papergrid

Date: / /

# III Probability.

→ Sample Space – Set of all possible outcomes

→ Event – Subset of Sample Space.

→ Probability – If the sample space of an experiment consists of finitely many outcomes that are equally likely, then the probability of an event A is given by:

$$\boxed{P(A) = \frac{\text{no: of points in } A}{\text{no: of points in } S}}$$

• General definition of Probability

Given a sample space S, with each event A of S, there is a no: P(A) called the probability of A, such that the following axioms of probability are satisfied:

1. $0 \leq P(A) \leq 1$ 2. $P(S) = 1$ 3. if $A \cap B = \phi$, then A & B are called Mutually exclusive.

If A and B are mutually exclusive

$$\boxed{P(A \cup B) = P(A) + P(B)}$$

if $A_1 \cap A_2 \cap A_3 \cap \cdots A_i = \phi$, then $P(A_1 \cup A_2 \cup A_3 \cup \cdots A_n)$
$= P(A_1) + P(A_2) + \cdots + P(A_n)$.

• Basic Theorems in Probability

1. $P(A^c) = 1 - P(A)$
2. For event A and B, in sample space S, that may/may not be mutually exclusive.

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i_1 < i_2} P(A_{i_1} \cap A_{i_2})$$
$$+ \sum_{i_1 < i_2 < i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \cdots$$

## • Conditional Probability:

$P(A|B) \rightarrow$ Probability of event A, given that the outcomes are only from B.

reduced sample set

$$\boxed{P(A|B) = \frac{P(A\cap B)}{P(B)}}$$

$P(S) = P(B)$ because given the outcomes are only from B.

$P(A\cap B)$ is the event, because event A, the outcomes in B.

$\therefore P(A|B) = \frac{n(A\cap B)}{n(B)} = \frac{P(A\cap B)}{P(B)}$ //.

tossing a coin, if we get H we can't get T

• If A and B are mutually exclusive. Then $A\cap B = \phi$, So $P(A|B) = P(\phi) = 0$

• If A and B are independent events i.e. Probability of occurence of A doesn't depend on the probability of occurence of B.

Then $P(A\cap B) = P(A)\cdot P(B) \Rightarrow P(A|B) = P(A)$
$P(B|A) = P(B)$ //.

Tossing a coin 'n' times: results of 3rd time doesn't depend on other tosses

## * Multiplication Theorem.

$$\boxed{\begin{aligned} P(A\cap B) &= P(A|B)\cdot P(B) \\ &= P(B|A)\cdot P(A) \end{aligned}}$$

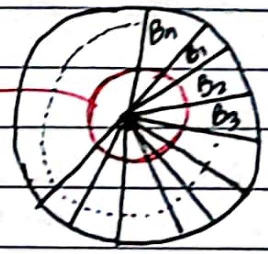


$P(A\cap B) = 0$ (when A and B are mutually exclusive events)

$P(A\cap B) = P(A)\cdot P(B)$ (when A and B are independent events)

## * Bayes Rule.

Consider 'n' mutually exclusive events $B_1, B_2, \cdots B_n$ of Sample space S.




$B_1 \cap B_2 \cap B_3 \cap \cdots B_n = \phi$

$B_1 \cup B_2 \cup B_3 \cup \cdots B_n = S$

$A = (A\cap B_1) \cup (A\cap B_2) \cup (A\cap B_3) \cup \cdots\cdots (A\cap B_n)$

$P(A) = P(A\cap B_1) + P(A\cap B_2) + \cdots\cdots + P(A\cap B_n)$

Total Probability

$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots \cdots + P(A|B_n)P(B_n)$

$$\boxed{P(B_i|A) = \frac{P(A\cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}}$$

(Q1) Four fair sided dice are rolled. Find the probability that sum = 22.

