Permutation and Combination—

When solving the problems that requires the number of ways that the people or items or the object can be arranged must first distinguish if order is ~~implemented~~ important.

a) If order dosen't matter it is combination (^nC_r)

b) If the order does matter then it is called as Permutation (^nP_r)

Example— Say we are talking about the no. of ways that the numbers 1, 2, 3 can be arranged— Combination

In how many ways can five runners can finish in a race.

⇒ There is 1st place, 2nd place.....etc the order in which the race is finished is important here hence it is permutation.

In how many ways can three comitee members be taken from a pool of 50 members.

⇒ The role of each member is same and dosent matter if it is first second or third. Since all the three members have same role. then it is combination.

In how many ways can a president and vice president from 20 people—
⇒ the roles of are not same the order is important ~~thing~~ therefore it is permutation.

How many ways can 5 cards be that for a poker hand from a desk of 52 cards.
⇒ the order in which they are received is not important therefore it is combination.

Counting—

Q Say that we have 3 pens and 2 marker pens.

a) Pick any one item from the above
⇒ there are five ways (3+2) to pickup anyone item from the five items.

b) Pick one pen and one marker.

$P_1 < \begin{smallmatrix} M_1 \\ M_2 \end{smallmatrix}$, $P_2 < \begin{smallmatrix} M_1 \\ M_2 \end{smallmatrix}$, $P_3 < \begin{smallmatrix} M_1 \\ M_2 \end{smallmatrix}$

Probability—

The random experiment happening under uncertain situation is called a random trial or random experiment or experiment of chance.

For ex— throwing of dice, tossing a unbiased coin or picking up a card from the pack of 52 cards.

Event – any question that we ask with regard to a random experiment.

Elementary event – Each outcome of the random experiment is called elementary outcome.
For ex – If coin is tossed head or tail is an elementary event.

Sample space – Sample space of a random experiment is the set of all possible outcomes.
For ex – If the coin is tossed then the sample space is {H, T} or in the case of die {1, 2, 3, 4, 5, 6}.

Let E denotes the random experiment and all possible outcomes then the sample space of E is $\{e_1, e_2, e_3 \dots e_n\}$

Example – If the two unbiased coins are tossed simultaneously

$S = \{HH, HT, TH, TT\}$, $n(S) = 4$.

when two random experiment having m outcomes say $e_1, e_2 \dots e_m$ and n outcomes say $p_1, p_2, \dots p_n$ respectively are conducted simultaneously then the sample space consist of mn elementary events

$$S = \{(e_1, p_1)(e_2, p_2) \dots (e_m, p_n)\}$$

$$S = \{(e_i, p_j) \mid i = 1, 2 \dots m,\ j = 1, 2 \dots n\}$$

Scanned by CamScanner

Universal event (Sure event)

Sample space S is collection of all the events. Since we know that A set is a subset of itself therlfore the sample space S can also be considered as an event with the experiment. Since every outcome belongs to and therefore the events are occurs always this event is called sure event or Universal event.

Impossible event—

Since empty set is subset of each set the ϕ set can be considered as representing a Impossible event. but there is no outcome of the experiment. Hence event represented by ϕ is called impossible event.

For example— throwing two dice simultaneously. Let the event B we defined as the sum of the no. of the faces is greater than or equals to two.

The event A is the sum of the no. greater than 12.

Scanned by CamScanner

Equally likely events –

Let S be a sample space of a random of variable if all the elementary events of S have the same chance of occurring then the events are said to be equally likely events.

Mutually exclusive events

Say that the event A and B are mutually exclusive if one event occur and other event cannot occur or two events cannot occur together.

Example – Throwing two dice simultaneously and note down sum of the two no's.

Event A ⇒ the sum of the two no's appearing on the dice is less than equals to 4. (≤ 4)

Event B ⇒ the sum of the two no's appearing on the dice is greater than or equal to 9. (≥ 9)

$A = \{2, 3, 4\}$,

$B = \{9, 10, 11, 12\}$

$A \cap B = \{\phi\}$.

Mutually exhaustive events —

Let S be a sample space of random experiment and $A_1, A_2 \ldots A_m$ be the events defined on the sample space of $A_1 \cup A_2 \cup \ldots A_m = S$ thus events are exhaustive if further $A_i \cap A_j = \phi$ for $i \neq j$ then the events are said to be mutually exclusive and exhaustive.

Combination of events —

The combination of events can be done by using the options and, or, NOT.

And $\Rightarrow \cup$, OR $\Rightarrow \cap$ Not $= \overline{A}$

or $\overline{A} = S - A$

Probability—

The ratio of favourable outcomes to the total possible outcomes is called probability.

Let S be the sample space and E be the event or some A be the event related to some space.

$$\text{Probability} = \frac{\text{The no. of favourable outcomes to } A}{\text{The total no. of events in the sample space}}$$

Axioms of probability—

Let S be a sample space and event A and event B be two mutually exclusive events then—

$0 \le P(A) \le 1$, $P(S) = 1$, $P(A \cup B) = P(A) + P(B)$

Example— What is the probability that a number selected from the numbers 1 to 20 is an even number when each of the given numbers is equally likely to be selected.

$S = \{1, 2, 3 \ldots\ldots 20\}$

$A = \{2, 4, 6, 8, \ldots\ldots 20\}$

$n(A) = 10$, $n(S) = 20$, $P(A) = \frac{10}{20} = \frac{1}{2}$

If $A_1, A_2, A_3 \ldots\ldots A_K$ are K mutually exclusive events then $(A_1 \cap A_2 \ldots \cap A_n = \emptyset)$

then $P(A_1 \cup A_2 \cup A_3 \ldots\ldots \cup A_K) = P\left(\bigcup_{i=1}^{k} A_i\right)$

$= P(A_1) + P(A_2) + \ldots\ldots + P(A_K)$

If $A_1, A_2, \ldots A_k$ are mutually exclusive or exhaustive events

$\bigcup_{i=1}^{k} A_i = S$ then $P(A_1) + P(A_2) + \ldots P(A_k)$

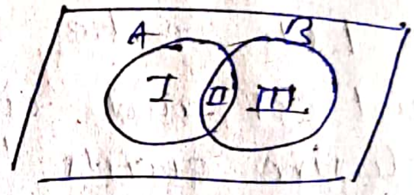
$= P(S) = 1$

theorem 1 (Law of Addition of probability)

If A and B be any two events associated with random experiment then the probability of E –

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$A \cup B = \underline{I + II + III}$



$P(I) = P(A) - P(A \cap B)$

$P(II) = P(A \cap B)$

$P(III) = P(B) - P(A \cap B)$

$P(A \cup B) = P(I) + P(II) + P(III)$

$= P(A) + P(B) - P(A \cap B)$ —— (*)

Note → If A and B are any mutually exclusive events then $P(A \cap B) = 0$

for example eqn (*) becomes

$$P(A \cup B) = P(A) + P(B)$$

Note If A, B, C are any events then

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

Note If A is any event associate with the random experiment then $P(\bar{A}) = 1 - P(A)$

Note If A is contained in B or A is subset of B then. $(A \subset B)$ $P(A) \leq P(B)$ i.e. then the happening of B is more likely then A.

Ex Two unbiased dice are tossed once find the probability of getting an even no. on the first dice or a total of eight.

~~$\frac{18}{36}$ ~~ $\frac{5}{36}$~~ let us Define the two events

A = ~~Even~~ getting the even no. on first die

B = sum is 8.

$P(A \cup B) = ?$

$$A = \left\{ \begin{matrix} (2,1)(2,2)(2,3)(2,4)(2,5)(2,6) \\ (4,1)(4,2)(4,3)(4,4)(4,5)(4,6) \\ (6,1)(6,2)(6,3)(6,4)(6,5)(6,6) \end{matrix} \right\}$$

$36\overline{)200}$

$n(A) = 18$

$B = \{(2,6)(3,5)(4,4)(5,3)(6,2)\}$

$n(B) = 5$ | $A \cap B = \{(2,6)(4,4)(6,2)\}$

$n(S) = 36$ | $n(A \cap B) = 3$

$P(A) = \frac{18}{36}$, $P(B) = \frac{5}{36}$, $P(A \cap B) = \frac{3}{36}$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$= \frac{18}{36} + \frac{5}{36} - \frac{3}{36} \Rightarrow \frac{20}{36}$

Ex If A, B, C are mutually exclusive and exhaustive event associated with random experiment.

$P(B) = 0.6 . P(A)$

$P(C) = 0.2 . P(A)$

$P(A) + P(B) + P(C) = 1$

$1.8 P(A) = 1$

$P(A \cup B \cup C) = P(A) + P(B) + P(C) = 1$

$= P(A) + 0.6P(A) + 0.2P(A) = 1$

$P(A) = \frac{5}{9}$

Ex The probability that at least one of the event occurs $P(A \cup B) = 0.8$ the prob that both the events occur simultaneous is 0.25 the find $P(\bar{A}) + P(\bar{B})$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$0.8 = 1 - P(\bar{A}) + 1 - P(\bar{B}) - 0.25$

$P(\bar{A}) + P(\bar{B}) = 2 - 0.25 - 0.8$

$P(\bar{A}) + P(\bar{B}) = 0.95$

Ex Prove that Boole's inequality

$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P(A_i)$

$i=1$ $P(A_1) = P(A_1)$

$i=2$ $P[A_1 \cup A_2] \leq P(A_1) + P(A_2)$

$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

$\leq P(A_1) + P(A_2)$

we prove the booth inequality by using the mathematical induction.

Assume that inequality is true for $i = k$

$P\left[\bigcup_{i=1}^{k} A_i\right] \leq \sum_{i=1}^{k} P(A_i)$

to prove that $i = k+1$

$P\left[\bigcup_{i=1}^{k+1} A_i\right] = P\left[A_{k+1} \cup \left(\bigcup_{i=1}^{k} A_i\right)\right]$

$= P(A_{k+1}) + P\left(\bigcup_{i=1}^{k} A_i\right) - P\left[A_{k+1} \cap \left[\bigcup_{i=1}^{k} A_i\right]\right]$

$\boxed{= P\left(\bigcup_{i=1}^{k+1} A_i\right)}$ $\leq P\left(\bigcup_{i=1}^{k} A_i\right) + P(A_{k+1})$

$$\cancel{\leq P\left[\bigcup_{i=1}^{k} A_i\right] + P\left[A_{k+1}\right]}$$

Conditional probability —

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)}$$

$$P\left(\frac{B}{A}\right) = \frac{P(A \cap B)}{P(A)}$$

$$\left\{\begin{array}{l} A \cap B = B \cap A \\ A \cap \emptyset = \emptyset \\ A \cup \emptyset = A \\ A \cap \bar{A} = \emptyset \\ (\overline{A \cup B}) = \bar{A} \cap \bar{B} \\ A \cap (A \cup B) = A \\ A \cup (A \cap B) = A \end{array}\right.$$

Let the probability of $B \neq 0$, given an event B the probability that event A happens given that event B has occured. is defined and denoted by.

Ex – Ten cards numbered 1 to 10 are placed in a hat. they are mixed then one card is pulled at random If the card is even number card. what is probability that it is divisible by 3.

$A = \{3, 6, 9\}$, $B = \{2, 4, 6, 8, 10\}$

$A \cap B = \{6\}$, $\boxed{P(A \cap B) = \frac{1}{10}}$

$\boxed{P(B) = \frac{5}{10}}$

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)} = \frac{1/10}{5/10} = \frac{1}{5}$$

Q A coin is tossed twice what is probability that heads appears in both tosses. given that heads appears ~~in~~ at least one of the tosses.

$$S = \{HH, HT, TH, TT\}$$

$$B = \{HH, HT, TH\}, \quad A = \{HH\}$$

$$\therefore A \cap B = \{HH\}$$

$$P(A \cap B) = \frac{1}{4}, \qquad P(B) = \frac{3}{4}$$

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)} = \frac{1/4}{3/4} = \frac{1}{3}$$

Q For any two events A and B probability of A is 0.5, B = 0.6 $P(A \cup B) = 0.8$ then find ~~A~~ $P\left(\frac{A}{B}\right)$ and $P\left(\frac{B}{A}\right)$

$$P(A \cap B) = 0.5 + 0.6 - 0.8$$

$$\therefore P(A \cap B) = 0.3$$

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.6} \Rightarrow \frac{1}{2}$$

$$P\left(\frac{B}{A}\right) = \frac{0.3}{0.5} = \frac{3}{5}$$

Q A die is thrown twice and the sum of the no. appearing is noted to be 8. what is the conditional probability that the no. is 5 is appeared at least once.

