Baye's Theorem:

Suppose that $A, B_1, B_2, \ldots, B_n$ are events from a sample space S. Suppose that $\bigcup_{i=1}^{n} B_i = S$ and that $B_i \cap B_j = \phi$ for all $i \neq j$. Suppose $P(A) > 0$, and $P(B_j) > 0$ for all j. Then for all j:

$$P(B_j | A) = \frac{P(A|B_j)\, P(B_j)}{\sum_{i=1}^{n} P(A|B_i) P(B_i)}$$

Proof :-

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

$$= \frac{P(A|B_j) P(B_j)}{P(A)} \quad \left[\text{By } P(A/B_j) = \frac{P(A\cap B_j)}{P(B_j)}\right] \quad \text{--- (1)}$$

A can be written as $A = \bigcup_{i=1}^{n} (A \cap B_i)$ $\left[\because \bigcup_{i=1}^{n} B_i = S \text{ \& } B_i \cap B_j = \phi\right]$

Then $P(A) = P\left(\bigcup_{i=1}^{n} (A \cap B_i)\right)$

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

[By addition theorem]

$$= \sum_{i=1}^{n} P(A \cap B_i)$$

$$= \sum_{i=1}^{n} P(A|B_i) P(B_i)$$

$$\therefore (1) \Rightarrow P(B_j|A) = \frac{P(A/B_j)\, P(B_j)}{\sum_{i=1}^{n} P(A/B_i)\, P(B_i)}$$

Example 1

Three machines M_1, M_2 and M_3 produce identical items f their respective output 5%, 4% & 3% of items re faulty. On a certain day, M_1 has produced 5% of the total output, M_2 has produced 30% ind M_3 the remainder. An item selected at random s found to be faulty. What are the chances that it was produced by the machine with he highest output?

ol^n:

et the event of drawing a faulty item from ny of the machines be A, and the event hat an item drawn at random was produced y M_i be B_i. We have to find $P(B_i/A)$ for which we proceed as follows:

	M_1	M_2	M_3	
$P(B_i)$	0.25	0.30	0.45	∵ Sum = 1
$P(A/B_i)$	0.05	0.04	0.03	
$P(B_i)P(A/B_i)$	0.0125	0.012	0.0135	Sum = 0.38
$P(B_i/A)$	$\frac{0.0125}{0.038}$	$\frac{0.012}{0.038}$	$\frac{0.0135}{0.038}$	by Baye's Theorom.

The highest output being from M_3, the required probability $= \frac{0.0135}{0.038} = 0.355$

* Theo

* There are three bags: first containing 1 white, 2 red, 3 green balls; second 2 white, 3 red, 1 green balls and third 3 white, 1 red, 2 green balls. Two balls are drawn from a bag chosen at random. These are found to be one white and one red. Find the probability that the balls so drawn came from the second bag. (13)

Sol^n:

Let B_1, B_2, B_3 pertain to the first, second, third bags chosen and A - the two balls are white & red.

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$$

$P(A|B_1) = P$ (a white and a red ball are drawn from first bag)

$$= \frac{{}^1C_1 \times {}^2C_1}{{}^6C_2} = \frac{2}{15}$$

|||ly

$$P(A|B_2) = \frac{{}^2C_1 \times {}^3C_1}{{}^6C_2} = \frac{2}{5}$$

$$P(A|B_3) = \frac{{}^3C_1 \times {}^1C_1}{{}^6C_2} = \frac{1}{5}$$

By Baye's theorem, $P(B_2|A) = \dfrac{P(B_2)\,P(A|B_2)}{P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)\,P(A|B_3)}$

$$= \frac{\frac{1}{3} \times \frac{2}{5}}{\frac{1}{3}\times\frac{2}{15} + \frac{1}{3}\times\frac{2}{5} + \frac{1}{3}\times\frac{1}{5}}$$

$$= \frac{6}{11}$$

You go to see the doctor about an ingrowing toenail. The doctor selects you at random to have a blood test for swine flue, which for the purposes of this exercise we will say is currently suspected to affect 1 in 10,000 people in Australia. The test is 99% accurate, in the sense that the probability of a false positive is 1%. The probability of a false negative is zero. You test positive. What is the new probability that you have a swine flue?

Sol^n:

Let $P(B)$ be the probability you have swine flu.

Let $P(T)$ be the probability of a positive test.

We want to know $P(B|T)$

From Baye's theorem $P(B|T) = \dfrac{P(T|B)\, P(B)}{P(T)}$

$$= \frac{P(T|B)\, P(B)}{P(T|B)\, P(B) + P(T|NB)\, P(NB)}$$

where $P(NB)$ is the probability of not having swine flu

$P(B) = \dfrac{1}{10{,}000} = 0.0001$ (a priori probability you have swine flu)

$P(NB) = 1 - 0.0001 = 0.9999$

$P(T|B) = 1$ (if you have swine flu the test is always positive)

$P(T|NB) = 0.01$ (1% chance of a false positive)

$$P(B|T) = \frac{1 \times 0.0001}{1 \times 0.0001 + 0.01 \times 0.9999} \approx 0.01$$

Given that the test was positive your chance of having swine flue

14

A factory uses three machines X, Y, Z to produce certain items.