Ans $(6-y_1)+(6-y_2)+(6-y_3)+(6-y_4) = 22$

$24-(y_1+y_2+y_3+y_4) = 22$

$y_1+y_2+y_3+y_4 = 2$ // $y_i$ can be from 0 to 5

$^{n+r-1}C_{r-1} = {}^{4+2-1}C_{4-1} = {}^{5}C_{3} = \frac{5!}{3!\,2!} = \frac{4\times5}{2} = 10$ //.

$\therefore \frac{10}{6^4}$ //

(Q2) A box contains 25 parts of which 10 are defective. Two parts are being drawn simultaneously in a random manner from the box. The prob. of both the parts being good is ——.

Ans $\frac{^{15}C_2}{^{25}C_2} = 7/20$ //

(Q3) A box contains 10 screws, 3 of which are defective. Two screws are drawn at random with replacement. The probability that none of the two screws is defective will be ——.

Ans $^{7}C_2/^{10}C_2 = \underline{7/15}$

Two screws are drawn at random one after the other with replacement. The probability that none of the 2 screws are defective is ——.

$\left(\frac{7\times7}{10\times10}\right) = \underline{0.49}$



* Card Classification

52 Cards → 26 Reds → 13 Hearts, 13 Diamonds

52 Cards → 26 Blacks → 13 Spades, 13 Clubs

A, K, Q, J, 10, 9, 8, ..., 2

Total face Cards → K, Q, J

$= 3\times4 = \underline{\underline{12}}$

(Q4) A no: is selected at random from the 1st 200 natural no:s, Find the probability that the no: is divisible by 6 or 8.

Ans ~~n(6 or 8)~~ $n(6\text{ or }8) = n(6) + n(8) - n(6\,\&\,8)$

$n(6) = 200/6 = 33$ $n(8) = 200/8 = 25$ $n(6\&8) = 200/\text{LCM}(6,8)$

$= \frac{200}{24} = \underline{\underline{8}}$

$\therefore n(6\text{ or }8) = 33+25-8 = \underline{50}$

$\frac{50}{200} = \underline{\underline{0.25}}$

(Q5) A ticket is selected at random from 100 tickets 00, 01, 02 ---- ---99, If X and Y denote the sum & product of the digits on the tickets respectively. The value of $P(X=9\,|\,Y=0)$ is ——.

Ans 00, 01, 02, 03, 04 ---- 10, 20, 30 --- 90 = 19 // (reduced sample space)

09, 90 = 2 //

$\underline{\underline{(2/19)}}$

Date:

(Q) Person A can solve 80% of Gate 2020 paper, Person B can solve 60%. The probability that atleast one of them will solve a problem from the question paper selected at random is ____.

Ans $P(A) = 0.8$ $\quad 1 - P(A' \cap B')$

$P(B) = 0.6$ $\quad 1 - P(A') \cdot P(B') = 1 - 0.2 \times 0.4 = \underline{0.92}$

(Q) A person throws two dice simultaneously. If the sum of the outcomes is 12, he offers lunch at a five star hotel with probability 2/3, if the sum is 7, he offers a lunch with probability 1/2, In all other cases he offers lunch with prob = 1/3.

(i) Find the prob. that lunch is offered ?

(ii) If the lunch is offered, the prob that sum of the outcomes is 12 is ____ ?

Ans (i) $P(LO/sum=12) = 2/3$ $\quad P(LO/sum=7) = 1/2$

$P(LO/sum = otherwise) = 1/3$.

$P(LO) = P(LO/sum=12) P(sum=12) + \ldots$

$\ldots = \underline{\underline{10/27}}$

(ii) $P(sum=12/LO) = P(B_i/A) = \dfrac{P(A|B_i) P(B_i)}{P(A)}$

$P(A) = 10/27$, $P(B_i) = \frac{1}{36}$, $P(A|B_i) = 2/3$.

$= \dfrac{2/3 \times 1/36}{10/27} = 1/20$ //.





Date:

* Random Variable – It is a function whose domain is sample space and whose range is some set of real no:s.