$A = \{(2,5)(1,5)(3,5)(4,5) \ldots \}$

$n(A) = 11$

$B = \{(3,5)(5,3)(6,2)(2,6)(4,4)\}$

$n(B) = 5$

$n(A \cap B) = 2$

$$\boxed{P\left(\frac{A}{B}\right) = \frac{2}{5}}$$

Q Prove that $P\left(\frac{\bar{A}}{\bar{B}}\right) = \frac{1-P(A \cup B)}{P(\bar{B})}$ if $P(\bar{B}) \neq 0$.

$$P\left(\frac{\bar{A}}{\bar{B}}\right) = \frac{P(\bar{A} \cap \bar{B})}{P(\bar{B})}$$

$$= \frac{P(\overline{A \cup B})}{P(\bar{B})}$$

$$= \frac{1 - P(A \cup B)}{P(\bar{B})}$$

Independent Events —

Two events A and B are said to be independent if the information that one of them occurs doesn't change the probability of occurrence of the other event i.e. that the prob. of event A doesn't depend upon occurrence and non-occurrence of event B.

Scanned by CamScanner

→ The events A and B are said to be independent events.

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P\left(\frac{A}{B}\right) = P(A)$$

$$P\left(\frac{B}{A}\right) = P(B)$$

→ If given that A and B are independent event. then show that ~~A and~~ $P(\bar{A} \cap \bar{B})$ is also independe

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$$

$$= 1 - P(A \cup B)$$

$$= 1 - [P(A) + P(B) - P(A \cap B)]$$

$$= 1 - P(A) - P(B) + P(A \cap B)$$

$$= 1 - P(A) - P(B) + P(A)P(B)$$

$$= 1 - P(A) - P(B)[1 - P(A)]$$

$$P(\bar{A} \cap \bar{B}) = [1 - P(A)][1 - P(B)]$$

Note – For the independency of 3 events A_1, A_2, A_3, we must have $P(A_i \cap A_j) = P(A_i)P(A_j)$ and $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ which means ~~favourab~~ pairwise independence of events A_1, A_2, A_3 doesn't imply independence of 3 events. i.e. if A_1, A_2, A_3 are pairwise independence –

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$
$$P(A_2 \cap A_3) = P(A_2) \cdot P(A_3)$$
$$P(A_3 \cap A_1) = P(A_3) \cdot P(A_1)$$

doesnt imply —

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

Q Given that A and B are independent events examine if the events.

a) $\bar{A} \cap B$

b) $A \cap \bar{B}$ or Independent.

a ⇒ $\bar{A} \cap B$

$$B = (A \cap B) \cup (\bar{A} \cap B)$$

$$P(B) = P(A \cap B) + P(\bar{A} \cap B)$$ {it is mutually exclusive

$$P(\bar{A} \cap B) = P(B) - P(A \cap B)$$
$$= P(B) - P(A)P(B)$$
$$= P(B)[1 - P(A)]$$
$$= P(\bar{A}) \cdot P(B)$$

You must write that as Ans

Ans ⇒ $A \cap B$ and $A \cap \bar{B}$ are mutually exclusive and $(A \cap B) \cup (A \cap \bar{B}) = P(\bar{A}) \cdot P(B)$

Q Roll two unbiased dies one red and one blue and consider the events

a) the red die lands on 4.

b) the sum on the dice is 9.

c) the blue die lands on an odd number.

determine which pair of events are independent.

Q Events E and F are indepent events find probability of P(F) if its is given that $P(E) = 0.4$ $P(E \cup F) = 0.55$. then find $P(E \cap F)$

$$P(E \cap F) = P(E) \cdot P(F)$$

$\Rightarrow$ $P(E \cup F) = P(E) + P(F) - P(E) \cdot P(F)$

$$0.55 = 0.4 + P(F) - 0.4P(F)$$

$$P(F) = 0.25$$

Q A and B are two independent events given that $P(A) = 0.4$ and $P(A \cup \bar{B}) = 0.7$ then find probability of B or P(B) ?

$\Rightarrow$ $P(A \cup \bar{B}) = P(A) + P(\bar{B}) - P(A) \cdot P(\bar{B})$

$$0.7 = 0.4 + P(\bar{B}) - 0.4 \cdot P(\bar{B})$$

$$0.6P(\bar{B}) = 0.3$$

$$P(\bar{B}) = \frac{1}{2}$$

Q Given that $P(A \cap \bar{B}) = \frac{1}{4}$ and $P(A \cup B) = \frac{3}{4}$ find P(A) and P(B)

$\Rightarrow$ $P(A \cap \bar{B}) = P(A) - P(A \cap B)$

Q The probability that a ~~stu~~ student P is fails in an examination is 0.25 and the ~~probability~~ same student Q fails in the same exam is
the probability that either P & Q are failed in the examination.

Baye's theorem —

If $E_1, E_2 \ldots E_n$ are mutually exclusive events of universal set S and $X \subseteq \bigcup_{i=1}^{n} E_i$, then,

$$P\left(\frac{E_i}{X}\right) = \frac{P(X/E_i)\, P(E_i)}{\sum_{i=1}^{n} P(X/E_i)\, P(E_i)}$$

Proof — $X \subseteq \bigcup_{i=1}^{n} E_i$

$$X = \bigcup_{i=1}^{n} (X \cap E_i)$$

$$P(X) = \sum_{i=1}^{n} P(X \cap E_i)$$

$$P(E_i/X) = \frac{P(E_i \cap X)}{P(X)}$$

$$= \frac{P(X/E_i)\, P(E_i)}{\sum_{i=1}^{n} P(X \cap E_i)}$$

$$= \frac{P(X/E_i)\, P(E_i)}{\Sigma}$$

Q A doctor has decided to prescribe two drugs A and B to the 200 patients. 50 patient get drug A and 50 patient get drug B and rest of them get both the drugs A and B. The 200 patient were choosen so that each had a chance of 80% getting heart attack if no drug is given. Drug A reduces the probability of getting heart attack by 35% drug B reduces the probability of getting heart attack by 20% and the two drugs when given work independently.

If a randomly selected patient has a heart attack what is the probability that the patient were given drug A, drug B or both the drugs.

Sol 200 Patients

Let E_1 event that a patient given drug A

E_2 - - - - B

E_3 - - - - both the drug

$P(E_1) = \frac{50}{200} = \frac{1}{4}$,

$P(E_2) = \frac{50}{200} = \frac{1}{4}$,

$P(E_3) = \frac{100}{200} = \frac{1}{2}$, Let H, be the event that the patient gets heart attack.

$P\left(\frac{H}{E_1}\right) = 0.8 \times 0.65 =$

$P\left(\frac{H}{E_2}\right) = 0.8 \times 0.8 =$

$P\left(\frac{H}{E_3}\right) = 0.8 \times 0.65 \times 0.8 =$

then,

$$P\left(\frac{E_1}{H}\right) = \frac{P(H/E_1)\, P(E_1)}{P\left(\frac{H}{E_1}\right) P(E_1) + P\left(\frac{H}{E_2}\right) P(E_2) + P\left(\frac{H}{E_3}\right) P(E_3)}$$

$$P\left(\frac{E_2}{H}\right) = \frac{P(H/E_2)\, P(E_2)}{\text{— " —}}$$

$$P\left(\frac{E_3}{H}\right) = \frac{P(H/E_3)\, P(E_3)}{\text{— " —}}$$

Q The completion of an construction project depends on whether the carpenters and plumbers working on the projects will on strike. The probabilities of delay in the completion of project are 100%, 80%, 40%, and 5% if both go on strike, carpenters alone go on strice, plumbers alone go on strike. and neither of them striks respetively also there is 60% chance that plumber ~~carpenters~~ strike if carpenter go on strike and if plumbers go on strike there is 80%

chance that carpenters would follow. It is known that the chance of plumbers strike is 10%.

a) Determine the probability of delay in the completion of the project.

b) If there is a delay in the completion of the project determine.

1) Probabity that both carpentors and plumbers strike.

2) The probability that the carpenters strike and plumbers do not strike.

3) Probability of carpenters strike.

Sol Let, P = event that plumbers strike

C = event that carpenter strikes

D = event that there is delay in completion of project.

$P\left(\frac{D}{PC}\right) = 100\% = 1$

$P\left(\frac{D}{\bar{P}C}\right) = 80\% = 0.8$

$P\left(\frac{D}{P\bar{C}}\right) = 40\% = 0.4$

$P\left(\frac{D}{\bar{P}\bar{C}}\right) = 5\% = 0.05$

$P\left(\frac{P}{C}\right) = 0.6$

$P\left(\frac{\bar{P}}{C}\right) = 0.4$

$P\left(\frac{C}{P}\right) = 0.3$

$P\left(\frac{\bar{C}}{P}\right) = 0.7$

$P(P) = 0.1$

Scanned by CamScanner

a)

$$P(D) = P(D \cap PC) + P(D \cap \bar{P}C) + P(D \cap P\bar{C}) + P(D \cap \bar{P}\bar{C})$$

$$= P\left(\frac{D}{PC}\right) P(PC) + P\left(\frac{D}{\bar{P}C}\right) P(\bar{P}C) + P\left(\frac{D}{P\bar{C}}\right) P(P\bar{C}) + P\left(\frac{D}{\bar{P}\bar{C}}\right) P(\bar{P}\bar{C})$$

$$\because P(PC) = P\left(\frac{C}{P}\right) \cdot P(P)$$

$$= P\left(\frac{P}{C}\right) \cdot P(C) \times$$

$$\because P(\bar{P}C) = P(\bar{P}/C)\ P(C)$$

$$\cancel{P(C/\bar{P})\ P(\bar{P})} \times$$

$$\because P(P\bar{C}) = \underline{P(\bar{C}/P)\ P(P)}$$

$$\cancel{P(P/\bar{C})\ P(\bar{C})} \times$$

$$\because P(\bar{P}\bar{C}) = P(\bar{P} \cap \bar{C})$$

$$= P(\overline{P \cup C})$$

$$= 1 - P(P \cup C)$$

$$= 1 - [P(P) + P(C) - P(PC)]$$

b) 1 $$P\left(\frac{PC}{D}\right) = \frac{P(D/PC)\ P(PC)}{P(D)}$$

b) 2 $$P\left(\bar{P}C/D\right) = \frac{P\left(\frac{D}{\bar{P}C}\right) P(\bar{P}C)}{P(D)}$$

b) 3 $$P\left(\frac{C}{D}\right) = P(CP \cup C\bar{P} / D)$$

$$= P\left(\frac{CP}{D}\right) + P\left(\frac{C\bar{P}}{D}\right)$$

Scanned by CamScanner

Q The two players namely A and B participate in a game of throwing two dice the first player who gets ~~some of~~ sum of 7 is awarded the price if A starts the game find the probabilities of they winning the game.

Sol $S = \{(1,6)(2,5)(3,4)(4,3)(5,2)(6,1)\}$

$$P(7) = \frac{6}{36}$$

$$P(\bar{7}) = \frac{30}{36}$$

$$p + q^2p + q^4p + \ldots$$

$$p[1 + q^2 + q^4 + \ldots]$$

$$p\left(\frac{1}{1-q^2}\right) \Rightarrow \frac{6}{11} \text{ wining A}$$

$$\frac{5}{11} \text{ wining B}$$

Q Bag I contains 3 Red Balls 4 Blacks while Bag II contains 5 Red 6 Black Balls. One ball is drawn at random from one of the bags and it is found to be red. find the probability that it was drawn from bag 2.

$E_1 = 7$ $E_2 = 11$

$P\left(\frac{R}{E_1}\right) = \frac{3}{7}$, $P\left(\frac{B}{E_1}\right) = \frac{4}{7}$ | $P\left(\frac{R}{E_2}\right) = \frac{5}{11}$, $P\left(\frac{B}{E_2}\right) = \frac{6}{11}$

$P(E_1) = \frac{1}{2}$ | $P(E_2) = \frac{1}{2}$

$$P\left(\frac{E_2}{R}\right) = \frac{P\left(\frac{R}{E_2}\right)P(E_2)}{P\left(\frac{R}{E_1}\right)P(E_1) + P\left(\frac{R}{E_2}\right)\cdot P(E_2)}$$

$$= \frac{5/11 \times \frac{1}{2}}{\frac{5}{11}\times\frac{1}{2} + \frac{3}{7}\times\frac{1}{2}} = \frac{35}{68}$$

Q Let A and B two possible outcomes of a experiment suppose $P(A)=0.4$, $P(A \cup B)=0.7$ and $P(B)=p$. ① for what choice of P are A and B are mutually exclusive.

② For what choice of P, A and B are independent.

① $P(A \cup B) = P(A) + P(B) - \underline{P(A \cap B)}$

$0.7 = 0.4 + p - \underline{0}$

$p = 0.3$

② $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$0.7 = 0.4 + p - (0.4)p$

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

Q An Urn contain four tickets marked with numbers. (112, 121, 211, 222) and one ticket is drawn at random let A_i ($i=1,2,3$) be the event that ith digit of the number of the ticket drawn is 1. Discuss the independency of event A_1, A_2, A_3.

Random Variable (r.v)

Random variable means a real no. associated with the outcome of random experiment it can take any one of the various possible values each with a definite probability.

For ex.– Consider a random experiment three tosses of a coin.