(i) Machine X produces 50 percent of the items of which 3 percent are defective.

(ii) Machine Y produces 30 percent of the items of which 4 percent are defective.

(iii) Machine Z produces 20 percent of the items of which 5 percent are defective.

Suppose a defective item is found among the output. Find the probability that it came from each of the machines.

	X	Y	Z	
$P(P_i)$	0.50	0.30	0.20	sum = 1
$P(D/P_i)$	0.03	0.04	0.05	
$P(P_i)P(D/P_i)$	0.015	0.012	0.010	sum = 0.037
$P(P_i/D)$	$\frac{0.015}{0.037}$	$\frac{0.012}{0.037}$	$\frac{0.010}{0.037}$	
	40.5%	32.5%	27.0%	

Scanned by CamScanner

Of the 3 men, the chances that a politician, a businessman and an academician will be appointed is a vice-chancellor of a university are 0.50, 0.30 and 0.20 respectively. Probability that research is promoted by these people if they are appointed as V.C are 0.3, 0.7 & 0.8 respectively.

(a) Determine the probability that research is promoted in the university,

(b) If research is promoted in the university, what is the probability that the V.C is an (i) academician? (ii) businessmann.

Soln

	P	B	A	
$P(A_i)$	0.50	0.30	0.20	sum = 1
$P(R/A_i)$	0.3	0.7	0.8	
$P(A_i)P(R/A_i)$	0.15	0.21	0.16	sum = 0.52
$P(A_i/R)$	$\frac{0.15}{0.52}$	$\frac{0.21}{0.52}$	$\frac{0.16}{0.52}$	
	0.28846	0.4038	0.30769	

(a) $P(R) = 0.52$

b) (i) $P(A_i/R) = 0.30769$

(ii) $P(A_i/R) = 0.4038$

15 16

* In a certain college 25% of boys and 10% of girls are studying mathematics. The girls constitute 60% of the student body.

(a) What is the probability that mathematics is being studied?

(b) If a student is selected at random and is found to be studying Mathematics, find the probability that the student is a (i) girl (ii) a boy.

Sol^n

	B.	G	
$P(B_i)$	0.4	$\cancel{\frac{60}{100}}$ 0.6	sum = 1
$P(M/B_i)$	0.25	0.10	
$P(B_i)P(M/B_i)$	0.1	0.06	sum = 0.16
$P(B_i/M)$	$\frac{0.10}{0.16}$	$\frac{0.06}{0.16}$	
	0.625	0.375	

(a) $P(M) = 0.16$

(b) (i) $P(B_i/M) = 0.375$

(ii) $P(B_i/M) = 0.625$

Scanned by CamScanner

* For a certain binary communication channel, the probability that a transmitted '0' is received as a '0' is 0.95 and the probability that a transmitted '1' is received as '1' is 0.90. If the probability that a '0' is transmitted is 0.4, find the probability that (i) a '1' is received (ii) a '1' was transmitted given that a '1' was received.

	0	1	
$P(B_i)$	0.4	0.6	sum=1
$P(R/B_i)$	$\cancel{0.95}$ 0.05	0.90	
$P(B_i)P(R/B_i)$	$\cancel{0.38}$ 0.02	0.54	sum = 0.56
$P(B_i/R)$	$\frac{0.02}{0.56}$	$\frac{0.54}{0.56}$	
	0.0357	0.9643	

(i) $\cancel{P(a) = 0.9643}$ $P(R) = 0.56$

ii) $P(B_i/R) = 0.9643$

Random Variables

A random variable, usually written X, is a variable whose possible values are numerical outcomes of a random phenomenon. There are two types of random variables, discrete and continuous.

A discrete random variable is one which may take on only a countable number of distinct values such as 0, 1, 2, 3, 4 . . . . . Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, ~~the~~ the Friday night attendance in a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

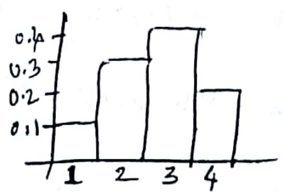
The probability distribution of a discrete
random variable is a list of probabilities
associated with each of its possible values.
It is also sometimes called the probability
function or the probability mass function.

Suppose a random variable X may take
k different values, with probability that $X = x_i$
defined to be $P(X = x_i) = p_i$. The probabilities p_i
must satisfy the following:

(i) $0 \le p_i \le 1$ for each i

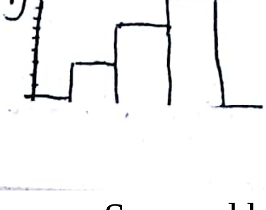
(ii) $p_1 + p_2 + \cdots + p_k = 1$.

Ex: Suppose a random variable X takes the values
1, 2, 3 or 4. The probabilities associated with each
outcome are as follows:



outcome:	1	2	3	4
probability:	0.1	0.3	0.4	0.2

Cummulative distribution function is a function
giving the probability that the random variable X
is less than or equal to x, for every value x.



outcome : 1 2 3 4

Random Variable

A random variable is a rule that assigns a numerical value to each possible outcomes of a probabilistic experiment.
We denote a random variable by a capital letter, say X.

example
X: the age of a randomly selected student in a class.