(Eg) X → represents no: of heads when a coin is tossed 3 times.
↳ Random variable.

$X=0 \rightarrow \{TTT\}$ ; $X=1 \rightarrow \{HTT, THT, TTH\}$
... etc.

A random variable maps the sample space to Real No:s.

Here X takes finitely many values (0, 1, 2, 3). A random variable that takes finitely many values is called Discrete Random Variable.

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| P(X) | 1/8 | 3/8 | 3/8 | 1/8 |

• $\sum_{i=0}^{3} P(x=i) = 1/8 + 3/8 + 3/8 + 1/8 = \underline{\underline{1}}$

• $P(X \le 2) = P(X=0) + P(X=1) + P(X=2) = 7/8$ //

$$\boxed{\begin{array}{l} f(x) = P(X=x) \rightarrow \text{Probability Mass Function (PMF)} \\ F(x) = P(X \le x) = \sum_{x_i \le x} P(x = x_i) \rightarrow \text{Cumulative Distribution Function (CDF)} \end{array}}$$

• Continuous Random Variable – A random variable x and its distribution are of continuous type if its cumulative distribution (CDF) F(x) is given by

$$\boxed{F(x) = P(X \le x) = \int_{-\infty}^{x} f(v)\,dv}$$

where f is called Probability density function (PDF) of X.

$$\boxed{f(x) = \frac{d}{dx} F(x)}$$

Date: / /

- Properties of CRV

1. Non-Negativity: The PDF $f(x)$ is a non negative function $(f(x) \geq 0)$
2. Normalization: The total area under the graph of the PDF is equal to unity

$$\boxed{\int_{-\infty}^{\infty} f(x)\,dx = 1}$$

- $P(X \leq x) = P(X < x) = \int_{-\infty}^{x} f(v)\,dv$
- $P(X \geq x) = P(X > x) = \int_{x}^{\infty} f(v)\,dv$
- $P(x_1 \leq X \leq x_2) = P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(v)\,dv$

All these are true for Continuous Random Variables. They may/may not be true for Discrete Random Variables.

Q. The amount of time in hrs that a computer functions before breaking down is a CRV X with PDF given by

$$f(x) = \begin{cases} \frac{e^{-x/100}}{100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that a computer will function b/w 50 & 150 hours before breaking down?

Ans $\int_{50}^{150} \frac{1}{100} e^{-x/100} = \underline{\underline{0.3834}}$

$$\int e^{ax} = \frac{e^{ax}}{a} + C$$





Date: / /

- Expected Value of Random Variable & its properties:

| Discrete Random Variable | Continuous Random Variable |
|---|---|
| $\boxed{E(x) = \sum_{\forall x} x P(X=x)}$ | $\boxed{E(x) = \int_{-\infty}^{\infty} x f(x)\,dx}$ $f(x) = PDF$ |
| Expected Value of (X) or Avg Value (or) Mean | Expected Value (or) Avg Value (or) Mean Value. |

- Properties of Expected Value

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

1. $E(aX) = aE[X]$
2. $E(ax+b) = aE(x)+b$
3. $E(k) = k$
4. $E(E(x)) = E(x)$
5. $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx \rightarrow$ Second Moment

$E(X^3) = \int_{-\infty}^{\infty} x^3 f(x)\,dx \rightarrow$ Third moment

$E(X^n) = \int_{-\infty}^{\infty} x^n f(x)\,dx \rightarrow$ $n^{th}$ moment.

6. $E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x)\,dx$ (Expected Value of function of random variable)

- Varience / Second Central Moment of a Random Variable.