$S = \{(HHH), (HHT)(HTH)(THH)(HTT)$
$(THT)(TTH)(TTT)\}$

Let ~~the~~ us consider the random variable x. which is the no. of heads obtain then X is a random variable which can take any one of the values, 0, 1, 2, 3

Outcome	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
Variable of X	3	2	2	2	1	1	1	0

Therefor the Random variable is a function which takes real values which are determined by the outcomes of the random experiment.

Discrete and Continuous Random Variable

If the random variable X assumes only a finite number or countably infinite sets of values is known as discrete random variables.

For. Ex – the no. of student in the MCA 1st.

If the random variable X assumes infinite or uncountable sets of values it is said to be continuous random variable.

For example. weight of a person.

In case of continuous random variable we usually takes the value in a particular interval not a point.

#

Probability Distribution Function Descrete Variable

Let us consider a descrete Random variable X which can takes the real values $x_1, x_2 \ldots x_n$ with each value of the variable x we associate a number.

$p(x = x_i) = p_i$, $i = 1, 2, 3 \ldots n$

which is known as the probability of x_i satisfying the following conditions

a) $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$

Probability Distribution Function Continuous Random Variable

Unlike the descrete probability distribution function of the descrete variable a continuous distr function cannot be present in a tabular form.

Probability density Function — (PDF)

Let X may a continuous random variable taking the values in the interval $[a, b]$ a function $p(x)$ is said to be probability density function if it satisfies the folln condn —

(i) $p(x) \geq 0$. $x \in [a, b]$

(ii) The total area under the probability curve is 1.

$$p(a \leq x \leq b) = 1$$

(iii) For two distinct numbers c and d in the interval (a, b) probability of

$p(c \leq x \leq d)$ = Area under the probabity curve b/w the ordinal $x = c$ and $x = d$.

Q State with reasons if the following distributions function are admissiable or not.

Not (1)

x	0	1	2
$p(x)$	0.3	0.2	0.5

Not (2)

x	-1	0	1
$p(x)$	0.4	0.4	0.3

Not (3)

x	-2	-1	0	1	2
$p(x)$	0.3	0.4	-0.2	0.2	0.3

Q A die is tossed twice getting an odd number is termed as success find the probability distribution of the no. of success.

→ $E = \{1, 3, 5\}$, $P(S) = \frac{3}{6}$, $P(F) = \frac{3}{6}$

X denotes the no. of success in two throws of die. then $X = \{0, 1, 2\}$

$P(x=0)$ = P (Failure in first throw & and Failure in second throw)

$= P(FF) \Rightarrow P(F) \cdot P(F)$

$= \frac{1}{2} \times \frac{1}{2} \Rightarrow \frac{1}{4}$

$P(x=1)$ = the probability of success in first throw of dice and failure in second throw or vice versa.

$= P(SF) + P(FS) \Rightarrow P(S) \cdot P(F) + P(F) P(S)$

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

$P(x=2)$ = Probability of success in both throw.

$P(x=2) = P(S)(S)$

$= \frac{1}{2} \times \frac{1}{2} \Rightarrow \frac{1}{4}$

x	0	1	2
$p(x)$	1/4	1/2	1/4

Q Obtain the probability distribution of X the no of tosses of a 3 coin. and no. of Head on face. side

x	Favourable event	No. of events	$P(X=x)$
0	(TTT)	1	1/8
1	(TTH) (THT) (HTT)	3	3/8
2	(THH) (HTH) (HHT)	3	3/8
3	(HHH)	1	1/8

Q Two dice are rolled at random obtain the probability distribution of the sum of the numbers on them.

Sum of the no. (x)	Favourable outcome	No. of outcome	$F(X=x)$
2	(1 1)	1	1/36
3	(1,2) (2,1)	2	2/36
4		3	3/36
5		4	
6		5	
7		6	
8		5	
9		4	
10		3	
11		2	
12		1	

Q From a lot of 12 items containing 3 defective items a sample of 4 items are taken are drawn randomly without replacement. Let a random variable X denotes the no. of defective items. Find the prob. distr of X.

X	0	1	2	3
P(X=x)	$\frac{^9C_4}{^{12}C_4}$	$\frac{^9C_3 \times {^3C_1}}{^{12}C_4}$	$\frac{^9C_2 \times {^3C_2}}{^{12}C_4}$	$\frac{^9C_1 \times {^3C_3}}{^{12}C_4}$

Cumulitive Probability Distribution –

If X is a random variable with the probability function $P(x)$ then the distribution function is denoted by $F(x)$ and defined by $F(x) = P(X \leq x)$

If X takes integer values viz $0, 1, 2, 3 \ldots x$ then $F(x) = P(x=1) + P(x=2) \ldots P(x=x)$

$F(x) = P(1) + P(2) + P(3) + \cdots P(x)$.

$F(x-1) = P(1) + P(2) + P(3) \cdots P(x-1)$

$$\boxed{F(x) - F(x-1) = P(x)}$$

If X is a continuous random with probability density function $p(x)$ then the cumulative distribution function is given by – $F(x) = P(X \leq x)$

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

Q A random variable X takes —

x	-2	-1	0	1	2	3
$P(x)$	0.1	k	0.2	$2k$	0.3	$3k$

Find ① $k = ?$

$F(n) - F(n-1)$

② $P(x<2)$

③ $P(-2<x<2)$

④ Cumulative distr of X

① $0.1 + k + 0.2 + 2k + 0.3 + 3k = 6k + 0.6 = 1$

$6k = 0.4$

$k = \frac{1}{15}$

② $P(x<2)$

$P(x=-2) + P(x=-1) + P(x=0) + P(x=1)$

$= 0.1 + \frac{1}{15} + 0.2 + \frac{2}{15} = \frac{3+2}{10} = \frac{1}{2}$

③ $P(-2<x<2)$

$P(x=-1) + P(x=0)$

$\frac{1}{15} + \frac{0.2}{10}$

④

$f(-2) = 0.1$

Q A random variable X takes the value 1, 2, 3 ..... and probability of $p(X=x) = \frac{1}{2^n}$, n = 1, 2, 3 ... find.

(1) p(X is odd)

(2) $p(x \leq 5)$

(3) p(x is divisible by 5)

(1) $p(X=1) + p(X=3) + p(X=5) + \ldots$

$\frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \ldots$

$\frac{1}{2}\left[1 + \frac{1}{2^2} + \frac{1}{2^4} + \ldots\right]$

$\frac{1}{2} \times \left(\frac{1}{1-\frac{1}{2^2}}\right) \rightarrow$ $a = 1$, $r = \frac{1}{2^2}$

(2) $P(X \leq 5) = P(x=1) + P(x=2) + \ldots P(x=5)$

$= \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots \frac{1}{2^5}$

(3) $p(x=5) + p(x=10) + p(x=15) + \ldots$

$\frac{1}{2^5} + \frac{1}{2^{10}} + \frac{1}{2^{15}} + \ldots$

$\frac{1}{2^5}\left[1 + \frac{1}{2^5} + \frac{1}{2^{10}} + \ldots\right]$

$\frac{1}{2^5}\left[\frac{1}{1-\frac{1}{2^5}}\right]$ Ans $a = 1$, $r = \frac{1}{2^5}$

Scanned by CamScanner

Q obtain the probability distribution of the no. of sixes of two roles of a die.

Q Two die are thrown simultaneously and getting a number less than 3 on a die is termed as success. Obtain the prob. distribution of the no. of success.

doubt

Q Two cards are drawn the first case successively with replacement or simultaneously (without replacement) from a pack of well shuffled desk of 52 cards find the prob. ~~of~~ distr of Aces.

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

Continuous Random Variable (Probability density Function)

If X is a continuous random variable and a and b are real constant and $a \leq b$ then in the case to calculate

$$P(a \leq x \leq b) = P(a < x \leq b)$$
$$= P(a \leq x < b)$$
$$= P(a < x < b)$$

$$P(a \leq x \leq b) = \int_a^b f(x)\,dx$$

A function can serve as probability density function of continuous random variable x. If its values satisfy the following condition.

① $f(x) \geq 0, \quad -\infty < x < \infty$

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

Q① If 'x' is a prob. density function. say

$$f(x) = \begin{cases} ke^{-3x} & \text{if } x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

① then find the value of k.

② $P(0.5 < x < 1)$

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

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

$$0 + \int_{0}^{\infty} ke^{-3x}dx = 1$$

$$\frac{k}{(-3)}\left(e^{-3x}\right)_0^{\infty} = 1$$

$$\frac{-k}{3}[0-1] = 1$$

$$\boxed{k=3}$$

(ii) $P(0.5 \leq x < 1) = \int_{0.5}^{1} f(x)dx$

$$= \int_{0.5}^{1} 3.e^{-3x}dx$$

$$= 3\left[\frac{e^{-3}}{3} - \frac{e^{-1.5}}{-3}\right]$$

Q If X is a continuous random variable and the value of its probability density function at t is f(t), then the ~~prob~~ cumulative distribution function is given by.

$$F(x) = P(X \leq x)$$

$$= \int_{-\infty}^{x} f(t)dt$$

Scanned by CamScanner

Theorem – If $f(x)$ and $F(x)$ are the values of probability density and prob. distribution function of the continuous random variable X, then,

$$p(a \le X \le b) = F(b) - F(a)$$

For any real constant a and b with $a \le b$. $f(x) = \frac{dF(x)}{dx}$

Note → $\boxed{F(-\infty) = 0}$ and $\boxed{F(\infty) = 1}$

Q. Find the cumulative distribution function of the random variable X of the data in Que No – 1 (Previous Page)

$$f(x) = \begin{cases} 3e^{-3x}, & x = 0 \\ 0 & \text{elsewhere} \end{cases}$$

(i) $F(x) = P(X \le x) = \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{\infty} f(t)\,dt$

$$= 0 + \int_{0}^{x} 3e^{-3t}\,dt$$

$$F(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-3x}, & x > 0 \end{cases}$$

(ii) $P(0.5 < x \le 1)$

$= F(b) - F(a)$

$=$

Q Find the prob. density function for the X whose distribution function is given by —

$$F(x) = \begin{cases} 0 & x \le 0 \\ x & 0 < x < 1 \\ 1 & x \ge 1 \end{cases}$$

and plot the graph.

Sol

On differentiating

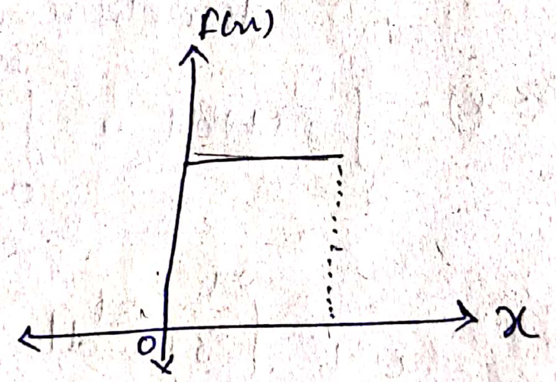
$$F(x) = \begin{cases} 0 & x \le 0 \\ 1 & 0 < x < 1 \\ 0 & x > 1 \end{cases}$$

or

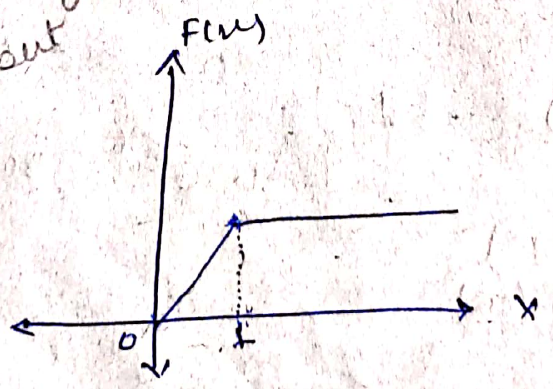
density funⁿ $f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$

Graph —

density



Distributⁿ



coutd.

