12\!:15$
- Since the distribution is symmetrical, all the moments of odd order about mean are zero, i.e. $\mu_{2n+1} = 0$, n =1, 2, 3, $\cdots$
- Moments of even order are given by:
  $\mu_{2n} = 1.3.5 ... (2n-1)\sigma^{2n}$, n =1, 2, 3, $\cdots$

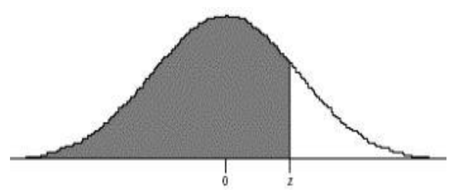
Putting n =1 and 2; $\mu_2 = \sigma^2$, $\mu_4 = 3\sigma^4$, $\beta_1 = \frac{u_3^2}{u_2^3} = 0$, $\beta_2 = \frac{\mu_4}{u_2^2} = \frac{3\sigma^4}{\sigma^4} = 3$

Using Normal Distribution tables: We first convert the variate x into standard normal variate z; using the relation $z = \frac{x-\mu}{\sigma}$ and find the limits of z -score corresponding to the given limits of the variate x. Normal Distribution tables are broadly of two types; either of the two may be used to determine the area (probability) for the standard normal variate z in the range $(-3.49 \leq z \leq 3.49)$.

The first type of table gives the area covered by standard normal variate z between the ordinates 0 to z as shown in adjoining figure. This table covers more than 0.499 units of area on positive side of the curve. If value of z is negative, we may use the symmetrical property of the normal curve, i.e area in the region $z < -a$ is same as area $z > a$ where a is any positive number within the range $(0, 3.49)$.



The second table gives the area covered by standard normal variate z from $-\infty$ to z as shown in the given figure. This table covers more than 0.998 units of area on the whole.



3.4.1 Moments of Normal Distribution

Result: Odd order moments about mean are zero

Proof: r^{th} moment about the mean $\bar{x}$ of a continuous distribution is given by:

$$\mu_r = \int_{-\infty}^{\infty} \frac{f(x-\bar{x})^r}{n}\, dx$$

$\therefore$ Odd order moments of normal distribution with mean μ are given by:

$$\mu_{2n+1} = \int_{-\infty}^{\infty} (x-\mu)^{2n+1} p(x) dx \quad \because \frac{f}{n} = p(x)$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n+1} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad \because p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{\sigma^{2n+1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n+1} e^{-\frac{1}{2}z^2} dz \text{ Putting } \frac{x-\mu}{\sigma} = z\,,\ dx = \sigma\, dz$$

$$= 0, \quad z^{2n+1} e^{-\frac{1}{2}z^2} \text{being an odd function of } z.$$

Result: Even order moments about mean are given by:

$\mu_{2n} = 1.3.5 \ldots (2n-1)\sigma^{2n}$, n =1, 2, 3, $\cdots$

Proof: $\mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n} p(x) dx$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \qquad \because p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{\sigma^{2n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^{2n} e^{-\frac{1}{2}z^2} dz \quad \text{Putting } \frac{x-\mu}{\sigma} = z \Rightarrow dx = \sigma\, dz$$

$$= \frac{2\sigma^{2n}}{\sqrt{2\pi}} \int_{0}^{\infty} z^{2n} e^{-\frac{1}{2}z^2} dz\,, \quad z^{2n} e^{-\frac{1}{2}z^2} \text{being even function of } z.$$

$$= \frac{2\sigma^{2n}}{\sqrt{2\pi}} \int_{0}^{\infty} (2t)^n e^{-t} (2t)^{-\frac{1}{2}} dt \text{ Putting } \frac{z^2}{2} = t \Rightarrow z dz = dt$$

$$= \frac{2^n\sigma^{2n}}{\sqrt{\pi}} \int_{0}^{\infty} e^{-t} t^{n-\frac{1}{2}} dt$$

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

Again changing n to $(n-1)$

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

$$\therefore \frac{\mu_{2n}}{\mu_{2n-2}} = 2\sigma^2 \frac{\Gamma\left(n+\frac{1}{2}\right)}{\Gamma\left(n-\frac{1}{2}\right)} = 2\sigma^2 \frac{\left(n-\frac{1}{2}\right)\Gamma\left(n-\frac{1}{2}\right)}{\Gamma\left(n-\frac{1}{2}\right)} = 2\sigma^2 \left(n-\frac{1}{2}\right)$$

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

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

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

$$\vdots$$

$\therefore\ \mu_{2n} = 1.3.5 \ldots (2n-5)(2n-3)(2n-1)\sigma^{2n}$, n =1, 2, 3, $\cdots$

Example16 For a normal distribution the mean 20 and the standard deviation 15, find $i.$ Q_1 and Q_3 $ii.$ Mean deviation $iii.$ the inter quartile range.