Varience $= E((X - E(x))^2)$

$= E(x^2 + E(x)^2 - 2xE(x)) = E(x^2) + E(E(x)^2) - 2E(x)E(x)$

$= E(x^2) + E(x)^2 - 2E(x)^2$

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

Correlation coefficient 'r' is $r = \dfrac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}$,



$\therefore$ Variance $= E\left((X-E(X))^2\right)$
$= E(X^2) - E(X)^2 = Var(X)$

Standard Deviation of X $= \sqrt{Var(X)} = \sqrt{E(X^2)-E(X)^2}$

Properties:
1. $Var(X) \geq 0$
2. $Var(aX) = a^2 Var(X)$
3. $Var(k) = 0$
4. $Var(aX \pm b) = a^2 Var(X)$
5. $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$
6. $Var(X-Y) = Var(X) + Var(Y) - 2Cov(X,Y)$

where Covariance $Cov(X,Y) = E(XY) - E(X)E(Y)$

- If X and Y are independent RVs, then

$Cov(XY) = E(XY) - E(X)E(Y)$
$= E(X)E(Y) - E(X)E(Y) = 0$

$\therefore Var(X+Y) = Var(X-Y) = \underline{Var(X) + Var(Y)}$

$Cov(X,Y) = E\left((x-E(x))(y-E(y))\right)$

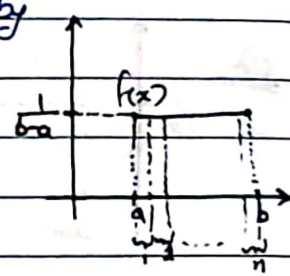




## * Probability Distribution functions

1. Uniform Random Variable. - The probability density function of a uniform random variable is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$




$$E(X) = \int_a^b x \frac{1}{b-a}\,dx = \left(\frac{x^2}{2}\right)_a^b \frac{1}{b-a} = \frac{a+b}{2}$$

$$E(X^2) = \int_a^b x^2 \frac{1}{b-a}\,dx = \frac{a^2+b^2+ab}{3}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{(b-a)^2}{12}$$

$E(X) = \dfrac{a+b}{2}$

$E(X^2) = \dfrac{a^2+b^2+ab}{3}$

$Var(X) = \dfrac{(b-a)^2}{12}$

- If we divide $[a,b]$ into n equal intervals.

Then for each interval $P(x) = \underset{(\text{height})}{\dfrac{1}{b-a}} \times \underset{(\text{width})}{\dfrac{b-a}{n}} = \underline{1/n}$ (area = probability for the interval)

$P(x_{n-1} < X < x_n) = 1/n$ (X is a uniform random variable).

Q1) A point is randomly selected with uniform probability in the XY plane within the rectangle with corners at (0,0), (1,0), (1,2), (0,2). If P is the length of the position vector of the point, the expected value of $P^2$ is

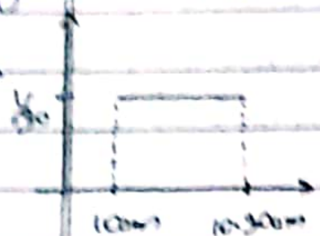
Ans $P^2 = X^2 + Y^2$ ; $E(X^2+Y^2) = E(X^2) + E(Y^2)$
$= \dfrac{1^2 + 0\times1 + 0^2}{3} + \dfrac{2^2 + 2\times0 + 0^2}{3} = 5/3$

* Let $X \sim U(0,1)$
$Y \sim U(0,2)$

($X \sim U(0,1)$ means, X uniformly in (0,1))

(Q2) A passenger arrives at a bus stop at 10:00 am, knowing the bus will arrive at some time uniformly distributed b/w 10am & 10:30am. What is the probability that he will have to wait longer than 10 minutes?

