Chapter –3

Elementary Concepts of Statistics

By

Dr V.D.Kulkarni

HUTATMA RAJGURU MAHAVIDYALAYA

RAJGURUNAGAR (PUNE)

Probability is defined as the ratio of number of cases in which events occurs to the total no. of cases.

A coin tossed , pro. That head on upper side =1/2

<u>Probability of independent events:-</u>

Consider two events occurs simultanously or succession,

One event occur in n_1 and other event occur in n_2

Total events= $n=n_1n_2$

Let m1 no. of ways in first event. And similarly m_2

Total favorable events= m_1m_2

Probability of composite p= $m/n = m_1m_2/n_1n_2$

$P= m_1/n_1 * m_2/n_2 = p_1*p_2$

The probability of occurrence of two events simultaneously is equal to product of probabilities of individual independent Events is called as multiplicative law of probability.

<u>Additive law of Probability :-</u> Two or more are said to be mutually exclusive if the occurrence of any one of them prevents occurrence of others.

Consider two events occurs exclusively, m1 be no. of ways Favorable events of first event and similarly m2. and n be total no. of ways in first and second events.

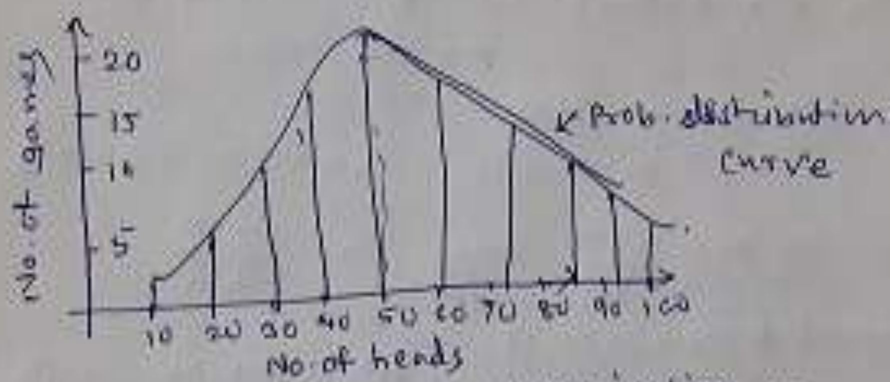
The events occur in first and second events =m= m1+m2

Total prob. =p =m/n= m1+m2/n=m1/n+m2/n =p1+p2

Probability Distribution :-

1

A game of tossing a coin 100 times. The prob. of getting head on upperside may be 50%.



Random walk Problem & Binomial distribution :-

A drunkard standing at lamppost on street

Let l be length of each step of drunkard

p be prob of drunkard moves right side by each step

q —— " —— left side by each step.

n_1 be total no. of steps on right side

n_2 be total no. of step on left side

Total prob $= p+q=1 \Rightarrow q = 1-p$

Total prob. on right side $= p \cdot p \cdot p \cdots n_1$ times $= p^{n_1}$

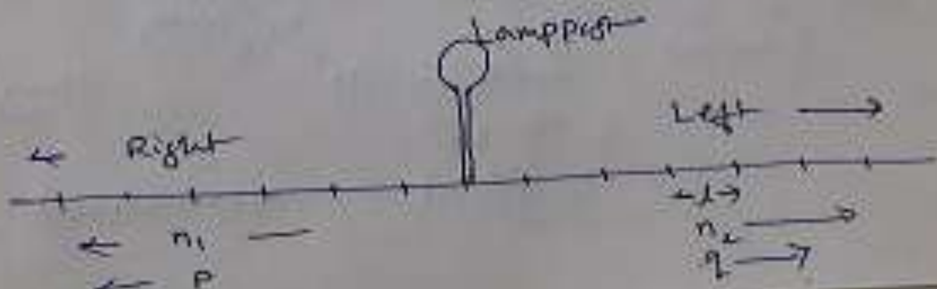
Total prob. on left side $= q \cdot q \cdots n_2$ times $= q^{n_2}$