Q the p.d.f of a random variable

1) is given by.

$$f(x) = \begin{cases} 1/5 & 2<x<7 \\ 0 & \text{elsewhere} \end{cases}$$

(i) Draw its graph and verify that the total area under the curve is 1

(ii) Find $P(3<x<5)$

2) $$f(y) = \begin{cases} 1/8(y+1) & \text{if } 2<y<4 \\ 0 & \text{e.w} \end{cases}$$

① $P(y<3.2)$ ② $P(2.9<Y<3.2)$

3) $$F(x) = \begin{cases} \frac{c}{\sqrt{x}} & 0<x<4 \\ 0 & \text{elsewhere} \end{cases}$$

(i) $c = ?$ (ii) $P(x<\frac{1}{4})$ and $P(x=1)$

Sol

2) (i) $P(y<3.2) = \int_{-\infty}^{3.2} F(x)\,dx$

$= \int_{2}^{3.2} \frac{1}{8}(y+1)\,dy =$

(ii) $P(2.9<y<3.2)$

$= \int_{2.9}^{3.2} F(x)\,dx \Rightarrow \int_{2.9}^{3.2} \frac{1}{8}(y+1)\,dy$

(3) (i) $\int_0^4 f(x)\,dx = 1$

$$\int_0^4 \frac{c}{\sqrt{x}}\,dx = 1 \Rightarrow c\int_0^4 x^{-1/2}\,dx = 1$$

$$c\left[\frac{x^{1/2}}{1/2}\right]_0^4 = 1 \Rightarrow 2c\left[4^{1/2}\right] = 1$$

$$\boxed{c = \frac{1}{4}}$$

(ii) $p\left(x<\frac{1}{4}\right)$ and $p(x>1)$

$$\int_{-\infty}^{1/4} f(x)\,dx \Rightarrow \int_{-\infty}^{1/4} \frac{1}{4\sqrt{x}} \Rightarrow \int_{-\infty}^{0} \frac{1}{4\sqrt{x}} + \int_0^{1/4} \frac{1}{4\sqrt{x}}$$

$$= 0 + \int_0^{1/4} \frac{1}{4\sqrt{x}}\,dx$$

$=$

$$p(x>1) \Rightarrow \int_1^{\infty} f(x)\,dx \Rightarrow \int_1^4 f(x)\,dx + \int_4^{\infty} f(x)\,dx$$

$$= \int_1^4 \frac{1}{4\sqrt{x}}\,dx$$

Q The distribution function of the random variable x is given by -

$$F(x) = \begin{cases} 0 & x < -1 \\ \frac{x+1}{2} & -1 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

(i) $P\left(-\frac{1}{2} < x < \frac{1}{2}\right)$

(ii) $P(2 < x < 3)$

Sol (i) $P\left(-\frac{1}{2} < x < \frac{1}{2}\right)$

$= F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right)$ $\{F(b) - F(a)\}$

$=$

$= 0.5$

(ii) $F(3) - F(2)$

$= 1 - 1$

$= 0$

Q Find the distribution function of random variable X whose density function is given by—

$$F(x) = \begin{cases} x & 0 < x < 1 \\ 2x & 1 \leq x \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

Also ~~the~~ plot the graph of prob. density function and distribution.

Scanned by CamScanner

Q The distribution function of random variable y is given by—

$$F(y) = \begin{cases} 1 - \frac{9}{y^2} & y > 3 \\ 0 & \text{elsewhere} \end{cases}$$

(i) $P(Y \le 5)$

(ii) $P(Y > 8)$

Q $$F(x) = \begin{cases} 1 - (1+x)e^{-x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

(i) $P(x \le 2)$

(ii) $P(1 < x < 3)$

(iii) $P(x > 4)$

Mathematical Expectation—

If X is a descrete random variable and $p(x)$ is the value of the probability distribution at X. The expected value of X is denoted and defined by—

$$E(X) = \sum_{i}^{n} x_i \, p(x_i)$$

Correspondingly if X is a continuous random variable and $f(x)$ is a value of its probability density at x. then expected value of $f(x)$ is—

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

Physical Interpretion of $P(x)$—

Let us consider the following freq. distr. of the random variable X.

x	x_1	x_2	x_3	.....	x_n
F	f_1	f_2	f_3		f_n

The mean of the distribution is given by.

$$\boxed{\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots\ldots f_n x_n}{N}}$$

out of total of N cases f_i cases are favourable to x_i of the total N cases.

Let,

$$\frac{f_i}{N} = P_i = P(X = x_i)$$

$$= x_1 \frac{f_1}{N} + x_2 \frac{f_2}{N} + x_3 \frac{f_3}{N} + \cdots x_n \frac{f_n}{N}$$

$$= x_1 P_1 + x_2 P_2 + x_3 P_3 + \cdots x_n P_n$$

$$= \sum_i x_i P_i$$

→ The mathematical expectation of random variable X is its arithmetic mean.

☆☆ #Note – The term expected value is unfortune ~~in~~ that it is not in sense a value which one expects to occur in a particular experiment.

#Note – If the mathematical expe. of the game of a player is zero which means the game is said to be fair. If E(x) of the game is greater than 1 then its said is biased to the player. and if E(x) is -ve the game is said to be biased against to the player.

Theorem—

1. $E(c) = c$

   expectation of a constant is always a constant.

2. $E(cX) = c\,E(X)$

3. $E(aX+b) = aE(X)+b$

   where a and b are two constant

Addition theorem—

If X and Y are two random variable then ~~which~~

$$\boxed{E(X+Y) = E(X)+E(Y)}$$

which means expected value of the sum of the two random variables is equals to sum of their expected value.

Multiplication theorem—

If X and Y are independent random variables then -

$$\boxed{E(XY) = E(X).\,E(Y)}$$

Q A lot of 12 television sets include two with white cords. If three of the sets are choosen at random for shipment to hotel. How many sets with white cords can the shipper expect to sent to the Hotel?

X	0	1	2
P(x)	$^{10}C_3/^{12}C_3$	$\frac{^{10}C_2\,^{2}C_1}{^{12}C_3}$	$\frac{^{10}C_1\,^{2}C_2}{^{12}C_3}$

$$E(x) = 0 \times \frac{^{10}C_3}{^{12}C_3} + 1 \times \frac{^{10}C_2\,^{2}C_1}{^{12}C_3} + 2 \times \frac{^{10}C_1 \times {^{2}C_2}}{^{12}C_3}$$

$$= 0.5$$

Ans ⇒ The 0.5 can not be shipped to the hotel. Therefore the expectation of E(x) on average pertaining to repeate a shipment made under the given condition.

Q Certain coded measurments of the pitch diameter of threads of a fitting have the prob. density function -

$$F(x) = \begin{cases} \frac{4}{\pi(1+x^2)} & 0<x<1 \\ 0 & \text{elsewhere} \end{cases}$$

Scanned by CamScanner

Find the expected value of this random variable.

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

$$= \int_0^1 x \cdot \frac{4}{\pi(1+x^2)}\, dx$$

$1+x^2 = t$
$2x\,dx = dt$
$x\,dx = \frac{1}{2} dt$

$\frac{1}{2}\int \frac{dt}{t}$

$\frac{1}{2}\log(1+x^2)$

$$= \frac{4}{\pi} \int_0^1 \frac{x}{1+x^2}\, dx$$

$$= \frac{4}{\pi} \times \frac{1}{2}\left[\log(1+x^2)\right]_0^1$$

$$\boxed{= \frac{2}{\pi}\log 2}$$

Theorem —

If X is a continuous random variable and $f(x)$ is a value of its probability density functⁿ then the expected value of $g(x)$ is given by —

$$E[g(x)] = \sum g(x) f(x)$$

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$$

Example —

If X is the no. of points rolled with a unbiased die find the expected value of.

$$g(x) = 2x^2 + 1.$$

$$E[g(x)] = E[2x^2+1]$$

$$= \sum_{i=1} (2x_i^2+1)\, p(x_i)$$

$$= [2(1)^2+1] \times \frac{1}{6} + [2(2)^2+1] \times \frac{1}{6} + \cdots + [2(6)^2+1] \times \frac{1}{6}$$

$$= \frac{94}{3}$$

<u>Q</u> If $f(x)$ is probability density ~~f(x)~~

$$f(x) = \begin{cases} e^{-x} & x = 0 \\ 0 & \text{elsewhere} \end{cases}$$

Find the expected value of $g(x) = e^{3x/4}$

<u>Sol</u> $$E[e^{3x/4}] = \int_0^\infty e^{3x/4} e^{-x}\, dx$$

$$= \int_0^\infty e^{-x/4}\, dx$$

$$= -4\left(e^{-x/4}\right)_0^\infty$$

$$= 4$$

Theorem –

If $c_1, c_2 \dots c_n$ are constants then the expected value of

$$E\left[\sum_{i=1}^{n} c_i g_i(x)\right] = \sum_{i=1}^{n} c_i \, E[g_i(x)]$$

Exercise –

Q A random variable X is defined as the sum of the faces when a pair of die is throw. Find the expected value of X.

Q What the expected value of the heads appearing when a fair coin is tossed 8 times.

Q (3) If the Probability density functio X is given by $F(x) = \begin{cases} 2(1-x) & 0 < x < 1 \\ 0 & e.w \end{cases}$

fine $E(X^r)$, and $E[(2x+1)^2]$

also find $E[(ax+b)^n]$?

MINOR – I

(3) Sol $E(x^r) = \int_{-\infty}^{\infty} x^r F(x)\, dx$

$$= \int_0^1 x^r \, 2(1-x) dx$$

$$= 2\int_0^1 \cancel{x^r} - 2\int_0^1 x^{r+1}\, dx$$

$$= 2\left[\frac{x^{r+1}}{r+1}\right]_0^1 - 2\left(\frac{x^{r+2}}{r+2}\right]_0^1$$

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

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

$$= \frac{2}{(r+1)(r+2)} \quad \underline{Ans} \qquad r \neq -1, \; r \neq -2$$

(2) $= E[(2x+1)^2]$

$= E[4x^2 + 4x + 1]$

$= 4E[x^2] + 4E[x] + 1$

$= 4 \times \frac{2}{3\times4} + 4 \times \frac{2}{2\times3} + 1$

$= \frac{2}{3} + \frac{4}{3} + 1 \Rightarrow \frac{9}{3} = 3$ Ans

(3) Find expectation of $E[(ax+b)^n] = ?$

$E[(ax+b)^n] = \Sigma (ax+b)^n \, p(x)$

$$= \sum_{i=0}^{n} {}^nC_i \, a^{x-i} b^i \, E(x^{n-1})$$

Q If the p.d.f X is given by —

$$F(x) = \begin{cases} \frac{1}{x\log 3} & 1<x<3 \\ 0 & \text{e.w.} \end{cases}$$

find a) $E(x)$
b) $E(x^2)$
c) $E(x^3)$
d) $E[x^3+2x^2-3x+1]$

Sol a) $E(x) = \int_{-\infty}^{\infty} x\,f(x)\,dx$

$$= \int_1^3 \frac{x}{x\log 3}\,dx$$

$$= \frac{2}{\log 3}$$

b) $E(x^2) = \int_{-\infty}^{\infty} x^2 F(x)\,dx \Rightarrow \int_1^3 \frac{x^2}{x\log 3}\,dx$

$$= \frac{1}{\log 3}\int_1^3 x\,dx \Rightarrow \frac{1}{\log 3}\left[\frac{x^2}{2}\right]_1^3$$

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

c) $E(x^3) = \int_1^3 \frac{x^3}{x\log 3} \Rightarrow \frac{1}{\log 3}\left(\frac{x^3}{3}\right]_1^3 = \frac{26}{3\log 3}$

d) $E[x^3+2x^2-3x+1]$

$$= E[x^3] + 2E[x^2] - 3E[x] + 1$$

$$= \frac{26}{3\log 3} + 2\times\frac{4}{\log 3} - 3\times\frac{2}{\log 3} + 1$$

$=$

$$F(x) = \begin{cases} x/2 & 0 < x \leq 1 \\ 1/2 & 1 < x \leq 2 \\ \frac{3-x}{2} & 2 < x < 3 \\ 0 & \text{e.w} \end{cases}$$

$$F(x) = x^2 - 5x + 3$$

Spoorthi

MOMENTS —

→ Moments about the origin —

The ~~cou~~ r^{th} moment about the origin of the random variable X is denoted by μ_r' is the expectation of X^r symbolically.

$$\mu_r' = E(X^r)$$

→ If X is a descrete random variable

$$\mu_r' = E(X^r) = \sum_x x^r p(x)$$

when $p(x)$ is probability distr func.

for random variable X.

→ If X is a continuous random variable

$$\mu_r' = E(X^r) = \int_{-\infty}^{\infty} x^r F(x)\,dx$$

~~# M~~

→ Mean of a distribution —

μ_1' is called mean of the distr of the random variable X or the mean of X.

$$\mu_1' = E(x) = \bar{x} = \mu$$

Moments about the mean (μ)

The r^{th} moment about the mean of a random variable X is denoted by μ_r. is the expected value of $(X-\mu)^r$.

$$\mu_r = E[(X-\mu)^r]$$

If X is a descrete random variable, μ_r is nothing but.

$$= \sum_x (x-\mu)^r p(x)$$

If X is a continuous random variable, μ_r is nothing but.

$$= \int_{-\infty}^{\infty} (x-\mu)^r F(x)dx$$

$$\mu_0 = 1$$

$$\mu_1 = 0$$

Variance —

μ_2 is called variance of the distribution X and is denoted by σ^2 or σ_X^2 or $Var(x)$ or $V(x)$

The +ve square root of variance is called standard deviation.

$$\mu_2 = E[(x-\mu)^2]$$

$$\mu_2 = E[x^2 + \mu^2 - 2\mu x]$$

$$\mu_2 = E[x^2] + \mu^2 - 2\mu E[x]$$

$$\mu_2 = E[x^2] - \mu^2 \cancel{= E[x^2]}$$

$$\sigma^2 = \mu_2 = E[x^2] - [E(x)]^2$$

and

$$\boxed{\sigma^2 = \mu_2' - \mu^2}$$

$$\boxed{Var(ax+b) = a^2 Var(x)}$$

$$\boxed{Var(c) = 0}$$

Moment generating functions–

$$M_x(t) = E[e^{tx}] = \sum_x p(x) e^{tx}$$

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \cdots + \frac{t^r x^r}{r!} + \cdots$$

$$= \sum p(x) + t\sum x p(x) + \frac{t^2}{2!}\sum x^2 p(x) + \cdots \frac{t^r}{r!}\sum p(x) x^r$$

$$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \cdots + \frac{t^r}{r!}\mu_r' + \cdots$$

$$\mu_r' = E[x^r]$$

__Note__ → the coefficient of $\frac{e^r}{r!}$ is μ_r'

i.e. r^{th} moment about the origin.