Ans

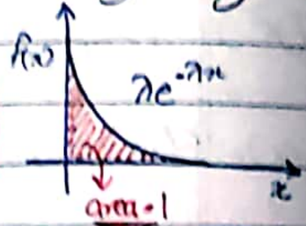



$$P(X>10) = \int_{10}^{30} \frac{1}{30}\,dx = \frac{20}{30} = \underline{\underline{2/3}}$$

## 2. Exponential Random Variable

The probability density function of a exponential R.V is given by:

$$\boxed{f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \quad \lambda > 0 \\ 0 & \text{otherwise} \end{cases}}$$




$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{\infty} f(x)dx = 0 + \int_0^\infty \lambda e^{-\lambda x}dx = \left(\frac{\lambda e^{-\lambda x}}{-\lambda}\right)_0^\infty$$

$$= -\left(e^{-\lambda x}\right)_0^\infty = -\left(e^{-\infty} - e^{0}\right) = -\left(\frac{1}{e^{\infty}} - 1\right) = \underline{+1}$$

becomes zero when $x \to \infty$ because exponential distribution dominates.

$$E(X) = \int_0^\infty x\lambda e^{-\lambda x}dx = \lambda\left[\frac{xe^{-\lambda x}}{-\lambda} - \int \frac{1\cdot e^{-\lambda x}}{-\lambda}dx\right]_0^\infty = \frac{1}{\lambda}$$

$$E(X^2) = \int_0^\infty x^2\lambda e^{-\lambda x}dx = \lambda\left[\frac{x^2 e^{-\lambda x}}{-\lambda} - 2\int \frac{xe^{-\lambda x}}{-\lambda}dx\right]_0^\infty = 2/\lambda^2$$

$$\boxed{\begin{aligned} E(X) &= 1/\lambda \\ E(X^2) &= 2/\lambda^2 \\ Var(X) &= 1/\lambda^2 \end{aligned}}$$





- In practice, the exponential distribution often arises as the distribution of the amount of time until some specific event occurs.

Eg1: The amount of time until an earthquake occurs.

Eg2: The amount of time until a new war breaks out.

Eg3: The amount of time until a telephone call you receive turns out to be a wrong number.

(Q1) The length of the shower on a tropical island during rainy season has an exponential distribution with parameter 2, time being measured in minutes. What is the probability that a shower will last more than 3 minutes?

Ans $f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & \text{others} \end{cases}$ $\quad \lambda = 2.$ $\quad \int_3^\infty f(x)dx = 2\int_3^\infty e^{-2x} = \frac{1}{e^6}$

$= \underline{\underline{0.0025}}$

## 3. Poisson Random Variable.

The probability mass function of Poisson R.V is given by:

$$\boxed{P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0,1,2,\ldots\ldots \ (x \text{ is discrete R.V}) \quad \lambda = \text{parameter } (\lambda > 0)}$$

$$\sum_{\forall x} P(X=x) = \underline{1} : \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ldots\ldots\right) = e^{-\lambda}.e^{+\lambda} = \underline{\underline{1}}$$

$$E(X) = \sum_{\forall x} x\,P(X=x) = \sum_{x=0}^{\infty} x\,\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\left(0 + \frac{\lambda}{1!} + \frac{2\lambda^2}{2!} + \cdots \frac{x\lambda^x}{x!}\right)$$

$$= e^{-\lambda}\cdot\lambda\left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \cdots\right] = e^{-\lambda}\cdot\lambda\cdot e^{\lambda} = \underline{\lambda}$$

$$E(X^2) = \sum_{\forall x} x^2 P(X=x) = e^{-\lambda}\sum_{\forall x}\frac{x^2\lambda^x}{x!}$$

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

$$= \lambda e^{-\lambda}\left(e^{\lambda} + \lambda e^{\lambda}\right) = \lambda(1+\lambda)$$

$$\boxed{\begin{aligned} E(X) &= \lambda \\ E(X^2) &= \lambda(\lambda+1) \\ Var(X) &= E(X) = \lambda \end{aligned}}$$

- Some examples of RV that generally obey the poisson probability distribution are.
  1. No. of misprints on a page of a book
  2. The no. of people in a community who survive the age of 100.
  3. The no. of wrong telephone numbers dialed in a day.
  4. The no. of customers entering a post office on a given day, etc.
- Events associated with arrival rate, defect rate and rare occurrence we can use Poisson Random variables.