& Total no. of steps $= N = n_1+n_2$

No. of distinct possibilities $= \dfrac{N!}{n_1!\, n_2!}$

The prob. $W_N(n_1)$ considering total no. of steps (N) is

$$W_N(n_1) = \frac{N!}{n_1!\, n_2!} p^{n_1} q^{n_2} \quad \text{--- Binomial distribution}$$



Fig(1)

$$P_N(m) = \frac{N!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!} p^{\left(\frac{N+m}{2}\right)} q^{\left(\frac{N-m}{2}\right)}$$

If $p = q = 1/2$

$$\boxed{P_N(m) = \frac{N!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!}\left(\frac{1}{2}\right)^N}$$ Prob.

Calculation of mean value of mean square deviation for Random walk problem:-

(a) Normalization:- We know that prob.

$$W_N(n_1) = \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1}$$

Take sum on both sides

$$\Sigma W_N(n_1) = \sum \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1}$$ ... using Binomial thm

$$= (p+q)^N$$

$$= 1^N = 1$$ ... Normalised

② Mean value of $n_1 = \overline{n_1}$:- The mean value of n_1 is

$$\overline{n_1} = \sum W(n_1)\, n_1$$

$$= \sum \frac{N!}{n_1!\,(N-n_1)!} p^{n_1} q^{N-n_1}\, n_1 \quad \text{—} ①$$

Since $\underline{n_1 p^{n_1} = p \frac{\partial}{\partial p}(p^{n_1})} = p\, n_1 p^{n_1 - 1}$

put in ①

$$\overline{n_1} = \sum \frac{N!}{n_1!\,(N-n_1)!}\, p\frac{\partial}{\partial p}(p^{n_1})\, q^{N-n_1}$$

$$= p\frac{\partial}{\partial p}\left[\sum \frac{N!}{n_1!\,(N-n_1)!} p^{n_1} q^{N-n_1}\right]$$

$$= p\frac{\partial}{\partial p}(p+q)^N$$

$$= Np\,(p+q)^{N-1}$$

$$\boxed{\overline{n_1} = Np}$$

Ily $\Rightarrow \boxed{\overline{n_2} = Nq}$

$$\overline{n_1} + \overline{n_2} = Np + Nq = N(p+q) = N$$

mean displacement is

$$\overline{m} = \overline{n_1 - n_2} = \overline{n_1} - \overline{n_2} = Np - Nq = N(p-q)$$

$$\text{or } p = q \Rightarrow \overline{m} = 0$$

(3) Mean square deviation or dispersion:-

The deviation in n_1 is

$$\Delta n_1 = n_1 - \overline{n_1}$$

$$(\Delta n_1)^2 = (n_1 - \overline{n_1})^2 = n_1^2 - 2n_1\overline{n_1} + (\overline{n_1})^2$$

The mean square deviation is

$$\overline{(\Delta n_1)^2} = \overline{n_1^2 - 2n_1\overline{n_1} + (\overline{n_1})^2} = \overline{n_1^2} - 2\overline{n_1}\,\overline{n_1} + (\overline{n_1})^2$$

$$= \overline{n_1^2} - (\overline{n_1})^2$$

We have to calculate mean square deviation

$$\overline{n_1^2} = \sum W(n_1)\, n_1^2$$

$$= \sum \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1^2 \quad \text{—(2)}$$

Since

$$n_1^2 p^{n_1} = \left(p \frac{\partial}{\partial p}\right) p^{n_1} = p \frac{\partial}{\partial p} n_1 p^{n_1-1} p$$

$$= p n_1 \frac{\partial}{\partial p} p^{n_1}$$

$$= n_1 p\, n_1 p^{n_1-1} = n_1^2 p^{n_1}$$

Eq^n (2) becomes

$$\overline{n_1^2} = \sum \frac{N!}{n_1!(N-n_1)!} \left(p \frac{\partial}{\partial p}\right)^2 p^{n_1} q^{N-n_1}$$