A discrete random variable can take only distinct, separate values.

example
X: number of heads when tossing 3 coins.

A continuous random variable can take any value in some interval.

example.
X: time a customer spends waiting in line at the store.

Probability function / probability mass function

A function f whose value for each real number x is given by $f(x) = P(X=x)$, s called the probability function of the random variable X.

xample :

Consider an experiment of tossing 3 coins.

Sample space = $\begin{Bmatrix} HHH, & HHT, & HTH, & HTT \\ TTT, & THT, & TTH, & THH \end{Bmatrix}$

X be defined as the number of heads.

$X = \{0, 1, 2, 3\}$

$P(X=0) = \frac{1}{8}$

$P(X=1) = \frac{3}{8}$

$P(X=2) = \frac{3}{8}$

$P(X=3) = \frac{1}{8}$

example

Consider an example of rolling two dies.

Sample space = $\{(i,j) \mid i,j = 1,2,3,4,5,6\}$

X be defined as the sum of the number on the dies

$$X = \{2,3,4,5,6,7,8,9,10,11,12\}$$

$$f(2) = P(X=2) = P\{(1,1)\} = \frac{1}{36}$$

$$f(3) = P(X=3) = P\{(1,2),(2,1)\} = \frac{2}{36}$$

$\therefore$

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Note: The probabilities of X in the table add upto 1.

Probability Distribution Function.

The function f is called a probability distribution function of the random variable X, if it satisfies the conditions:

(i) $f(x) \geq 0 \quad (P(X=x_i) \geq 0)$

(ii) $\sum f(x) = 1 \quad (\sum P(X=x_i) = 1)$

Mean or Expected Value, Variance, Standard Deviation

The mean or expected value of a random variable X is defined as

$$E(X) = \mu = \sum P(X=x_i)\, x_i = \sum p_i x_i$$

The variance of a random variable X is defined as

$$Var(X) = \sigma^2 = \sum P(X=x_i)(x_i-\mu)^2$$

$$= \sum p_i (x_i-\mu)^2$$

$$\left[\sum p_i(x_i-\mu)^2 = \sum p_i(x_i^2 - 2\mu x_i + \mu^2) = \sum p_i x_i^2 - 2\mu \sum p_i x_i + \mu^2 \sum p_i = \sum p_i x_i^2 - 2\mu \times \mu + \mu^2 \times 1 = \sum p_i x_i^2 - \mu^2\right]$$

Standard Deviation of a random variable X is defined as

$$SD(X) = \sigma$$

$$= \sqrt{\sigma^2}$$

$$= \sqrt{Var(X)}$$

ation

Let X be the number that comes up on a roll of one die. Compute the mean, variance and standard deviation of X.

<u>Sol^n</u> $X = \{1, 2, 3, 4, 5, 6\}$.

$f(x_1) = f(1) = P(X=1) = \frac{1}{6} = p_1$ $f(x_4) = f(4) = P(X=4) = \frac{1}{6} = p_4$

$f(x_2) = f(2) = P(X=2) = \frac{1}{6} = p_2$ $f(x_5) = f(5) = P(X=5) = \frac{1}{6} = p_5$

$f(x_3) = f(3) = P(X=3) = \frac{1}{6} = p_3$ $f(x_6) = f(6) = P(X=6) = \frac{1}{6} = p_6$

$$\text{mean} = \sum p_i x_i = p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 + p_5 x_5 + p_6 x_6$$

$$= \frac{1}{6}\times 1 + \frac{1}{6}\times 2 + \frac{1}{6}\times 3 + \frac{1}{6}\times 4 + \frac{1}{6}\times 5 + \frac{1}{6}\times 6$$

$$\mu = 3.5.$$

$$\text{variance} = \sum p_i (x_i - \mu)^2$$

$$= p_1(x_1-3.5)^2 + p_2(x_2-3.5)^2 + p_3(x_3-3.5)^2 + p_4(x_4-3.5)^2 + p_5(x_5-3.5)^2 + p_6(x_6-3.5)^2$$

$$= \frac{1}{6}(1-3.5)^2 + \frac{1}{6}(2-3.5)^2 + \frac{1}{6}(3-3.5)^2 + \frac{1}{6}(4-3.5)^2 + \frac{1}{6}(5-3.5)^2 + \frac{1}{6}(6-3.5)^2$$

$$\sigma^2 = 2.917.$$

$$\text{Standar deviation} = \sqrt{\text{variance}}$$

$$= \sqrt{2.917}$$

$$1.708$$

* Find the mean and variance of the no of tails when 2 coins are tossed.