Q1 An observer counts 240 vehicles/hr at a specific highway location. Assume that the vehicle arrival at the location is poisson distributed. Find the probability of having one vehicle arriving over a 30 second time interval.

Ans $\frac{240\text{ vehicles}}{60\times60\text{ s}} = \frac{1\text{ vehicle}}{15\text{ sec}} = 2\text{ vehicles}/30\text{ sec} = \lambda = 2.$

$P(X=x) = \frac{e^{-2}2^x}{x!}$, $P(X=1) = \frac{e^{-2}\times 2}{1!} = \frac{2}{e^2} = \underline{0.27}$





Q2 Suppose 300 misprints are distributed randomly throughout a book of 500 pages. The probability that a given page contains 2 or more misprints is ____.

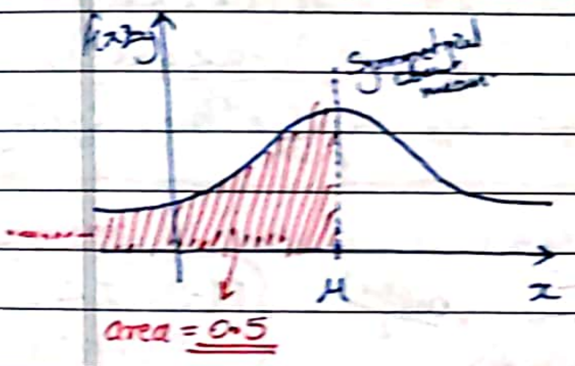
Ans $\frac{300}{500} = \frac{3}{5} = \lambda$. $f(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-3/5}(3/5)^x}{x!}$

$$1 - P(X<2) = 1 - \left(\frac{e^{-3/5}(3/5)^0}{0!} + \frac{e^{-3/5}(3/5)^1}{1!}\right)$$

$= \underline{0.122}$

## 4. Normal or Gaussian Random Variable.

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty} \quad \text{(PDF)}$$




$P(X<\mu) = P(X>\mu) = 0.5$

$f(x) \geq 0$

$\int_{-\infty}^{\infty} f(x)\,dx = 1$

$\int_{-\infty}^{\mu} f(x)\,dx = \int_{\mu}^{\infty} f(x)\,dx = \frac{1}{2}$

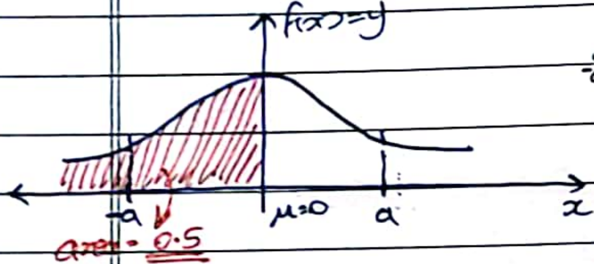
papergrid

Date: / /

When a Normal RV has $\mu = 0$ and $\sigma^2 = 1$, then we call it Standard Normal Random Variable.

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

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




$Z \sim N(0,1)$

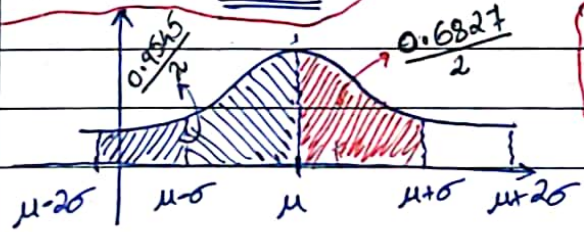
($Z$ is a Normal RV with $\mu = 0, \sigma^2 = 1$)

- Properties → • $P(Z > a) = P(Z < -a)$
  - $P(-a < Z < a) = 2P(0 < Z < a)$
    $= 2P(-a < Z < 0) //$

- $\boxed{P(-1 < Z < 1) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 2\int_{0}^{1} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 0.6827}$