Solution: $i.$ For a normal distribution

$$Q_1 = \mu - \tfrac{2}{3}\sigma = 20 - \tfrac{2}{3}(15) = 10$$

$$Q_3 = \mu + \tfrac{2}{3}\sigma = 20 + \tfrac{2}{3}(15) = 30$$

$ii.$ Mean deviation is $\frac{4}{5}\sigma = \frac{4}{5}(15) = 12$

$iii.$ The inter quartile range is $Q_3 - Q_1 = 20$

Example 17 Find p, mean and the standard deviation of the normal distribution given by $y = pe^{-\left(\frac{x^2}{8} - x + 2\right)}$

Solution: Rewriting $\left(\frac{x^2}{8} - x + 2\right)$ as $\frac{1}{2}\left(\frac{x^2-8x+16}{4}\right) = \frac{1}{2}\left(\frac{x-4}{2}\right)^2$

$$\therefore y = pe^{-\frac{1}{2}\left(\frac{x-4}{2}\right)^2}$$

Comparing with $y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, we get $\mu = 4, = 2$, $p = \frac{1}{\sigma\sqrt{2\pi}} = \frac{1}{2\sqrt{2\pi}}$

Example18 If X is a random variable with mean 20 and standard deviation 5, find the probabilities that (i) $15 \le X \le 25$ (ii) $|X - 20| > 5$

Solution: Given that X is a random variable with mean $\mu = 20$ and S.D. $\sigma = 5$.

i.e. $X \sim N(100\,,\ 25)$ and $z = \frac{X-\mu}{\sigma} = \frac{X-20}{5} \ \Rightarrow X = 20 + 5z$

(i) $P(15 \le X \le 25\,)$

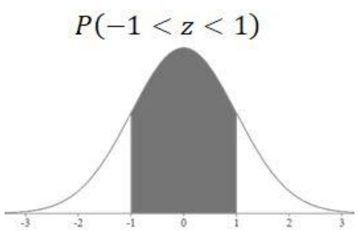
$= P(15 < 20 + 5z < 25)$

$= P\left(\frac{15-20}{5} < z < \frac{25-20}{5}\right)$

$= P(-1 < z < 1) = 2P(0 < z < 1)$

$= 2(0.3413)$ Using normal table

$= 0.6826$



(ii) $P(|X - 20| > 5) = 1 - P(|X - 20| \le 5)$

Now $P(|X - 20| \le 5) = P(-5 \le X - 20 \le 5) = P(15 \le X \le 25)$

$= \ 0.6826$ using part (i)

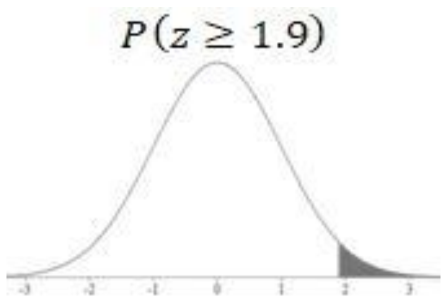
$\therefore P(|X - 20| > 5) = 1 - 0.6826 = 0.3174$

Example19 Fifteen hundred candidates appeared in a certain examination, having maximum marks as 100. It was found that the marks are normally distributed with mean 55 and standard deviation as 10.5. Determine approximately the number of candidates who were passed with distinction, i.e. 75% and above marks.

Solution: Let the random variable X denote the marks obtained out of 100.

Then X is a random variable with mean $\mu = 55$ and S.D. $\sigma = 10.5$.

i.e. $X \sim N(55\ ,\ 10.5^2)$ and $z = \frac{X-\mu}{\sigma} = \frac{X-55}{10.5}$



$\Rightarrow X \ = 55 + 10.5z$

$P(X \geq 75)$

$= P(55 + 10.5z \geq 75)$

$= P\left(z \geq \frac{75-55}{10.5}\right) = P(z \geq 1.9)$

$= 0.5 - P(0 < z < 1.9)$

$= \ 0.5 - 0.4713 = 0.0287$

$\therefore$Number of candidates who passing with distinction is:

$1500 \times 0.0287 = 43.05$ i.e. 43 approximately

Example 20 The daily wages of 1000 workers are normally distributed with mean 100$ and with a standard deviation 5$. Estimate the number of workers whose daily wages will be: (i) between 100$ and 105$ (ii) between 96$ and 105$

(iii) more than 110$ (iv) more than 110$

(v) Also estimate the daily wages of 100 highest paid workers.