Sol^n:

Sample space = {HH, HT, TH TT}

$X = \{0, 1, 2\}$

$f(0) = P(X=0) = P\{TT\} = \frac{1}{4} = P_1$, III^{ly} $f(1) = \frac{2}{4} = P_2$, $f(2) = \frac{1}{4} = P_3$

x_i	0	1	2
$f(x_i)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$$\mu = \sum p_i x_i$$

$$= \frac{1}{4}\times 0 + \frac{2}{4}\times 1 + \frac{1}{4}\times 2$$

Mean $\underline{\mu = 1}$

$$\sigma^2 = \sum p_i (x_i - \mu)^2$$

$$= \sum p_i x_i^2 - \mu^2$$

$$= \frac{1}{4}\times 0^2 + \frac{2}{4}\times 1^2 + \frac{1}{4}\times 2^2 - 1^2$$

Variance $\sigma^2 = \frac{1}{2}$

S.D $\sigma = \frac{1}{\sqrt{2}}$

* If X is a random variable with $P(X=x) = \frac{1}{2^x}$ (4)
where $x = 1, 2, 3, \ldots \infty$, find
(i) $P(X)$ (ii) $P(X = \text{even})$ (iii) $P(X = \text{divisible by } 3)$

(i) $P(X) = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots + \cancel{\frac{1}{4}} \ldots$

which is a geometric progression
with $a = \frac{1}{2}$ & $r = \frac{1}{2}$

$$S_\infty = \frac{a}{1-r} = \frac{1/2}{1-1/2} = 1$$

(ii) $P(X = \text{even}) = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \ldots$

here $a = \frac{1}{2^2}$, $r = \frac{1}{2^2}$

$$\therefore S_\infty = \frac{a}{1-r} = \frac{1/2^2}{1-1/2^2} = \frac{1/4}{3/4} = \frac{1}{3}$$

(iii) $P(X = \text{divisible by } 3) = \frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \ldots$

here $a = \frac{1}{2^3}$, $r = \frac{1}{2^3}$

$$\therefore S_\infty = \frac{a}{1-r} = \frac{1/2^3}{1-\frac{1}{2^3}} = \frac{1/8}{7/8} = \frac{1}{7}$$

Verify if the following function can be a probability distribution function. Find the mean if they are so.

(a) $f(x) = \frac{x}{24}$ where $x = \{0, 1, 2, 3, 4\}$

(b) $f(x) = \frac{15 - x^2}{24}$ where $x = \{0, 1, 2, 3, 4\}$

Sol^n

(a) $f(x) = \frac{x}{24}$

(i) $f(x) = \{0, \frac{1}{24}, \frac{2}{24}, \frac{3}{24}, \frac{4}{24}\}$ for $x = \{0, 1, 2, 3, 4\}$

$f(x) \geq 0$, for every x

(ii) $\Sigma f(x) = 0 + \frac{1}{24} + \frac{2}{24} + \frac{3}{24} + \frac{4}{24}$

$= \frac{10}{24} \neq 1$

$\therefore$ It is not a PDF.

(b) $f(x) = \frac{15 - x^2}{24}$

(i) $f(x) = \{\frac{15}{24}, \frac{14}{24}, \frac{11}{24}, \frac{6}{24}, -\frac{1}{24}\}$ for $x = \{0, 1, 2, 3, 4\}$

$f(x)$ is negative when $x = 4$.

ie, $f(x) < 0$ $(\not> 0)$

Hence it is not a PDF

* Show that the distribution represents a probability distribution. Find the mean & S.D. (5)

x	0	1	2	3	4	5
$f(x)$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{5}{32}$	$\frac{1}{32}$

Soln:

(i) $f(x) \geq 0$ for every x.

(ii) $\Sigma f(x) = \frac{1}{32} + \frac{5}{32} + \frac{5}{16} + \frac{5}{16} + \frac{5}{32} + \frac{1}{32}$

$= \frac{1+5+10+10+5+1}{32}$

$= 1$

Hence the distribution represents a PDF.

$\mu = \Sigma p_i x_i$

$= \frac{1}{32}\times 0 + \frac{5}{32}\times 1 + \frac{5}{16}\times 2 + \frac{5}{16}\times 3 + \frac{5}{32}\times 4 + \frac{1}{32}\times 5$

$= \frac{80}{32}$

mean $\mu = 2.5$

$\sigma^2 = \Sigma p_i x_i^2 - \mu^2$

$= \frac{1}{32}\times 0 + \frac{5}{32}\times 1 + \frac{5}{16}\times 4 + \frac{5}{16}\times 9 + \frac{5}{32}\times 16 + \frac{1}{32}\times 25 - (2.5)^2$

var $\sigma^2 = 1.25$

$\sigma = \sqrt{1.25}$

SD $\sigma =$

Scanned by CamScanner

1 The PDF of random variable X is given by the table

x	0	1	2	3	4	5	6
$P(X=x)=f(x)$	k	$3k$	$5k$	$7k$	$9k$	$11k$	$13k$

For what value of k is it a valid probability istribution. Also find (a) $P(X \geq 5)$ (b) $P(3 \leq X \leq 6)$

For a valid PDF $\sum P(X) = 1$

ie, $k + 2k + 5k + 7k + 9k + 11k + 13k = 1$

$$\Rightarrow 49k = 1 \Rightarrow k = 1/49$$

; The table can be written as

x	0	1	2	3	4	5	6
$f(x)$	$\frac{1}{49}$	$\frac{3}{49}$	$\frac{5}{49}$	$\frac{7}{49}$	$\frac{9}{49}$	$\frac{11}{49}$	$\frac{13}{49}$

a) $P(X \geq 5) = P(X=5) + P(X=6)$

$$= \frac{11}{49} + \frac{13}{49}$$

$$= \frac{24}{49}$$

b) $P(3 \leq X \leq 6) = P(X=3) + P(X=4) + P(X=5) + P(X=6)$

$$= \frac{7}{49} + \frac{9}{49} + \frac{11}{49} + \frac{13}{49}$$

$$= \frac{40}{49}$$

* A discrete random variable has the probability function as follows. (6)