- $\boxed{P(-2 < Z < 2) = 0.9545}$

$\boxed{P(-3 < Z < 3) = 0.9973}$






papergrid

Date: / /

If $X$ is a normal random variable with mean $= \mu$, variance $= \sigma^2$

Then $\boxed{Z = \frac{X - \mu}{\sigma}}$ is also a Normal RV. with $\mu = 0$ and $\sigma^2 = 1$ $\quad Z \sim N(0,1)$

- $E(Z) = E\left(\frac{X}{\sigma} - \frac{\mu}{\sigma}\right) = E\left(\frac{X}{\sigma}\right) - E\left(\frac{\mu}{\sigma}\right) = \frac{1}{\sigma}E(X) - \frac{\mu}{\sigma} = \frac{\mu}{\sigma} - \frac{\mu}{\sigma}$
  $= 0 //$

- $Var(Z) = Var\left(\frac{X}{\sigma} - \frac{\mu}{\sigma}\right) \quad (Var(aX+b) = a^2 Var(X))$
  $= \frac{1}{\sigma^2} Var(X) = \frac{\sigma^2}{\sigma^2} = 1 //$

- $Y = aX + b$ is also a Normal Random Variable.
  with $\boxed{\begin{array}{l} \mu = E(Y) = a\mu + b = E(aX+b) \\ \sigma^2 = Var(Y) = a^2\sigma^2 = a^2 Var(X) \end{array}}$

(Q1) If the mass of 300 students is normally distributed with mean 68 kgs and standard deviation 3 kgs. How many students have mass (i) Greater than 72 kg (ii) Less than / equal to 64 kg.

Area under normal curve from $z=0$ to $z=1.33$ is $0.4082$

Ans $\mu = 68$, $\sigma = 3$. $\quad Z = \frac{X - 68}{3}$ ($X$ = mass)

(i) $P(\text{mass} > 72) = P\left(\frac{\text{mass} - 68}{3} > \frac{72 - 68}{3}\right)$

$= P(Z > 4/3) = P(Z > 1.33) = 0.5 - 0.4082 = 0.0918$

$300 \times 0.0918 \approx 28$ students

$$(ii)\ P\left(Z \le \frac{64-68}{3}\right) = P(Z \le -4/3) = P(Z \ge 4/3) = 0.0918 \Rightarrow 28 \text{ students}$$

papergrid

Date: / /

## 5. Bernoulli & Binomial Random Variable.

- **Bernoulli Random Variable** $\Rightarrow$ Trial/error experiment, whose outcome can either be success $(x=1)$ or failure $(x=0)$.

$$\boxed{P(X=1) = p \quad ; \quad P(X=0) = 1-p \qquad 0<p<1}$$

- **Binomial Random Variable** - Suppose 'n' independent trials each of which results in a success with probability 'p' or failure with probability $(1-p)$ are to be performed. If X represents the no. of successes that occur in the 'n' trials then X is said to be Binomial RV with parameters $(n, p)$. Its probability mass function is given by.

$$\boxed{P(X=x) = {}^nC_x\, p^x (1-p)^{n-x}}$$

Choose x successes.

(Bernoulli's RV is a special case of Binomial RV when $n=1$)

$$\sum_{x=0}^{n} {}^nC_x\, p^x (1-p)^{n-x} = \binom{p=a}{1-p=b} =$$

$${}^nC_0 a^0 b^n + {}^nC_1 a^1 b^{n-1} + \cdots {}^nC_n a^n b^0 = (a+b)^n$$

$$= (p+1-p)^n = 1^n = 1$$

$$\boxed{\begin{aligned} E(X) &= \text{mean} = np \\ Var(X) &= np(1-p) = npq \end{aligned}}$$

papergrid

## * Normal Approximation to Binomial Random Variable.

When 'n' is large, a binomial random variable with parameters 'n' and 'p' will have approximately the same distribution as a normal random variable with the same mean & variance as the binomial. If X denotes the no. of success that occur when 'n' independent trials each resulting in a success with probability 'p' are performed then for any $a<b$.

$$\boxed{P\left(a \le \frac{X-np}{\sqrt{np(1-p)}} \le b\right) = P(a \le Z \le b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz}$$

(Q1) Assume that 4 percent of the population over 65 years old has some disease. If a random sample of 3500 people over 65 years is taken, then the probability that fewer than 150 of them have the disease is ______.