$$\overline{n_1^2} = \left(p\frac{\partial}{\partial p}\right)^2 \left[\sum \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1}\right]$$

$$= \left(p\frac{\partial}{\partial p}\right)^2 (p+q)^N$$

$$= \left(p\frac{\partial}{\partial p}\right)\left(p\frac{\partial}{\partial p}\right) (p+q)^N$$

$$= \left(p\frac{\partial}{\partial p}\right) Np(p+q)^{N-1}$$

$$= p\left[N(p+q)^{N-1} + Np(N-1)(p+q)^{N-2}\right]$$

$$= p\left[N + N^2p - Np\right] = Np\left[1 + Np - p\right]$$

$$= Np\left[Np + q\right] = N^2p^2 + Npq$$

$$\overline{n_1^2} - (Np)^2 = Npq \qquad \because \overline{n_1} = Np$$

$$\overline{n_1^2} - (\overline{n_1})^2 = Npq$$

$$\boxed{\overline{(\Delta n_1)^2} = Npq}$$

Dispersion in terms of displacement

$$\Delta m = m - \overline{m}$$

$$= (2n_1 - N) - (2\overline{n_1} - N)$$

$$= 2(n_1 - \overline{n_1}) = 2\Delta n_1$$

$$\overline{(\Delta m)^2} = 4\overline{(\Delta n_1)^2}$$

$$\boxed{\overline{(\Delta m)^2} = 4Npq}$$

If $p = q = 1/2$

$$\boxed{\overline{(\Delta m)^2} = N}$$

(8)

Probability Distribution for large scale N :-

Prob. distribution $\Rightarrow$ $W(n_1) = \frac{N!}{n_1!\,(N-n_1)!} p^{n_1} q^{N-n_1}$ --- (1)

The condition of maxima,

I Derivative = Zero & II Derivate = −ve

We have to find behavior of $W(n_1)$ near maximum

$$n_1 = \bar{n}_1 + \eta$$

According to Taylor's series expansion

$$\ln W(n_1) = \ln W(\bar{n}_1 + \eta)$$

$$= \ln W(\bar{n}_1) + B_1\eta + \tfrac{1}{2} B_2 \eta^2 + \cdots \quad (2)$$

$$B_1 = \frac{d \ln W(\bar{n}_1)}{dn_1} \quad \& \quad B_2 = \frac{d^2 \ln W(\bar{n}_1)}{dn_1^2}$$

According to condition of maxima,

$$B_1 = 0 \quad \& \quad B_2 = -|B_2|$$

Eq^n (2)

$$\ln W(n_1) = \ln W(\bar{n}_1) - \tfrac{1}{2}|B_2|\eta^2$$

$$\Rightarrow \ln W(n_1) - \ln W(\bar{n}_1) = -\tfrac{1}{2}|B_2|\eta^2$$

$$\boxed{W(n_1) = W(\bar{n}_1)\, e^{-\frac{1}{2}|B_2|\eta^2}} \quad (3)$$

Taking ln of eqn (1) (9)

$$\ln W(n_1) = \ln\left(\frac{N!}{n_1!\,(N-n_1)!}\, p^{n_1} q^{N-n_1}\right)$$

$$= \ln N! - \ln n_1! - \ln(N-n_1)! + n_1 \ln p + (N-n_1)\ln q.$$

$$\frac{d}{dn}\ln n! = \ln n.$$

$$\frac{d}{dn_1}\ln W(n_1) = 0 - \ln n_1 + \ln(N-n_1) + \ln p - \ln q = 0 \quad (4)$$