X	0	1	2	3	4	5	6	7
P(X)	0	k	$2k$	$2k$	$3k$	$3k^2$	$2k^2$	$7k^2+k$

Find (a) k, (b) $P(X<3)$ (c) $P(2<x\leq 5)$

Sol^n

(a) For a PDF $P(X)=1$

i.e, $P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)+P(X=6)+P(X=7)=1$

or, $0+k+2k+2k+3k+3k^2+2k^2+7k^2+k=1$

$\Rightarrow 12k^2+9k-1=0$

$\Rightarrow k=\dfrac{-9\pm\sqrt{129}}{24}$

$k=0.98$

Hence the table takes the form,

X	0	1	2	3	4	5	6	7
P(X)	0	0.98	1.96	1.96	2.94	2.8812	1.9208	7.7028

(b) $P(X<3) = P(X=0)+P(X=1)+P(X=2)$

$= 0+0.98+1.96$

$= 2.94$

(c) $P(2<x\leq 5) = P(X=3)+P(X=4)+P(X=5)$

$= 1.96+2.94+2.8812$

$= 7.7812$

: Two cards are drawn randomly, simultaneously from a well shuffled deck of 52 cards. Find the variance for the no of aces.

Sol^n Sample sapace = 52 cards.

X : drawing aces ~~together~~.

$X = \{0, 1, 2\}$

$$P(X=0) = \frac{^{48}C_2 \times {^4C_0}}{^{52}C_2} = \frac{1128}{1326}$$

no aces

$$P(X=1) = \frac{^{48}C_1 \times {^4C_1}}{^{52}C_2} = \frac{192}{1326}$$

1 ace

$$P(X=2) = \frac{^{48}C_0 \times {^4C_2}}{^{52}C_2} = \frac{6}{1326}$$

2 aces

X	0	1	2
P(X)	$\frac{1128}{1326}$	$\frac{192}{1326}$	$\frac{6}{1326}$

variance $\sigma^2 = \Sigma p_i x_i^2 - \mu^2$

$$= \Sigma p_i x_i^2 - (\Sigma p_i x_i)^2$$

$$= \left[\frac{1128}{1326} \times 0^2 + \frac{192}{1326} \times 1^2 + \frac{6}{1326} \times 2^2\right] - \left[\frac{1128}{1326} \times 0 + \frac{192}{1326} \times 1 + \frac{6}{1326} \times 2\right]^2$$

$$= \frac{216}{1326} - \left(\frac{204}{1326}\right)^2 = 0.1392$$

* 5 defective bulbs are accidently mixed with (7)
twenty good ones. It is not possible to just look at a bulb & tell whether or not it is defective. Find the mean of the number of defective bulbs if 4 bulbs are drawn at random from this lot.

Sol^n:-

X : no of defective bulbs.

$X = \{0, 1, 2, 3, 4\}$.

Sample space = 5 defective bulbs + 20 good bulbs
= 25 bulbs.

$$P(X=0) = \frac{{}^{20}C_4 \times {}^{5}C_0}{{}^{25}C_4} = \frac{969}{2530}$$

4 nondefective + 0 defective

$$P(X=1) = \frac{{}^{20}C_3 \times {}^{5}C_1}{{}^{25}C_4} = \frac{1140}{2530}$$

3 non defective + 1 defective

$$P(X=2) = \frac{{}^{20}C_2 \times {}^{5}C_2}{{}^{25}C_4} = \frac{380}{2530}$$

2 non defective + 2 defective

$$P(X=3) = \frac{{}^{20}C_1 \times {}^{5}C_3}{{}^{25}C_4} = \frac{40}{2530}$$

1 non defective + 3 defective

$$P(X=4) = \frac{{}^{20}C_0 \times {}^{5}C_4}{{}^{25}C_4} = \frac{1}{2530}$$

0 non defective + 4 defective

$$\text{mean } \mu = \Sigma p_i x_i = \frac{969}{2530} \times 0 + \frac{1140}{2540} \times 1 + \frac{380}{2530} \times 2 + \frac{40}{2530} \times 3 + \frac{1}{2530} \times 4.$$

2024 — 0.8

If X is a discrete random variable taking value $1, 2, 3, \ldots$ with $p(x) = \frac{1}{2}\left(\frac{2}{3}\right)^x$.