Solution: Let the random variable X denote the daily wages in dollars.

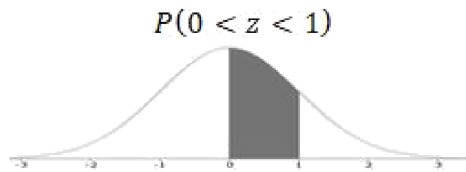
Then X is a random variable with mean $\mu = 100$ and S.D. $\sigma = 5$.

i.e. $X \sim N(100\ ,\ 25)$ and $z = \frac{X-\mu}{\sigma} = \frac{X-100}{5}$

$\Rightarrow X = 100 + 5z$

(i) $P(100 < X < 105)$

$= P(100 < 100 + 5z < 105)$

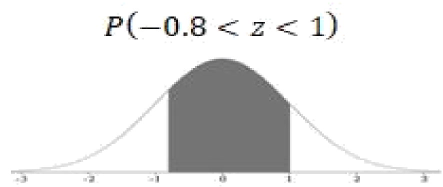


$= P\left(\frac{100-100}{5} < z < \frac{105-100}{5}\right)$

$= P(0 < z < 1) = 0.3413$

using Z table

(ii) $P(96 < X < 105)$



$= P(96 < 100 + 5z < 105)$

$= P\left(\frac{96-100}{5} < z < \frac{105-100}{5}\right)$

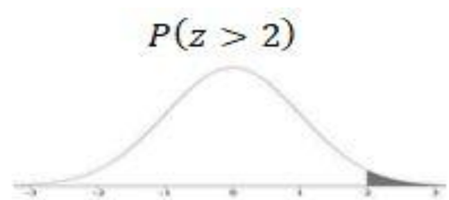
$= P(-0.8 < z < 1)$

$= P(0 < z < 0.8) + P(0 < z < 1)$

$=$ $0.2881 + 0.3413 = 0.6294$ using Z table

(iii) $P(X > 110)$

$= P(100 + 5z > 110)$



$= P\left(z > \frac{110-100}{5}\right)$

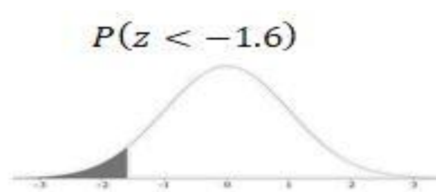
$= P(z > 2)$

$= 0.5 - P(0 < z < 2)$

$=$ $0.5 - 0.4772 = 0.0228$ using Z table

(iv) $P(X < 92)$

$= P(100 + 5z < 92)$



$= P\left(z < \frac{92-100}{5}\right)$

$= P(z < -1.6)$

$= P(z > 1.6)$ (Symmetry)

$= 0.5 - P(0 < z < 1.6)$

$= \ 0.5 - 0.4452 = 0.0548$ using Z table

(v) Proportion of 100 highest paid workers is $\frac{100}{1000} = \frac{1}{10} = 0.1$

To determine $X = r$ such that $P(X > r) = 0.1$

When $X = r, z = \frac{r-100}{5} = z_1$ (say)

$\therefore P(z > z_1) = 0.1$

$\Rightarrow P(0 < z < z_1) = 0.5 - 0.1 = 0.4$

From normal distribution table, $z_1 = 1.28$ approx.

$\therefore z = \frac{r-100}{5} = 1.28$

$\Rightarrow r = 100 + 5 \times 1.28 = 106.4$

Hence the lowest daily wages of 100 highest paid workers are 106.4$

Example 21 If in a normal distribution 7% of the items are under 35 and 89% are under 63. What are the mean and standard deviations of the distribution?

Solution: Let X be normally distributed with mean μ and standard deviation σ such that $z = \frac{X-\mu}{\sigma}$. Also given that $P(X < 35) = 0.07$ and $P(X < 63) = 0.89$

Let $z_1$be a standard normal variate corresponding to $X = 35$,

$\therefore \ z_1 = \frac{35-\mu}{\sigma}$ and $P(\ z_1 < 35) = 0.07$

$\Rightarrow \ \ P(-z_1 > 35) = 0.07$

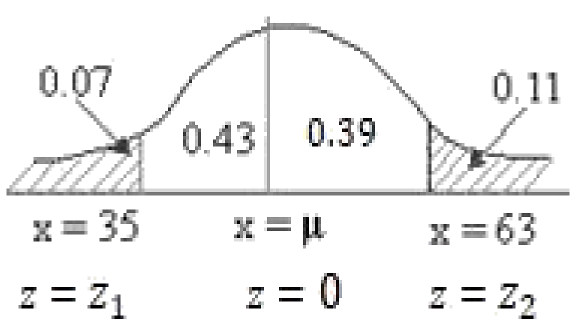
$\Rightarrow P(0 < -\ z_1 < z) = 0.5 - 0.07 = 0.43$