__Q__ Find the moment generating function of the random variable X whose probability density function is given by

$$F(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & e.w. \end{cases}$$

and use it to find an expression for μ_r'.

$$M_x(t) = E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

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

$$= \int_{0}^{\infty} e^{-x(1-t)}\, dx$$

$$= \frac{-1}{(1-t)} \left[e^{-x(1-t)} \right]_0^{\infty}$$

$$= \frac{1}{1-t}(0-1) \Rightarrow \frac{1}{1-t}$$

$$M_x(t) = \frac{1}{1-t}$$

$$= (1-t)^{-1}$$

$$= 1 + t + t^2 + t^3 + \dots t^r + \dots$$

mul and div by $r!$

$$\therefore \; = 1 + t\mu_1' + \frac{t^2}{2}\mu_2' + \dots \frac{t^r}{r!}\mu_r' + \dots$$

$$= 1 + t(1) + \frac{t^2}{2!}2! + \frac{t^3}{3!}3! + \dots \frac{t^r}{r!}r! + \dots$$

On comparing

$$\boxed{\mu_r' = r!}$$

Theorem

Theorem –

$$\mu_r' = \left.\frac{d^r M_x(t)}{dt^r}\right|_{t=0}$$

$$M_x(t) = \frac{1}{1-t}$$

$$\mu_1' = \left.\frac{d}{dt}\left[(1-t)^{-1}\right]\right|_{t=0}$$

$$= (-1)(1-t)^{-2}(-1)$$

$$= 1$$

Q Given that X has the probability distribution $P(x) = \frac{1}{8}\binom{3}{x}$ $x = 0, 1, 2, 3$ find the moment generating function used to determine μ_1', μ_2' and the variance?

$$M_x(t) = E[e^{tx}] = \sum_{x=0}^{3} e^{tx} \frac{1}{8} {}^3C_x$$

$$= \frac{1}{8}\left[e^{t0}\,{}^3C_0 + e^{t}\,{}^3C_1 + e^{2t}\,{}^3C_2 + e^{3t}\,{}^3C_3\right]$$

$$= \frac{1}{8}\left[1 + 3e^{t} + 3e^{2t} + e^{3t}\right]$$

$$= \frac{1}{8}(1+e^{t})^3$$

$$\mu_1' = \left\{ \frac{d}{dt}\left[\frac{1}{8}(1+e^t)^3\right]\right\}_{t=0}$$

$$= \left[\frac{1}{8} * 3(1+e^t)^2 . e^t\right]_{t=0}$$

$$= \frac{3}{8} * 4 \Rightarrow \frac{3}{2}$$

$$\mu_2' = \left\{ \frac{d^2}{dt}\left[\frac{1}{8}(1+e^t)^3\right]\right\}_{t=0}$$

$$= \left\{ \frac{d}{dt}\left[3(1+e^t)^2 . e^t\right]\right\}_{t=0}$$

$$= \left\{\frac{1}{8}\left[6(1+e^t)e^{2t} + 3(1+e^t)^2 . e^t\right]\right\}_{t=0}$$

$$= \frac{1}{8} \times [12+12] \Rightarrow 3$$

$$\mu = \sum x P(x)$$

$$= \sum_{x=0}^{3} x \frac{1}{8} {}^3C_x$$

$$= \frac{1}{8}\left[0 + 1 \cdot {}^3C_1 + 2 \cdot {}^3C_2 + 3 \, {}^3C_3\right]$$

$$= \frac{1}{8}$$

$$\text{Var} = \mu_2' - \cancel{\mu^2} \quad 3 - 1/64$$

$$= 3 - \cancel{9}/4$$

$$= 3/4$$

Theorems On Moment Generating func

① $M_{x+a}(t) = e^{at} M_x(t)$

$$M_{x+a}(t) = E[e^{t(x+a)}]$$
$$= E[e^{tx+at}]$$
$$= e^{at} E[e^{tx}]$$
$$= e^{at} M_x(t)$$

② $M_{bx}(t) = M_x(bt)$

$$\star \quad M_{bx}(t) = E[e^{tbx}] = M_x(bt)$$

③ $M_{\frac{x+a}{b}}(t) = e^{a/b\, t} M_x\left(\frac{t}{b}\right)$

$$= E\left[e^{t\left(\frac{x+a}{b}\right)}\right]$$
$$= E\left[e^{\frac{tx}{b}+\frac{at}{b}}\right]$$
$$= e^{\frac{ab}{b}} E\left[\frac{tx}{b}\right]$$
$$= e^{\frac{at}{\cancel{b}}} M_x\left(\frac{t}{b}\right)$$

Q. Find the moment generating function -

$$f(x) = \frac{1}{2}e^{-|x|}, \quad -\infty < x < \infty$$

$$M_x(t) = E(e^{tx})$$

$$= \int_{-\infty}^{\infty} e^{tx} \times \frac{1}{2} e^{-|x|} \, dx$$

$$= \frac{1}{2}\left\{ \int_{-\infty}^{0} e^{tx} e^{x} dx + \int_{0}^{\infty} e^{tx} e^{-x} dx \right\}$$

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

$$= \frac{1}{2}\left\{ \left[\frac{1}{1+t} \left[e^{x(1+t)} \right]_{-\infty}^{0} \right\} + \frac{1}{2}\left\{ \frac{-1}{(1-t)} \left[e^{-(1-t)x} \right]_{0}^{\infty} \right\}$$

$$= \frac{1}{2}\left\{ \frac{1}{1+t}(1-0) \right\} + \frac{1}{2}\left\{ \frac{-1}{(1-t)}(0-1) \right\}$$

$$= \frac{1}{2(1+t)} + \frac{1}{2(1-t)}$$

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

Q Find the moment generating function of the exponential distr function –

$$f(x) = \frac{1}{c} e^{-x/c}, \quad 0 \le x < \infty, \quad c > 0$$

Hence find mean and standard deviation.

Sol
$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$= \frac{1}{c} \int_{-\infty}^{\infty} e^{tx} \cdot e^{-x/c}\,dx$$

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

$$= \frac{1}{c} * \frac{-1}{\left(\frac{1}{c}-t\right)} \left[e^{-\left(\frac{1}{c}-t\right)x}\right]_0^{\infty}$$

$$= \frac{1}{c} * \frac{-c}{1-ct} [0-1]$$

$$= \frac{1}{1-ct}$$

$$= (1-ct)^{-1}$$

Q An unbiased coin is tossed twice if x denotes the no. of heads that appear find the moment generating funct' of x. also find ~~var~~ mean and variance.

	(TT)	(TH) (HT)	(HH)
x	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

$$M_x(t) = e^{tx}\frac{1}{4} + e^{tx}\frac{1}{2} + e^{tx}\frac{1}{4}$$

$$= \sum e^{tx} p(x) \qquad \underline{e^{tx}}(1)$$

$$= e^{tx}.$$

Scanned by CamScanner

Binomial Distribution –

Its a distribution associated with repetition of independent trials of an experiment each trial has possibly two outcomes generally called as success and failure such a trial is known as Bernoulli's trial, the sample space for Bernoulli's trial is.

Examples –

① tossing an unbiased coin (H, T)

② throwing a die (getting even or odd)

③ Performance of a student in an exam (Pass or Fail)

④ Birth of child (Male or Female)

An experiment consisting of a repeated no. of Bernoulli's trial is called Binomial distribution. A Binomial experiment must ~~have~~ posses the following properties

(1) there must be fixed no. of trials (n) all the trials must have identical probabilities of success

(2) the trials must be independent of each other.

Definition of Binomial random variable

Let X be the no. of success in 'n' repeated independent Bernouli's trials with probability 'p' is success then X is called Bernouli's random variable with parameter denoted by —

$$X \rightarrow B(x, n, p)$$

The prob. mass function, or prob. distr. functn of the Binomial distribution of given X is given by

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

$q = 1-p$

Note -

① $\sum_{x=0}^{n} P(X=x) = \sum_{x=0}^{n} {}^{n}C_{x} p^{x} q^{n-x} = (p+q)^{n} = 1$

② Assume that n-trials constitute an experiment E if experiment E is repeated N times then the frequency function of the binomial distribution is defined by -

$$F(x) = N P(X=x)$$
$$= N \, {}^{n}C_{x} p^{x} q^{n-x}$$
$$x = 0, 1, 2 \ldots x$$

③ The mean and variance of the binomial variance.

$$\text{mean} = \mu = E(x) = \sum x \, p(x)$$
$$= \sum_{x=0}^{n} x \, {}^{n}C_{x} p^{x} q^{n-x}$$
$$= \sum x \frac{n!}{(n-x)! \, x!} p^{x} q^{n-x}$$
$$= \sum \frac{n(n-1)!}{(n-x)! \, (x-1)!} p \cdot p^{x-1} q^{n-x}$$

Scanned by CamScanner

$$= np \sum \frac{(n-1)!}{(x-1)!\,[(n-1)-(x-1)]!} \cdot p^{x-1} q^{[(n-1)-(x-1)]}$$

$$= np \sum_{x} {}^{(n-1)}C_{(x-1)}\, p^{x-1} q^{[(n-1)-(x-1)]}$$

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

$$= np(1)$$

$$\boxed{\mu = np}$$

<u>Variance</u> - $\mu_2' - \mu^2$

$$\mu_2' = E(x^2) = \sum x^2 p(x)$$

$$= \sum x^2\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum [x(x-1)+x] \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$

$$= \sum \frac{x(x-1)\,n!}{(n-x)!\,x!} p^x q^{n-x} + \sum x\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum \frac{\cancel{x(x-1)}\,n(n-1)(n-2)!}{(n-x)!\,\cancel{x(x-1)}(x-2)!} p^x q^{n-x} + np$$

$$= n(n-1)p^2 \sum \frac{(n-2)!}{[(n-2)-(x-2)]!\,(x-2)!} p^{x-2} q^{[(n-2)-(x-2)]} + np$$

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

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

$$\boxed{\mu_2' = n(n-1)p^2 + np}$$

then,

$$Var(x) = \sigma^2 = \mu_2' - \mu^2$$

$$= n(n-1)p^2 + np - n^2p^2$$

$$= \cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2}$$

$$= np(1-p)$$

$$\boxed{var(x) = npq}$$

$$\boxed{\sigma = \sqrt{npq}}$$

Scanned by CamScanner

Moment generating function of a Binomial distribution —

$$M_x(t) = E[e^{tx}] = \sum_x e^{tx} p(X=x)$$

$$= \sum_{x=0}^{n} e^{tx}\, {}^nC_x\, p^x q^{n-x}$$

$$= \sum {}^nC_x (pe^t)^x q^{n-x}$$

$$\boxed{M_x(t) = (pe^t+q)^n}$$

$$\boxed{\because \mu_r' = \frac{d^r M_x(t)}{dt^n}}$$

$$\mu_1' = \left[n(pe^t+q)^{n-1} \cdot pe^t\right]_{t=0}$$

$$\boxed{\mu_1' = np = \mu}$$

$\mu_2' = ?$

$\mu_3' = ?$

Q For a binomial distribution with parameters $n=5$, $p=0.3$ find the prob. of getting (1) at least 3 success (2) at most 3 success (3) exactly 3 failures.

$$p = \frac{3}{10}, \quad q = \frac{7}{10}, \quad n = 5$$

(a) ~~$p(x=0)+p(x=1)+p(x=2)+p(x=3)$~~

(b)

P.T.O

(1) $P(X \geq 3) = P(X=3) + P(X=4) + P(X=5)$

$= {}^5C_3 (0.3)^3 (0.7)^2 + {}^5C_4 (0.3)^4 (0.7)^1 + {}^5C_5 (0.3)^5$

$=$

(2) $P(X \leq 3)$

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

$= {}^5C_0 (0.3)^0 (0.7)^5 + {}^5C_1 (0.3)^1 (0.7)^4 + {}^5C_2 (0.3)^2 (0.7)^3$
$+ {}^5C_3 (0.3)^3 (0.7)^2$

$=$

(3)

Scanned by CamScanner

Q Comment the following the mean of binomial distr is 3 and variance is 4.