Find P(X being an odd no) by first establishing that $P(x)$ is a probability function.

$$\Sigma P(X = x_i) = \sum_{x=1}^{\infty} \frac{1}{2}\left(\frac{2}{3}\right)^x$$

$$= \frac{1}{2}\left[\frac{2}{3} + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^3 + \ldots\right]$$

is a G.P with $a = \frac{2}{3}$, $r = \frac{2}{3}$

$$= \frac{1}{2}\left[\frac{\frac{2}{3}}{1 - \frac{2}{3}}\right]$$

$$= 1$$

$\therefore$ $p(x)$ is a probability function.

$$P(X = \text{x being an odd no}) = \sum_{x=1,3,5} \frac{1}{2}\left(\frac{2}{3}\right)^x$$

$$= \frac{1}{2}\left[\frac{2}{3} + \left(\frac{2}{3}\right)^3 + \left(\frac{2}{3}\right)^5 + \ldots\right]$$

is a G.P with $a = \frac{2}{3}$, $r = \left(\frac{2}{3}\right)^2$

$$= \frac{1}{2}\left[\frac{\frac{2}{3}}{1 - \left(\frac{2}{3}\right)^2}\right]$$

$$= \frac{3}{5}$$

A Continuous Random Variables is one which takes an infinite number of possible values. Continuous random variables are usually measurements. Examples include, height, weight, the amount of sugar in an orange, the time required to run a mile.

A continuous random variable is not defined at specific values. Instead, it is defined over an interval of values, and is represented by the area under a curve. The probability of observing any single value is equal to 0, since the number of values which may be assumed by the random variable is infinite.

Suppose a random variable X may take all values over an interval of real numbers. Then the probability that X is in the set of outcomes A, P(A) is defined to be the area above A and under a curve. The curve, which represents a function p(x), must satisfy the following:

(i) the curve has no negative values ($p(x) \geq 0, \forall x$)

(ii) the total area under the curve is equal to 1.

A curve meeting these requirements is known as a density curve.

Continuous random variable is a random variable which takes an uncountably infine number of possible value or which can take any value in some interval.

example:
Rainfall in a particular area

Probability density function of a continuous random variable X with sample space $S=(-\infty,\infty)$ is an integrable function $f(x)$ satisfying the following:

(i) $f(x)$ is positive everywhere in $S=(-\infty,\infty)$
i.e., $f(x) > 0$ for all x in $S=(-\infty,\infty)$

(ii) The area under the curve $f(x)$ in $S_{(-\infty,\infty)}$ is 1.
i.e., $\int_{-\infty}^{\infty} f(x)\,dx = 1$ [or $\int_S f(x)\,dx = 1$]

If $f(x)$ is the p.d.f (probability density function) of x, then the probability that x belongs to A, where A is some interval (a,b), is given by the integral of $f(x)$ over that interval,
i.e., $P(X \in A) = \int_A f(x)\,dx = \int_a^b f(x)\,dx$

Cumulative density function (Cumulative distribution ~~function~~ function) of a continuous random variable.

(is defined as

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \quad \text{for } -\infty < x < \infty$$

Mean, Variance, Standard deviation

The expected value or mean of a continuous random variable X is: $\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$

The variance of a continuous random variable X is

$$Var(X) = \sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

Alternatively, you can still use the shortcut formula for the variance, $\sigma^2 = E(X^2) - \mu^2$ with $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$.

The Standard deviation of a continuous random variable X is:

$$\sigma = \sqrt{Var(X)}$$

* Verify whether $f(x) = \begin{cases} 6x(1-x) & 0<x<1 \\ 0 & \text{elsewhere} \end{cases}$

is a probability density function and hence find $P(\frac{2}{3}<x<1)$

Sol^n:

(i) $f(x) \geq 0$ in the given interval.

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

$$= 0 + \int_{0}^{1} (6x-6x^2)\,dx + 0$$

$$= \left.\frac{6x^2}{2} - \frac{6x^3}{3}\right]_0^1$$

$$= 1$$

Hence verified.

$$P\left(\frac{2}{3}<x<1\right) = \int_{2/3}^{1} f(x)\,dx$$

$$= \int_{2/3}^{1} (6x-6x^2)\,dx$$

$$= \left.\frac{6x^2}{2} - \frac{6x^3}{3}\right]_{2/3}^{1}$$

$$= 3-2-\frac{4}{3}+\frac{16}{27} = \frac{7}{27}$$

A continuous random variable has the density function

$$f(x) = \begin{cases} k x^2 & -3 < x < 3 \\ 0 & \text{elsewhere.} \end{cases}$$

Find k and hence find $P(x<3)$, $P(x>1)$

Sol^n: For a density function, we have $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

i.e, $\int_{-\infty}^{-3} f(x)\,dx + \int_{-3}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx = 1$

i.e, $0 + \int_{-3}^{3} k x^2\,dx + 0 = 1$

ie, $\left.\frac{k x^3}{3}\right]_{-3}^{3} = 1 \Rightarrow k\left[\frac{27}{3} - \frac{(-27)}{3}\right] = 1 \Rightarrow k = \frac{1}{18}$

$P(x<3) = \int_{-\infty}^{3} f(x)\,dx = \cancel{\int_{-\infty}^{3} k x^2\,dx = \frac{k x^3}{3}}$

$= \int_{-\infty}^{-3} f(x)\,dx + \int_{-3}^{3} f(x)\,dx$

$= 0 + \int_{-3}^{3} k x^2\,dx$

$= \left.\frac{k x^3}{3}\right]_{-3}^{3}$

$= \frac{1}{18}\left[\frac{27}{3} - \frac{(-27)}{3}\right]$

$= 1$

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

$= \int_{1}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx$

$= \int_{1}^{3} k x^2\,dx + 0$

$= \left.k\frac{x^3}{3}\right]_{1}^{3}$

$= \frac{1}{18}\left[\frac{27}{3} - \frac{(-1)}{3}\right]$

$= \frac{14}{27}$

ction

(10)

* Verify that $f(x) = 3x^2$ is a probability density function for $0 < x < 1$

Sol^n: (i) $f(x) > 0$ for $0 < x < 1$.