$\Rightarrow -z_1 = 1.48$ From normal distribution table

or $z_1 = \frac{35-\mu}{\sigma} = -1.48$

$\Rightarrow \mu - 1.48\sigma = 35$ $\cdots$ ①

Again let $z_2$be a standard normal variate corresponding to $X = 63$



$\therefore\ z_2 = \frac{63-\mu}{\sigma}$ and $P(z_2 < 63) = 0.89$

$\Rightarrow P(0 < z_2 < z) = 0.89 - 0.5 = 0.39$

$\Rightarrow z_2 = \frac{63-\mu}{\sigma} = 1.23$ from normal distribution table

$\Rightarrow \mu + 1.23\sigma = 63 \ \cdots$ ②

Solving ① and ②, we get: $\mu = 109.6, \sigma = 50.4$

Example21: Fit a normal curve to the following distribution

x	2	4	6	8	10
f	1	4	6	4	1

Solution: Let x be normally distributed with mean μ and standard deviation σ such that $z = \frac{X-\mu}{\sigma}$.

$$\mu = \frac{\sum fx}{\sum f} = \frac{2+16+36+32+10}{16} = 6$$

Also $$\sigma = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2} = \sqrt{\frac{(4+64+216+256+100)}{16} - 36} = \sqrt{40-36} = 2$$

Again x has to be a continuous variable to follow normal distribution, therefore taking x as mid value of an interval (x_1, x_2)

x	(x_1, x_2)	(z_1, z_2)	$P(z) \equiv$Area under the curve in the interval (z_1, z_2)	Theoretical frequency $16 \times P(z)$
2	(1, 3)	(-2.5, -1.5)	0.0606	0.9696
4	(3, 5)	(-1.5, -0.5)	0.2417	3.8672
6	(5, 7)	(-0.5, 0.5)	0.3829	6.1264
8	(7, 9)	(0.5, 1.5)	0.2417	3.8672
10	(9, 11)	(1.5, 2.5)	0.0606	0.9696

Exercise 3

1. In 256 sets of 12 tosses of a coin, in how many cases can one expect 8 heads and 4 tails?
2. In a precision bombing attack, there are 50% chances that any one bomb will hit the target. If two direct hits are required to destroy the target completely, how many bombs must be dropped to provide a 99% or more chances to completely destroy the target.
3. Comment on the statement: For a binomial distribution mean is 5 and standard deviation is 3.

4. If the probability of a defective item is 0.02, find the probability that at most 5 defective items will be found out in a box of 200 items.
5. Six coins are tossed 6400 times. Using Poisson distribution, find the approximate probability of getting six heads 2 times.
6. In a certain factory making a machine part, probability of it being defective is 0.002. If the part is supplied in packs of 10, use Poisson distribution to calculate the approximate number of packets containing no defective and one defective machine part in a consignment of 10,000 packets.
7. A car hire firm has 2 cars, which are hired on daily basis. The number of demands for a car on each day follows 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, given that $e^{-1.5} = 0.2231$.
8. If in a Poisson distribution $P(X = 0) = P(X = 1) = a$, show that $a = \frac{1}{e}$.
9. If X is a random variable with mean 30 and standard deviation 5, find the probabilities that $(i)\ X > 45\ (ii)\ |X - 30| \geq 5$
10. If in a normal distribution 31% of the items are under 45 and 8% are over 64. What are the mean and standard deviation of the distribution?
11. In an intelligence test given to 1000 children, the average score is 42 with the standard deviation 24. Find $i.$ the number of children whose score exceeds 60

    $ii.$ the number of children whose score lie between 20 and 40.
12. Assuming mean height of the soldiers to be 68.22 inches with a variance of 10.8 square inches, how many soldiers in the regiment of 1000would you expect to be over 6 feet tall, given that area under the standard normal curve between $z = 0$ to $z = 1.15$ is 0.3746.
13. Fit a binomial distribution for the given data:

x	0	1	2	3	4	5	6	7	8
f	0	5	9	22	25	26	14	4	1

Answers

1. 31
2. 11
3. Variance cannot be greater than mean in case of a binomial distribution.
4. 0.7845
5. 5000 e^{-100}
6. 9802, 196
7. 0.2231, 0.1913

9. 0.0014 , 0.3174
10. 50, 10
11. 227, 287
12. 125
13. 0.4, 3.3, 11.6, 23.2, 29, 23.2, 11.6, 3.3, 0.4