$\mu = 3 = np$

$\sigma^2 = 4 = npq$

$q = 4/3 > 1$

$\because$ The data cannot be of binomial distr, since q is > 1.

Q The mean and variance of binomial distr are 8 and 6 find prob. of ~~P(x~~ $P(X \geq 2)$

$\mu = np = 8$, var $= npq = 6$

$q = \frac{3}{4}$, $p = \frac{1}{4}$, $n = 32$

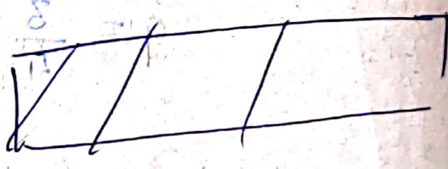
$P(X \geq 2) \Rightarrow 1 - P(X < 2)$

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

Scanned by CamScanner

Q A coin is known to come up heads 3 times as often as tails. This coin is tossed 3 times. Let X be the no. of heads that appear. Write down the prob. distribution of X.

Q Two players A and B play tennis games. Their chance of winning a game are in the ration 3:2 find the chance of player A atleast 2 games out of 4 games played.



$p = 3q$

$p = 3(1-p)$

$p = 3 - 3p$

$4p = 3$

$p = 3/4 \quad q = 1/4 \quad n = 3$

$$\boxed{pq\left[\frac{d\mu_k}{dp} + nk\mu_{k-1}\right] = \mu_{k+1}}$$

Recurrence formula for Binomial Distribution –

⇒ Recurrence relation for Binomial distribution –

$$\mu_K = E[(x-\mu)^k] \text{ ——— } (1)$$

$$= \sum_{x=0}^{n} (x-\mu)^k \, {}^nC_x p^x q^{n-x}$$

$$\mu_K = \sum_{x=0}^{n} (x-np)^k \, {}^nC_x p^x (1-p)^{n-x}$$

On diff^r. w.r.t p –

$$\frac{d\mu_K}{dp} = k(x-np)^{k-1}(-n)\,{}^nC_x p^x (1-p)^{n-x}$$

$$+\sum_{x=0}^{n} (x-np)^k \left[xp^{x-1}(1-p)^{n-x} + p^x(n-x)(1-p)^{n-x-1}(-1)\right]$$

$$\frac{d\mu_K}{dp} = -nk\sum (x-np)^{k-1}\,{}^nC_x p^x q^{n-x}$$

$$+\sum {}^nC_x (x-np)^k p^{x-1} q^{n-x-1}\left[x(1-p) - p(n-x)\right]$$

From (1)

$$= -nk\,\mu_{K-1} + \sum {}^nC_x (x-np)^k p^{x-1} q^{n-x-1}(x-np)$$

$$= -nk\,\mu_{K-1} + \sum (x-np)^{k+1}\,{}^nC_x p^{x-1} q^{n-x-1}$$

$$= -nk\,\mu_{K-1} + \frac{\sum}{pq} (x-np)^{k+1}\,{}^nC_x p^x q^{n-x}$$

$$= -nk\,\mu_{K-1} + \frac{1}{pq}\mu_{K+1}$$

P.T.O

$$\boxed{\mu_0 = 1}$$
$$\boxed{\mu_1 = 0}$$

$\boxed{k=1}$

$$\mu_2 = pq\left[\frac{d\mu_1}{dp} + n(1)\mu_0\right]$$

$$= pq[0+n]$$

$$\boxed{\mu_2 = npq} \Rightarrow \boxed{\mu_2 = np(1-p)}$$

$\boxed{K=2}$

$$\mu_3 = pq\left[\frac{d\mu_2}{dp} + n(2)\mu_1\right]$$

$$= pq\left[n(1-p-p) + 2n(0)\right]$$

$$= npq(1-2p)$$

$$= np(1-p)(1-2p)$$

Q. Assume that half of the population is vegetarians so that the chance of an individual being a vegeterial is of. Assume that 100 investigators takes samples of 10 individuals each to see whether they are vegeterian. So how many investigator would you expect to report that 3 people or less were vegetarian.

$x = 10,\ p = \frac{1}{2},\ q = \frac{1}{2},\ N = 100$

$P(X = x) = N\ {}^{n}C_{x}\, p^{x} q^{n-x}$

$= 100 \times {}^{10}C_{x} \left(\frac{1}{2}\right)^{x} \left(\frac{1}{2}\right)^{10-x}$

$P(x \leq 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$

Q① If the prob. distⁿ of a binomial distribution with parameter p and n is denoted by.

$$X \sim B(n, p)$$

then

$$P(X = x) = B(x, n, p)$$

prove that

① $$p(x+1) = \frac{n-x}{x+1} \cdot \frac{p}{q} \, p(x)$$

②

Q An unbiassed coin is tossed 8 times and the no. of heads noted. the experiment is repeated 256 times and following freq. distⁿ is obtained. calculate eradical freq.

No. of heads	0	1	2	3	4	5	6	7	8
Freq.	2	6	30	52	67	56	32	10	1

1 Sol

$$p(x) = {}^nC_x \, p^x q^{n-x}$$

$$p(x+1) = {}^nC_{x+1} \, p^{x+1} q^{n-x-1}$$

$$\frac{P(x+1)}{P(x)} = \frac{{}^nC_{x+1}}{{}^nC_x} \cdot \frac{p}{q} \Rightarrow \frac{\dfrac{\cancel{n!}}{(n-x-1)!\,(x+1)!}}{\dfrac{\cancel{n!}}{(n-x)!\,x!}}$$

$$\frac{{}^nC_{x+1}}{{}^nC_x} = \frac{\dfrac{n!}{(n-x-1)!\,(x+1)!}}{\dfrac{n!}{(n-x)!\,x!}} \Rightarrow \frac{(n-x)\,\cancel{(n-x-1)!}\,\cancel{x!}}{\cancel{(n-x-1)!}\,(x+1)\,\cancel{x!}}$$

$$\boxed{\frac{p(x+1)}{\cancel{p(x)}} = \frac{(n-x)}{(x+1)} \times \frac{p}{q}\, p(x)}$$

q2

$$\cancel{256}\, p(1) = \frac{8-0}{0+1} \times \frac{1/2}{1/2} \times \cancel{256}$$

$$p(1) = 8$$

Poission Distribution —

Poission distribution is a deserete distribution and poission distr is the limiting case of Binomial distr. Poission distr can be used under the following conditions

1. Each trial results two mutually exclusive outcomes say success and failure.
2. n the number of trials are very large $n \to \infty$
3. The constant probability of success is very small i.e. $np = \lambda$
4. $np = \lambda$ where λ is +ve real no. is finite.

$\boxed{p = \frac{\lambda}{n}}$, $q = 1 - p = \boxed{1 - \frac{\lambda}{n} = q}$

5

$\frac{e^{-\lambda} \lambda^{x}}{\lfloor x}$ $\frac{e^{-\lambda} \lambda^{x}}{\lfloor x}$ $\frac{e^{-\lambda} \lambda^{x}}{\lfloor x}$

$\frac{e^{-\lambda} \lambda^{x}}{\lfloor x}$

Example -

$\Rightarrow$ The no. of defective items out of lots produced in a manufacturing factory.

$\Rightarrow$ No. of printing mistakes on each page of a book.

Definition - Let X be descrete random variable taking the values. If the prob. mass function.

$$\boxed{P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!}}$$

then X is said to follow poission distr with parameter λ and denoted by $P(\lambda)$

$\Rightarrow$ Poission distribution is limiting case of Binomial distribution.

$$P(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}$$

$$= \frac{n!}{(n-x)!\,x!} \cdot \left(\frac{\lambda}{n}\right)^{x}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

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

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

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

$$\therefore \lim_{n\to\infty} p(X=x) = \frac{\lambda^x}{x!} 1 \cdot e^{-\lambda} \cdot 1$$

$$\boxed{p(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}}$$

Note → the above results means that we may compute Binomial prob. by using the corresponding Poission prob. whenever n is very large and p is very small.

Mean And Variance of Poission dst^r.

Mean (μ)

$$\mu = E(x) = \sum_x x\,p(x)$$

$$= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} \frac{\cancel{x}\, e^{-\lambda}\lambda^x}{\cancel{x}(x-1)!}$$

$$= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

Scanned by CamScanner

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$\left\{ \begin{array}{l} x-1=y \Rightarrow x=y+1 \\ x=1,\ y=0 \end{array} \right.$$

$$= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!}$$

$$= \lambda \cancel{e^{-\lambda}} \cdot \cancel{e^{\lambda}}$$

$$\boxed{\mu_1 = \lambda}$$

<u>Variance</u>

$$\sigma^2 = \mu_2' - \mu_1^2$$

$$\mu_2' = E(x^2) = \sum_x x^2 p(x)$$

$$= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} [x(x-1)+x] \frac{e^{-\lambda} \cdot \lambda^x}{x(x-1)!}$$

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

$$= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^x}{(x-2)!} + \lambda$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda$$

$$\mu_2' = \lambda^2 + \lambda$$

Variance $= \mu_2' - \mu_1^2$

$$\sigma^2 = \lambda^2 + \lambda - \lambda^2$$

$$\sigma^2 = \lambda$$

$$\boxed{\text{Stand. devi} = \sqrt{\lambda}}$$

Moment generating functn of a poission distribution –

$$M_x(t) = E[e^{tx}] = \sum_{x=0}^{\infty} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \cdot \frac{\lambda^x}{x!} e^{-\lambda}$$

$$= \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} e^{-\lambda}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda t}$$

$$\boxed{E[e^{tx}] = e^{\lambda[e^t - 1]}}$$

Find μ_2', μ_3'

$$\mu_1' = \frac{d}{dt}\left[e^{\lambda(e^t - 1)}\right]\Big|_{t=0}$$

$$= \left\{e^{\lambda(e^t-1)} \cdot \lambda \cdot e^t\right\}\Big|_{t=0}$$

$$= \lambda$$

Q Find μ_2', μ_3' from moment generating function.

Q After correcting the proofs of the 50 pages of the book it is found that on the average there are 3 errors per 5 pages. Use poission probabilities and estimate the no. of pages with 0, 1, 2, 3 errors in the whole book of thousand pages.

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

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

$$P(X=0) = \frac{1000 \times e^{-3/5} \cdot 0!}{0!}$$

$$= 1000 \times 0.5488$$

$$= 548.8 \approx 549 \text{ Pages}$$

$$P(X=1) = \frac{1000 \times e^{-3/5} \cdot \frac{3}{5}}{1}$$

$$= 1000 \times$$

Q A factory employing a large number of workers finds that over a period of time the average absenties rate is 3 workers per shift. Calculate the probability that in the given shift —

(1) Exactly 2 will be absent.

(2) More than 4 will be absent.

Thumb rule to use poission distribution as Binomial distribution

$$n \geq 100$$

$$np \leq 10$$

Recurrance relation for the moments of the poission distribution with parameter λ.

$$\mu_r = E(x-\mu)^r$$

$$= E[x-\lambda]^r$$

$$= \sum_{x=0}^{\infty} (x-\lambda)^r P(X=x)$$

$$\mu_r = \sum_{x=0}^{\infty} (x-\lambda)^r \frac{e^{-\lambda} \lambda^x}{x!}$$

$$\frac{d\mu_r}{d\lambda} = \sum \frac{1}{x!} \Big\{ r(x-\lambda)^{r-1}(-1)e^{-\lambda}\lambda^x + (x-\lambda)^r \left[-e^{-\lambda}\lambda^x + e^{-\lambda} x . \lambda^{x-1}\right]$$

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

$$= -r\,E\left[(x-\lambda)^{r-1}\right]+E\left[(x-\lambda)^{r+1}\right]$$

$$= -r\,\mu_{r-1}+\frac{1}{\lambda}\mu_{r+1}$$

$$\boxed{\mu_{r+1}=\lambda\left[\frac{d\mu_r}{d\lambda}+r\mu_{r-1}\right]}$$

$\because \mu_0 = 1$
$\mu_1 = 0$

\# $\mu_2 = \lambda\left[\frac{d\mu_1}{d\lambda}+\mu_0\right]$

$\mu_2 = \lambda[0+1] \Rightarrow \boxed{\mu_2=\lambda}$

\# $\mu_3 = \lambda\left[\frac{d\mu_2}{d\lambda}+2\mu_1\right]$

$= \lambda[1+2\times 0] \Rightarrow \boxed{\mu_3=\lambda}$

\# $\mu_4 = \lambda\left[\frac{d\mu_3}{d\lambda}+3\mu_2\right]$

$$\boxed{\mu_2 = \lambda[1+3\lambda]}$$

Q Calculate ~~to~~ poission distribution for the following frequency distribution

2
142
156
69
21
6
394

x	0	1	2	3	4	5
F	142	156	69	27	5	1
Fx	0	156	138	81	20	5

2
156
138
63
25
382

$$\text{mean} = \lambda = \bar{x} = \frac{\Sigma Fx}{\Sigma F} = \cancel{\frac{382}{394}}$$

$$\lambda = \frac{400}{400}$$

$$\lambda = 1$$

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-1}(1)^{x}}{x!} \quad x = 0, 1, 2, 3, 4, 5, 6$$

x	0	1	2	3	4	5	6
P(X=x)	~~0.3679~~ 147.15	147.15	73.58	24.52	6.131	1.225	0.2
~	147	147	74	25	6	1	0

Q On the record of 10 Indian army caps kept over 20 years the following data were obtained the no. of death caused by a horse. Calculate the periodical poission frequencies.