(ii) $\int_0^1 3x^2\,dx = x^3\Big]_0^1 = 1 - 0 = 1$.

Hence verified.

* Let X be a continuous random variable whose probability desity function is: $f(x) = \frac{x^3}{4}$ for an interval $0 < x < c$. What is the value of the value of the constant that makes $f(x)$ a valid probability density function?

Sol^n $\cancel{(ii) \int_0^c \frac{x^3}{4}\,dx = \frac{1}{4} \times \frac{x^4}{4}\Big]_0^c}$

For $f(x)$ to be a p.d.f, we should have $\int_0^c f(x)\,dx = 1$

ie, $\int_0^c \frac{x^3}{4}\,dx = 1$

ie, $\frac{1}{4} \times \frac{x^4}{4}\Big]_0^c = 1$

$\Rightarrow \frac{c^4}{16} = 1$

$c^4 = 16$

$c = \pm 2$

but $c = -2$ is not possible

$\therefore c = 2$

Suppose X is a continuous random variable with the following probability density function:

$f(x) = 3x^2$ for $0<x<1$. Find the mean and variance of X.

Sol:

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

$$= \int_{-\infty}^{0} x.f(x)\,dx + \int_{0}^{1} x.f(x)dx + \int_{1}^{\infty} x.f(x)dx$$

$$= 0 + \int_{0}^{1} x \times 3x^2 dx + 0$$

$$= \int_{0}^{1} 3x^3 dx$$

$$= \left.\frac{3x^4}{4}\right]_0^1$$

$$\mu = \frac{3}{4}$$

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

$$= \int_{0}^{1} x^2 f(x)dx - \mu^2$$

$$= \int_{0}^{1} x^2 \times 3x^2 dx - \left(\frac{3}{4}\right)^2$$

$$= \int_{0}^{1} 3x^4 dx - \left(\frac{3}{4}\right)^2$$

$$\left.\frac{3x^5}{5}\right]_0^1 - \frac{9}{16}$$

$$\sigma^2 = \frac{3}{5} - \frac{9}{16}$$

$$\sigma^2 = \frac{3}{80}$$

* Find the mean & variance for

$$f(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

Sol^n

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

$$= \int_{-\infty}^{0} x\, f(x)\,dx + \int_{0}^{\infty} x\, f(x)\,dx$$

$$= 0 + \int_{0}^{\infty} 2x e^{-2x}\,dx$$

$$= 2\left[x \times \frac{e^{-2x}}{-2} - 1 \times \frac{e^{-2x}}{4}\right]_0^{\infty}$$

$$= 2\left[0 - 0 + 0 + \frac{1}{4}\right]$$

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

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

$$= \int_{-\infty}^{0} x^2 \cdot f(x)\,dx + \int_{0}^{\infty} x^2 \cdot f(x)\,dx - \mu^2$$

$$= 0 + \int_{0}^{\infty} x^2 \cdot 2e^{-2x}\,dx - \left(\frac{1}{2}\right)^2$$

$$= 2\left[x^2 \times \frac{e^{-2x}}{-2} - (2x) \times \frac{e^{-2x}}{4} + (2) \times \frac{e^{-2x}}{-8}\right]_0^{\infty} - \frac{1}{4}$$

$$\sigma^2 = 2 \times \frac{2}{+8} - \frac{1}{4}$$

$$= +\frac{1}{2} - \frac{1}{4}$$

$$\sigma^2 = \frac{1}{4}$$

* Find the mean and variance of the probability density function: $f(x) = e^{-|x|}$

Sol^n

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

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

$$= \int_{-\infty}^{0} x\, f(x)\,dx + \int_{0}^{\infty} x\, f(x)\,dx$$

$$= \int_{-\infty}^{0} x \times e^{x}\,dx + \int_{0}^{\infty} x \times e^{-x}\,dx$$

$$= \left[x \times e^{x} - 1 \times e^{x}\right]_{-\infty}^{0} + \left[x \times \frac{e^{-x}}{-1} - 1 \times \frac{e^{-x}}{1}\right]_{0}^{\infty}$$

$$= -1 + 1$$

$$\underline{\mu = 0}$$

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

$$= \int_{-\infty}^{0} x^2 f(x)\,dx + \int_{0}^{\infty} x^2 f(x)\,dx - 0$$

$$= \int_{-\infty}^{0} x^2 \times e^{x}\,dx + \int_{0}^{\infty} x^2 \times e^{-x}\,dx$$

$$= \left[x^2 \times e^{x} - 2x \times e^{x} + 2 \times e^{x}\right]_{-\infty}^{0}$$

$$+ \left[x^2 \times \frac{e^{-x}}{-1} - 2x \times \frac{e^{-x}}{1} + 2 \times \frac{e^{-x}}{-1}\right]_{0}^{\infty}$$

$$= 2 + 2$$

$$\underline{\sigma^2 = 4}$$

In a certain city, the daily consumption of electric power (in million kwatt/hr) is random variable X having the probability density function (12)

$$f(x) = \begin{cases} \frac{1}{9} x e^{-x/3} & x > 0 \\ 0 & x < 0 \end{cases}$$

If the cities power plant has a daily capacity of 12 million kw/hr, what is the probability that this power supply will be insufficient on any given day.