Area under normal curve b/w $z=0$ and $z=0.86$ is $0.3051$.

Ans $= {}^{3500}C_0 (0.04)^0 (0.96)^{3500} + {}^{3500}C_1 (0.04)(0.96)^{3499} + \cdots$
$\cdots\ {}^{3500}C_{149} (0.04)^{149} (0.96)^{3500-149}$

(p = 0.04, 1-p = 0.96)

n is large, so we cannot use Binomial RV. We have approximation

1. Poisson approx. of Binomial RV $(n \to \infty, p \to 0)$
2. Normal approx. of Binomial RV.

In Q normal curve diagram, solve with the approx.

$\mu = np = 3500 \times 0.04 = 140$

$\sigma^2 = npq = 140 \times 0.96 = 134.4$

$P(X < 150) \quad (X \sim N(140, 134.4))$

$$Z = \frac{X-\mu}{\sigma} = \frac{X-140}{\sqrt{134.4}} = \frac{X-140}{11.59}$$

$$P\left(Z < \frac{150-140}{11.59}\right) = P(Z < 0.86) = 0.3051 + \frac{1}{2}$$

$$= 0.8051$$

## * Poisson approximation to Binomial Random Variable.

When 'n' is large and 'p' is small then Binomial distribution is very closely approximated by poisson distribution.

$$\boxed{\begin{array}{l} P(X=x) = {}^{n}C_{x}\, p^{x} (1-p)^{n-x} \\ P(X=x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!} \quad \lambda \text{ is mean of poisson RV} \\ \qquad\qquad\qquad\quad \lambda = np \end{array}}$$

Poisson distribution is a limiting case of Binomial distribution as $n \to \infty$ and $p \to 0$.

## * Sum of Independent Random Variables

- If X and Y are independent poisson RVs with respective parameters $\lambda_1$ and $\lambda_2$ then $X+Y$ has a poisson distribution with parameter = $\boxed{\lambda_1 + \lambda_2}$



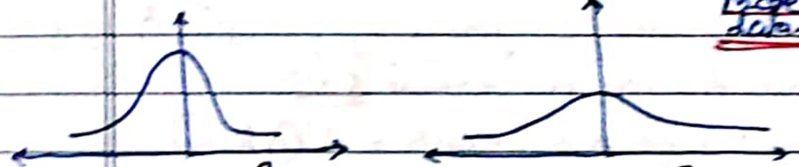
- If X and Y are independent Binomial RVs with respective parameters $(n,p)$ and $(m,p)$ then $X+Y$ is a Binomial RV with parameters = $\boxed{(n+m), p}$

- If X and Y are independent RVs with respective parameters $(\mu_1, \sigma_1^2)$ and $(\mu_2, \sigma_2^2)$ then $X+Y$ is Normal RV with parameters = $\boxed{\mu_1+\mu_2,\ \sigma_1^2+\sigma_2^2}$

## * Significance of Variance

| | Mean E(x) | Variance (x) |
|---|---|---|
| $X=0$ $P(X=0)=1$ | 0 | 0 |
| $Y \sim U(-1,1)$ | $\frac{-1+1}{2} = 0$ | $(1--1)^2/12 = 1/3$ |
| $X \sim U(-100,100)$ | $\frac{-100+100}{2} = 0$ | $(100+100)^2/12 = \frac{200^2}{12}$ |

Means all the same = 0 | More the variance, More the spread of data.



$\sigma_1^2$ $\qquad$ $\sigma_2^2$

$$\underline{\underline{\sigma_2^2 > \sigma_1^2}}$$