At $n_1 = \bar{n}_1$

$$\ln\left[\frac{(N-\bar{n}_1)p}{\bar{n}_1 q}\right] = 0 \Rightarrow \frac{(N-\bar{n}_1)p}{\bar{n}_1 q} = 1 \Rightarrow (N-\bar{n}_1)p = \bar{n}_1 q$$

$$\Rightarrow \bar{n}_1(p+q) = Np$$

$$\Rightarrow \boxed{\bar{n}_1 = Np}$$

II Derivative = −ve $\qquad \because \bar{n}_1 = Np$

$$B_2 = \frac{d^2 \ln W(n_1)}{dn_1^2} = -\frac{1}{n_1} - \frac{1}{N-n_1}$$

$$= -\frac{1}{Np} - \frac{1}{N-Np} = -\frac{1}{N}\left(\frac{1}{p} + \frac{1}{1-p}\right)$$

$$= -\frac{1}{N}\left(\frac{1}{p} + \frac{1}{q}\right) = -\frac{1}{N}\left(\frac{q+1}{pq}\right) = -\frac{1}{Npq}$$

Eqⁿ 3 Using condition of normalisation

$$w(n_1) = w(\bar{n}_1) \int_{-\infty}^{\infty} e^{-\frac{1}{2}|B_2|\eta^2} \, d\eta = 1$$

Using the integration formula

$$\int_{-\infty}^{\infty} x^n e^{-\alpha x^2} \, dx = \Gamma\frac{n+1}{2} \, \alpha^{-\left(\frac{n+1}{2}\right)}$$

$$w(\bar{n}_1) \sqrt{\frac{2\pi}{|B_2|}} = 1 \Rightarrow w(\bar{n}_1) = \sqrt{\frac{|B_2|}{2\pi}}$$

Eqⁿ 3

$$w(n_1) = \frac{1}{\sqrt{2\pi Npq}} \, e^{-\frac{1}{2}\frac{1}{Npq}\eta^2} \quad \because \eta = n_1 - \bar{n}_1$$

$$\boxed{w(n_1) = \frac{1}{\sqrt{2\pi Npq}} \, e^{-\frac{1}{2Npq}(n_1 - Np)^2}} \rightarrow \text{Gaussian approx ec}$$

Gaussian Prob. Distribution:-

Prob. $P(m)$ for large value of N.

$$n_1 = \frac{N+m}{2}$$

$$P(m) = W\left(\frac{N+m}{2}\right) = \frac{1}{\sqrt{2\pi Npq}} \exp\left[-\frac{(n_1 - Np)^2}{2Npq}\right]$$

$$P(m) = \frac{1}{\sqrt{2\pi Npq}} \exp\left[-\frac{[m - N(p-q)]^2}{8Npq}\right]$$

The actual displ. $x = ml$

$\Rightarrow \Delta x = \Delta m l \Rightarrow \Delta x = 2l$

$$P(x)\,dx = P(m)\frac{dx}{2l}$$

$$P(x)\,dx = \frac{1}{\sqrt{2\pi Npq}} \exp\left[-\frac{[ml - N(p-q)l]^2}{8Npq\,l^2}\right]\frac{dx}{2l}$$

$x = ml$, $\mu = N(p-q)l$ & $\sigma = 2\sqrt{Npq}\cdot l$

$$\boxed{P(x)\,dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx}$$ Gaussian Prob. distribut[ion]

Calculation mean value & mean square deviation (12)

(1) Normalisation :-

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

Put $x-\mu = y \Rightarrow dx = dy$

$$\int_{-\infty}^{\infty} p(x)\,dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty} e^{-y^2/2\sigma^2}\,dy \quad -(1)$$

$$\int_{-\infty}^{\infty} x^n e^{-\alpha x^2}\,dx = \Gamma\left(\frac{n+1}{2}\right)\alpha^{-\left(\frac{n+1}{2}\right)}$$

$$n=0 \ \& \ \alpha = \frac{1}{2\sigma^2} \Rightarrow \Gamma\left(\tfrac{1}{2}\right)\left(\frac{1}{2\sigma^2}\right)^{-1/2} = \sqrt{\pi}\,(2\sigma^2)^{1/2} = \sqrt{2\pi}\cdot\sigma$$

Eqn (1)

$$\int_{-\infty}^{\infty} p(x)\,dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\,\sqrt{2\pi}\cdot\sigma = 1 \quad \text{Normalised}$$