No. of deaths	0	1	2	3	4
Frequencies	109	65	22	3	1

$\cancel{x+y=6}$

$x+2y=8$

Hyper Geometric Distribution –

⇒ Sampling with replacement

If each object is selected from the sample would have been replaced before the next one is drawn.

⇒ The trials are not Independent.

⇒ Suppose we are interested in no. of defectives in a sample of size n units drawn from a lot containing N units of which 'a' are defective.

⇒ Sampling without replacement.

No. of ways in which 'x' success (defective) can be choosen is ${}^aC_x$.

⇒ The No. of ways in which '$n-x$' failures (Not defectives) be choosen as $\binom{N-a}{n-x}$

⇒ Hence the no. of ways 'x' success and '$n-x$' failure (non defective) can be choosen is $\binom{a}{x}\binom{N-a}{n-x}$

Scanned by CamScanner

the no. of ways n objects can be choosen from N objects is $\binom{N}{n}$ if all the possibilities are equally likely then for sampling without replacement the probability of getting x success in 'n' trials is given by

$$\boxed{h(x,n,a,N) = \frac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}}}$$

$x = 0, 1, 2, \ldots n$.

when $x \leq a$, $(n-x) \leq (N-a)$

→ In general, $h(x,n,a,N) \rightarrow B(x,n,p)$ with, $p = \frac{a}{N}$ when $N \rightarrow \infty$

Tip → thumb rule to use the Binomial distr as an approximation of Hyper Geo. distr if —

$$\frac{n}{N} \leq 0.05$$

Mean and Variance of Hyper Geometric distribution—

$$\boxed{\text{Mean} = \mu = \frac{n \cdot a}{N}}$$

$\because$ $n \rightarrow$ sample size

$N \rightarrow$ population size

$a \rightarrow$ no. of success

$\Rightarrow$ Show that the mean of hypergeometric distr is given by—

$$\boxed{\mu = n \cdot \frac{a}{N}} \qquad n - x \leq N - a$$

$$\mu = E(x) = \mu_1' = \sum_x x P(X = x)$$

$$= \sum_{x=0}^{n} x \frac{{}^{a}C_{x} \, {}^{(N-a)}C_{(n-x)}}{{}^{N}C_{n}}$$

$$\Rightarrow \frac{a}{\binom{N}{x}} \sum_{x=1}^{\infty} \binom{a-1}{x-1} \binom{N-a}{n-x}$$

$$\left\{ x \cdot {}^{a}C_{x} = x \frac{a!}{(a-x)! \, x!} = \frac{a!}{(a-x)(x-1)!} = \frac{a \cdot (a-1)!}{(a-x)! \, (x-1)!} = \frac{a \, (a-1)!}{(x-1)!} \right.$$

★ let $y = x - 1$

$x = 1, \; y = 0$

$x = x, \; y = x - 1 = m$

$$\Rightarrow \frac{a}{\binom{N}{n}} \sum_{y=0}^{m} \binom{a-1}{y} \binom{N-a}{m-y}$$

Scanned by CamScanner

$$= \sum_{r=0}^{m} \binom{a}{r}\binom{b}{m-r}$$

$$= \binom{a+b}{m}$$

$$= \frac{a}{\binom{N}{n}} \binom{N-1}{m}$$

$$= \frac{a}{N!} \times (\cancel{N}-n)! \times n! \times \frac{(N-1)!}{(\cancel{N}-n)! \times (n-1)!}$$

$$\boxed{\mu_1 = \frac{a \cdot n}{N}}$$

<u>Variance</u> - Proove that the variance of hyper geometric distribution is —

$$\sigma^2 = \frac{n \cdot a (N-a)(N-n)}{N^2 (N-1)}$$

$$\because \sigma^2 = \mu_2' - \mu_1^2$$

$$\mu_2' = E(x^2) = \sum \frac{x^2 \binom{a}{x} \binom{N-a}{n-x}}{\binom{N}{n}}$$

$$= a \sum \frac{x \binom{a-1}{x-1} \binom{N-a}{n-x}}{\binom{N}{n}}$$

$$\Rightarrow \frac{a}{\binom{N}{n}} \sum_{x=1}^{n} [(x-1)+1] \binom{a-1}{x-1} \binom{N-a}{n-x}$$

$$= \frac{a}{\binom{N}{n}} \sum_{x=1}^{n} (x-1) \binom{a-1}{x-1} \binom{N-a}{n-x}$$

$$+ \frac{a}{\binom{N}{n}} \sum_{x=1}^{n} \binom{a-1}{x-1} \binom{N-a}{n-x}$$

$$\left\{ \begin{array}{l} \frac{(x-1).(a-1)!}{(a-x)!\,(x-1)!} \\ = \frac{(a-1)(a-2)!}{(a-x)(x-2)!} \end{array} \right.$$

I^{st} term

$$= \frac{a(a-1)}{\binom{N}{n}} \sum_{x=2}^{n} \binom{a-2}{x-2} \binom{N-a}{n-x}$$

$$+ \frac{a(a-1)}{\binom{N}{n}} \sum_{x=1}^{n} \binom{a-2}{x-2} \binom{N-a}{n-x}$$

2^{nd} term

Let $y = x-2$ in 1^{st} term

$x = 2 \Rightarrow y = 0$

$x = n \Rightarrow y = x-2 = m$

$y = x-1$ in 2^{nd} term

$x = 1,\ Z = 0$

$x = n \Rightarrow Z = n-1 = k$

$$\Rightarrow \frac{a(a-1)}{\binom{N}{n}} \sum_{y=0}^{m} \binom{a-2}{y} \binom{N-a}{m-y} + \frac{a}{\binom{N}{n}} \sum_{z=0}^{k} \binom{a-1}{z} \binom{N-a}{k-z}$$

$$\Rightarrow \frac{a(a-1)}{\binom{N}{n}} \binom{N-2}{m} + \frac{a}{\binom{N}{n}} \binom{N-1}{k}$$

$$\Rightarrow \frac{a(a-1)}{\binom{N}{n}} \binom{N-2}{n-2} + \frac{a}{\binom{N}{n}} \binom{N-1}{n-1}$$

$$= \frac{a(a-1)}{N!} \cdot \frac{\cancel{(N-n)!}\, n!\, \cancel{(N-2)!}}{\cancel{(N-n)!}\,(n-2)!}$$

$N(N-1)\cancel{(N-2)!}$

$$+ \frac{a}{N!} \cdot \frac{\cancel{(N-n)!}\, n!\, (N-1)!}{\cancel{(N-n)!}\,(n-1)!}$$

$$\boxed{\mu_2' = \frac{a(a-1)}{N(N-1)} \cdot n(n-1) + \frac{a}{N} \cdot n}$$

then,

$$\sigma^2 = \mu_2' - \mu_1^2$$

$$= \frac{a(a-1)n(n-1)}{N(N-1)} + \frac{an}{N} - \frac{n^2a^2}{N^2}$$

$$= n \cdot \frac{a}{N}\left[\frac{(a-1)(n-1)}{(N-1)} + 1 - n \cdot \frac{a}{N}\right]$$

$$= n \cdot \frac{a}{N}\left[\frac{N(a-1)(n-1) + N(N-1) - na(N-1)}{N(N-1)}\right]$$

$$\boxed{\sigma^2 = \frac{n \cdot a(N-a)(N-n)}{N^2(N-1)}}$$

Q. Among the 120 applicants for a job only 80 are actually qualified if five of the applicants are randomly selected for ~~test~~ an indepth interview find the probability that only two of 5 will be qualified for the job by using -

(1) the formula for the hypergeometric distr.

(2) the for^n for the Binomial distr with $p = \frac{80}{120}$ as an approximation.

Q. A carton contains light bulbs in which defective are 3 total 24 bulbs. what is the prob. that if a sample of 6 is choosen at random from the cartoon of bulbs will be defective.

Continuous distribution —

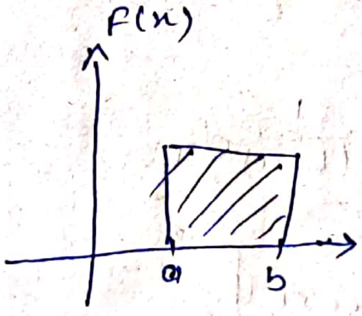
a- Uniform distr.
b- Exponential distr.
c- Normal distr.

UNIFORM DISTRIBUTION — (Rectangular Distribution)

Let x be a continuous random variable assuming all values in the interval (a, b) a, and b are finite constant. the prob. density funct^n of is given by —

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & , \text{elsewhere} \end{cases}$$

then x is said to be distributed uniformely for the interval a, b, and the distr. is called uniform distribut^n



Property —

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

$$= \int_{-\infty}^{0} 0\,dx + \int_{a}^{b} \frac{1}{b-a}\,dx + \int_{0}^{\infty} 0\,dx$$

$$= 1$$

Scanned by CamScanner

Note —

If x is uniformely distributed random in the open interval (a, b) if (α, β) is sub interval of (a, b) then the probability of —

$$\boxed{P(\alpha < x < \beta) = \frac{\beta - \alpha}{b - a}}$$

→ Cumulative distribution function of the Uniform distribution

It is given by

$$F(x) = P(X \leq x)$$

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

$$= \int_{a}^{x} f(x)\, dx$$

$$= \int_{a}^{x} \frac{1}{(b-a)}\, dx.$$

$$\boxed{F(x) = \frac{x - a}{b - a}}$$

Scanned by CamScanner

Mean and Variance of the Uniform Distribution —

⇒ **Mean —**

$$\mu = E(x) = \mu_1' = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$= \int_a^b x \cdot \frac{1}{b-a}\, dx$$

$$= \left[\frac{x^2}{2}\right]_a^b \frac{1}{(b-a)}$$

$$\boxed{\mu = \frac{a+b}{2}}$$

⇒ Variance —

$$\sigma^2 = \mu_2' - \mu^2$$

$$\mu_2' = E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$$

$$= \int_a^b \frac{x^2}{(b-a)}\, dx$$

$$= \frac{(b^3 - a^3)}{} \times \frac{1}{3(b-a)}$$

$$\mu_2' = \frac{b^2 + ab + a^2}{3}$$

then

$$\text{Varianc} = \frac{b^2 + ab + a^2}{3} - \frac{(a+b)^2}{4} \Rightarrow \frac{4a^2 + 4b^2 + 4ab - 3a^2 - 3b^2 - 6ab}{12}$$

$$= \frac{a^2 + b^2 - 2ab}{12}$$

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

$$\boxed{\text{Variance} = \frac{(a-b)}{\sqrt{12}}}$$

Moment generating function of the uniform distribution —

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$M_x(t) = \int_a^b e^{tx} \cdot \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a} \cdot \frac{1}{t} (e^{tx})_a^b$$

$$\boxed{M_x(t) = \frac{1}{t(b-a)} \left[e^{tb} - e^{ta} \right]}$$

Cumulative distribution function of the random variable X is given by —

$$F(x) = \begin{cases} 0 & x \le 0 \\ \frac{x}{2\pi} & 0 < x \le 2\pi \\ 1 & x > 2\pi \end{cases}$$

find the p.d.f of X and show that it has unifor distr in interval $(0, 2\pi)$ also find prob. of $\left(\frac{\pi}{4} \le x \le \frac{\pi}{2} \right)$.

$$f(x) = \frac{dF(x)}{dx}$$

$$f(x) = \begin{cases} 0 & x \le 0 \\ \frac{1}{2\pi} & 0 < x < 2\pi \\ 0 & x > 2\pi \end{cases}$$

$$f(x) = \begin{cases} \frac{1}{2\pi} & 0 < x < 2\pi \\ 0 & \text{elsewhere} \end{cases}$$

Scanned by CamScanner

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{2\pi} \frac{1}{2\pi}\,dx = 1$$

then

$$P\left(\frac{\pi}{4} \le x \le \frac{\pi}{2}\right) = \frac{\frac{\pi}{2} - \frac{\pi}{4}}{2\pi - 0}$$

$$=$$

Exponential distribution

X is a continuous random variable with prob. density function $F(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$, $x > 0$ $\theta > 0$ then

X ~~has~~ do have exponential distr.