Sol^n:

The power supply will be insufficient on any given day means the power consumption is more than the supply. i.e, $P(X > 12)$

$$P(X > 12) = \int_{12}^{\infty} f(x)\,dx$$

$$= \int_{12}^{\infty} \frac{1}{9} x e^{-x/3}\,dx$$

$$= \frac{1}{9}\left[x \times \frac{e^{-x/3}}{\left(\frac{-1}{3}\right)} - 1 \times \frac{e^{-x/3}}{\left(\frac{1}{9}\right)}\right]_{12}^{\infty}$$

$$= \frac{1}{9}\left[36 e^{-4} + 9 e^{-4}\right]$$

$$= 5 e^{-4}$$

* The length of time (in minutes) that a certain lady speaks on telephone is found to be a random variable with probability function

$$f(x) = \begin{cases} A e^{-x/5} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

i) Find A

ii) Find the probability that she will speak on the phone (b) more than 10 min

c) less than 5 min

d) between 5 & 10 min.

Sol^n: (a) Given $f(x)$ is p.d.f $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx = 1 \Rightarrow 0 + \int_{0}^{\infty} A e^{-x/5}\,dx = 1$$

$$\Rightarrow A \left. \frac{e^{-x/5}}{\left(-\frac{1}{5}\right)} \right]_0^{\infty} = 1 \Rightarrow 5A = 1 \Rightarrow A = \frac{1}{5}$$

b) $$P(x > 10) = \int_{10}^{\infty} f(x)\,dx = \int_{10}^{\infty} \frac{1}{5} e^{-x/5}\,dx = \frac{1}{5} \times \left. \frac{e^{-x/5}}{-\frac{1}{5}} \right]_{10}^{\infty} = e^{-2} = 0.1353$$

c) $$P(x < 5) = \int_{-\infty}^{5} f(x)\,dx = \int_{0}^{5} \frac{1}{5} e^{-x/5}\,dx = \frac{1}{5} \left. \frac{e^{-x/5}}{-\frac{1}{5}} \right]_{0}^{5} = -e^{-1} + 1 = 0.6322$$

d) $$P(5 < x < 10) = \int_{5}^{10} f(x)\,dx = \int_{5}^{10} \frac{1}{5} e^{-x/5}\,dx = \frac{1}{5} \left. \frac{e^{-x/5}}{-1/5} \right]_{5}^{10} = -e^{-2} + e^{-1}$$

$$= 0.2325$$

The Monty Hall Game show - Problem.

In a TV Game show, a contestant selects one of three ~~color~~ doors; behind one of the doors there is a prize, and behind the other two there are no prizes. After the contestant selects a door, the game-show host opens one of the remaining doors, and reveals that there is no prize behind it. The host then asks the contestant whether they want to switch their choice to the other ~~to~~ unopened door, or stick to their original choice. Is it probabilistically advantageous for the contestants to swith doors, or is the probability of winning the prize the same whether they stick or switch? (Assume that the host selects a door to open, from those available, with equalty probability).

Sol^n: Let A, B, C correspond to the prize being behind the selected, opened and remaining door respectively, and let H_B denote the event that the host opens door B.

We want to compare $P(A|H_B)$ (Stick) with $P(C|H_B)$ (Switch).

Now $P(A) = P(B) = P(C) = \frac{1}{3}$,

given $P(H_B|A) = \frac{1}{2}$, $P(H_B|B) = 0$ & $P(H_B|C) = 1$.

By Baye's theorem,

$$P(A|H_B) = \frac{P(H_B|A)\,P(A)}{P(H_B)} = \frac{P(H_B|A)\,P(A)}{P(H_B|A)P(A) + P(H_B|B)\,P(B) + P(H_B|C)\,P(C)}$$

$$= \frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{3} + 0\times\frac{1}{3} + 1\times\frac{1}{3}} = \frac{1}{3}$$

Ill^{ly} $P(C|H_B) = \frac{P(H_B|C)\,P(C)}{P(H_B)} = \frac{1\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{3} + 0\times\frac{1}{3} + 1\times\frac{1}{3}} = \frac{2}{3}$

So it is advantageous to switch.

Bayes' theorem is also known as the formula for the probability of causes, i.e, probability of a particular (cause) B_r given that event A has happened (already).

$P(B_i)$ is 'a priori probability' known even before the experiment, $P(A|B_i)$ "likelihoods" and $P(B_i/A)$ 'posteriori probabilities' determined after the result of the experiment.