② Mean value of x :-

$$\bar{x} = \int_{-\infty}^{\infty} x\,(P(x))\,dx$$

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

Put $x-\mu = y \Rightarrow x = \mu + y \Rightarrow dx = dy$

$$\bar{x} = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty}(y+\mu)\, e^{-y^2/2\sigma^2}\,dx$$

$$\bar{x} = \frac{1}{\sqrt{2\pi}\cdot\sigma}\left[\underbrace{\int_{-\infty}^{\infty} y\, e^{-y^2/2\sigma^2}\,dy}_{0} + \mu\int_{-\infty}^{\infty} e^{-y^2/2\sigma^2}\,dy\right]$$

$$\bar{x} = \mu \cdot \frac{1}{\sqrt{2\pi}\cdot\sigma}\,\sqrt{2\pi}\,\sigma \Rightarrow \boxed{\bar{x} = \mu}$$

③ Mean square deviation or dispersion:-

$$\overline{(\Delta x)^2} = \overline{(x-\bar{x})^2} = \overline{(x-\mu)^2}$$

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

Put $x-\mu = y \Rightarrow dx = dy$

② mean value of x :- ①

$$\bar{x} = \int_{-\infty}^{\infty} x\,(P(x))\,dx$$

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

Put $x-\mu = y \Rightarrow x = \mu + y \Rightarrow dx = dy$

$$\bar{x} = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty} (y+\mu)\, e^{-y^2/2\sigma^2}\,dx$$

$$\bar{x} = \frac{1}{\sqrt{2\pi}\cdot\sigma}\left[\underbrace{\int_{-\infty}^{\infty} y\, e^{-y^2/2\sigma^2}\,dy}_{0} + \mu\int_{-\infty}^{\infty} e^{-y^2/2\sigma^2}\,dy\right]$$

$$\bar{x} = \mu \cdot \frac{1}{\sqrt{2\pi}\cdot\sigma}\,\sqrt{2\pi}\,\sigma \Rightarrow \boxed{\bar{x} = \mu}$$

③ mean square deviation or dispersion:-

$$\overline{(\Delta x)^2} = \overline{(x-\bar{x})^2} = \overline{(x-\mu)^2}$$

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

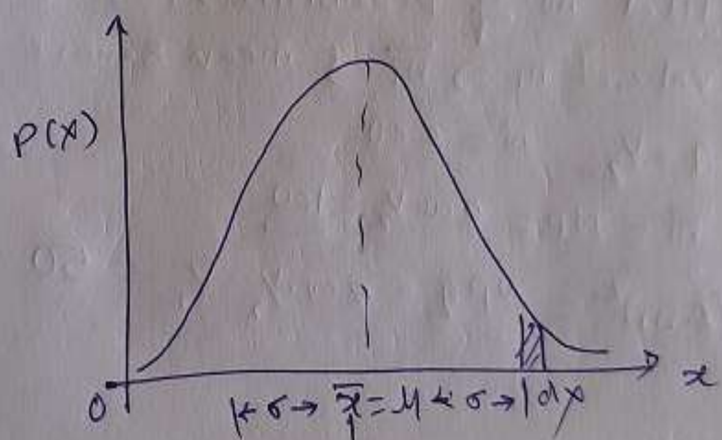
Put $x-\mu = y \Rightarrow dx = dy$

2

$$\overline{(\Delta x)^2} = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty} y^2 e^{-y^2/2\sigma^2}\, dy$$

$$= \frac{1}{\sqrt{2\pi}\cdot\sigma}\,\frac{\sqrt{\pi}}{2}\,(2\sigma^2)^{3/2}$$

$$\overline{(\Delta x)^2} = \sigma^2$$



Reference: A Text book of Thermodynamics and Statistical Physics by V.K.Dhas, Dr.S.D.Aghav, Dr.P.S. Tambade, B.M.Laware : Nirali Prakashan, Pune (Second edition) November-2017