$$\Gamma(a+1) = \int_0^\infty e^{-x} x^a \, dx = \text{gamma of } (a+1)$$

$$\Gamma(3+1) = \int_0^\infty e^{-x} x^3 \, dx$$

$\Gamma(a+1) = a!$ ∵ If a is an integer

ex – $\Gamma 5 = \Gamma(4+1) = 4!$

① $\Gamma(a+1) = a\Gamma a$

② $\Gamma \frac{1}{2} = \sqrt{\pi}$

③ $\Gamma \frac{3}{2} = \Gamma\left(\frac{1}{2}+1\right) = \frac{1}{2}\Gamma\frac{1}{2} = \frac{1}{2}\sqrt{\pi}$

④ $\Gamma a = \frac{\Gamma(a+1)}{a}$

⑤ $\Gamma\left(\frac{-1}{2}\right) = \frac{\Gamma\left(\frac{-1}{2}+1\right)}{\left(\frac{-1}{2}\right)} = -2\sqrt{\pi}$

ex :- $\Gamma \frac{5}{2} = \Gamma\left(\frac{5}{2}+1-1\right) = \Gamma\left(\frac{3}{2}+1\right) = \frac{3}{2}\Gamma\frac{3}{2}$

Mean and Variance of the Exponential Distribution —

$$\begin{array}{cc} D & I \\ x & e^{-x/\theta} \\ 1 & (-\theta)e^{-x/\theta} \\ 1 & (\theta^2)e^{-x/\theta} \end{array}$$

$x.e^{-x/\theta}$

Diff In

$= -x\theta e^{-x/\theta} - \theta^2 e^{-x/\theta}$

Mean —

$$\text{Mean} = \mu_1' = E(X) = \int_{-\infty}^{\infty} x F(x)\, dx$$

$$= \int_0^{\infty} x \cdot \frac{1}{\theta} e^{-x/\theta} dx$$

$$= \frac{1}{\theta} \int_0^{\infty} x . e^{-x/\theta} dx$$

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

$$= \frac{1}{\theta} \left[(0-0) - (0 - \theta^2) \right]$$

$$\boxed{\text{mean} = \frac{\theta^2}{\theta} = \theta}$$

Variance – $\sigma^2 = \mu_2' - \mu^2$

$$\mu_2' = E(X^2) = \int_{-\infty}^{\infty} x^2 F(x) dx$$

$$= \int_0^{\infty} x^2 \left(\frac{1}{\theta} e^{-x/\theta} \right) dx$$

$$= \frac{1}{\theta} \int_0^{\infty} x^2 e^{-x/\theta} dx$$

$$\begin{array}{cc} D & I \\ x^2 & e^{-x/\theta} \\ 2x & (-\theta)e^{-x/\theta} \\ 2 & (\theta^2) e^{-x/\theta} \\ 0 & -\theta^3 e^{-x/\theta} \end{array}$$

$$= \frac{1}{\theta} \left[-\theta x^2 e^{-x/\theta} - 2x\theta^2 e^{-x/\theta} - 2\theta^3 e^{-x/\theta} \right]$$

$$= \frac{1}{\theta} \left[(0-0-0) - (-2\theta^3) \right]$$

$$\boxed{\mu_2' = 2\theta^2}$$

Scanned by CamScanner

Variance $= \mu_2' - \mu^2$

$= 2\theta^2 - \theta^2$

$$\boxed{\text{Variance} = \theta^2}$$

Moment generating function for Exponential distribution.

$$M_x(t) = E(e^{tx}) = \int_0^\infty e^{tx} \cdot \frac{1}{\theta} e^{-x/\theta} dx$$

$$= \frac{1}{\theta} \int_0^\infty e^{-\left(\frac{1}{\theta} - t\right) x} dx$$

$$= \frac{1}{\theta} \times \frac{-1}{\left(\frac{1}{\theta} - t\right)} \left[e^{-\left(\frac{1}{\theta} - t\right)x} \right]_0^\infty$$

$$= \frac{-1}{\theta} \times \frac{1}{\left(\frac{1}{\theta} - t\right)} [0 - 1]$$

$$= \frac{1}{\theta\left(\frac{1}{\theta} - t\right)}$$

$$= \frac{1}{1 - \theta t}$$

$$\boxed{= (1 - \theta t)^{-1}}$$

Q Suppose the duration of the telephone calls handled by the exchange can be reasonabally assumed to have exponential distr with parameter $\theta = 2$ time being measured in minutes what is the probability that a telephone call would last atleast two minutes, and that a call which has lasted 2 mins, lasts further 2 mins. (2 more min = 4 mins).

$$\theta = 2$$

$$P(X>2) = \int_2^\infty \frac{1}{2} e^{-x/2}\, dx$$

$$\cancel{P\left(\frac{2<x<4}{x>2}\right) = P\left(\frac{A}{B}\right) = \frac{P(A\cap B)}{P(B)}}$$

$$P\left(\frac{A}{B}\right) = \frac{P(A\cap B)}{P(B)} = \frac{P(X>4)}{P(X>2)}$$

$$P(X>4) = \int_4^\infty \frac{1}{2} e^{-x/2}\, dx$$

Normal Distribution —
(Gaussian distribution)

Examples — where they follow

1. The population
2. The Heights of people
3. Blood preasure.
4. Size of the things produced by machine.

Eqn of Normal probability distr curve —

If X is a continuous random variable following normal distribution with mean μ and standard deviation σ then, its p.d.f is given by:

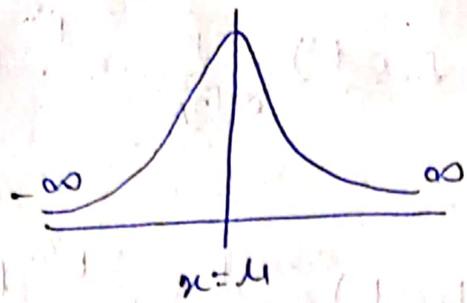
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty$$

$$\boxed{F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}}$$

where $\pi = \frac{22}{7}$, $\sqrt{2\pi} = 2.5066$

$e = 2.718$

Scanned by CamScanner



Area under curve —

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

$\therefore$ put $\frac{x-\mu}{\sigma\sqrt{2}} = t$, $\frac{dx}{\sigma\sqrt{2}} = dt$

$$= \int_{-\infty}^{\infty} \frac{1}{\cancel{\sigma}\cancel{\sqrt{2}}\sqrt{\pi}} e^{-t^2} \cdot \cancel{\sigma}\cancel{\sqrt{2}}\, dt$$

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

$$= \frac{1}{\sqrt{\pi}} \times \sqrt{\pi} \Rightarrow 1$$

Mean and variance of Normal distribution —

$\Rightarrow$ **MEAN** —

$$\mu_1 = E(x) = \int_{-\infty}^{\infty} x\, F(x)\, dx$$

$$= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]} dx$$

$\therefore$ put $\frac{x-\mu}{\sqrt{2}\sigma} = t$, $dx = \sqrt{2}\sigma\, dt$

$x \to \infty$, $t \to \infty$, $x \to -\infty$
$t \to -\infty$

$$= \int_{-\infty}^{\infty} (\mu + \sqrt{2}\sigma t) \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-t^2} \cdot \sqrt{2}\sigma \, dt$$

$$= \int_{-\infty}^{\infty} (\mu + \sqrt{2}\sigma t) \cdot \frac{1}{\sqrt{\pi}} \cdot e^{-t^2} dt$$

$$= \frac{\mu}{\pi} \int_{-\infty}^{\infty} e^{-t^2} dt + \frac{\sqrt{2}\sigma}{\pi} \int_{-\infty}^{\infty} t \cdot e^{-t^2} dt$$

$\because$ put $t^2 = y \Rightarrow 2t\,dt = dy \Rightarrow t\,dt = \frac{dy}{2}$

$t \to \infty$ then $y \to \infty$, $t \to -\infty$, $y \to \infty$

$\therefore SR = \int_{-\infty}^{\infty} \frac{e^{-y}}{2} dy = 0 \quad = \cancel{2\int_0^{\infty} e^{-y} dy}$

$$= \frac{\mu}{\sqrt{\pi}} \sqrt{\pi} + 0$$

$$\boxed{\text{Mean} = \mu}$$

VARIANCE —

$$\sigma^2 = \mu_2' - \mu^2$$

$$\mu_2' = E(x^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

$$= \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\left[\frac{(x-\mu)^2}{2\sigma^2}\right]} dx$$

$\left\{ \text{put } \frac{x-\mu}{\sqrt{2}\sigma} = t, \; dx = \sqrt{2}\sigma\,dt \right\}$

$$= \int_{-\infty}^{\infty} (\mu + \sqrt{2}\sigma t)^2 \cdot \frac{1}{\cancel{\sigma}\cancel{\sqrt{2}}\sqrt{\pi}} \cdot e^{-t^2} \cdot \cancel{\sqrt{2}}\cancel{\sigma}\, dt$$

$$= \int_{-\infty}^{\infty} (\mu^2 + 2\sqrt{2}\sigma\mu t + 2\sigma^2 t^2) \cdot \frac{1}{\sqrt{\pi}} e^{-t^2} dt$$

$$= \frac{\mu^2}{\sqrt{\pi}} \cdot \sqrt{\pi} + \frac{2\sqrt{2}\sigma}{\sqrt{\pi}} \cdot \int_{-\infty}^{\infty} t e^{-t^2} dt + \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt$$

$$= \mu^2 + 2 \cdot \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} t^2 e^{-t^2} dt$$

$$\left\{ \text{put } t^2 = y,\quad t = y^{\frac{1}{2}} \Rightarrow 2t\,dt = dy,\ t\,dt = \frac{dy}{2} \right\}$$

$$= \mu^2 + \frac{4\sigma^2}{\sqrt{\pi}} \int_0^{\infty} \frac{\sqrt{y}\, e^{-y}}{2} dy$$

$$= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \int_0^{\infty} e^{-y} y^{1/2} dy$$

$$= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \Gamma\left(\frac{1}{2} + 1\right)$$

$$= \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \times \frac{1}{2} \times \sqrt{\pi}$$

$$\boxed{\mu_2' = \mu^2 + \sigma^2}$$

4th
> moment
< moment

denoted by - $N(\mu, \sigma^2)$

P.T.O

then variance,

$$\sigma^2 = \mu^2 + \sigma^2 - \mu^2$$

$$\sigma^2 = \sigma^2$$

$\therefore$ Variance is $\sigma^2 = 1$

Standard Normal distribution —

If X is a continuous random variable following a normal distr with mean μ and s.d σ then the random variable Z is defined by.

$$Z = \frac{X - E(X)}{\sigma} = \frac{x - \mu}{\sigma}$$

is called standard normal variate

Mean —

$$E(Z) = E\left(\frac{x-\mu}{\sigma}\right)$$

$$= \frac{1}{\sigma}[E(X) - \mu]$$

$$= \frac{1}{\sigma}[\mu - \mu]$$

$$\boxed{\text{Mean} = 0}$$

Variance of standard normal variate

$$Var(z) = var\left(\frac{x-\mu}{\sigma}\right)$$

$$= \frac{1}{\sigma^2} var(x-\mu)$$

$$= \frac{1}{\sigma^2}[\sigma^2 - 0]$$

$$\boxed{Var(z) = 1}$$

Mean

Variance

$$\boxed{\begin{array}{l} X \sim N(\mu, \sigma^2) \\ Z \sim N(0, 1) \end{array}}$$

Hence the p.d.f of standard normal variate is given by –

$$\boxed{\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}}$$

for $-\infty < z < \infty$

Note –

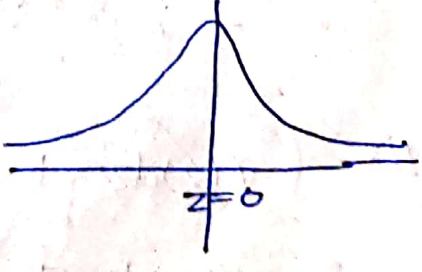
1. $\phi(z)$ is symmetric about $z = 0$.

becoz, $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$

$\phi(-z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$

$\phi(z) = \phi(-z)$

$x = \mu,\ z = 0.$



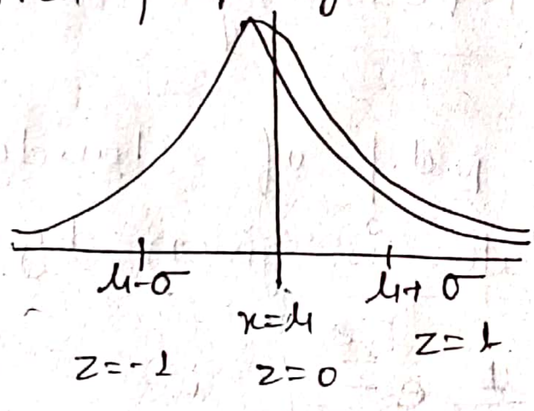
2. $P(Z \le z) = \int_{-\infty}^{z} \phi(z)\, dz = \Phi(z)$

The normal distribution is catrectorized by two parameter

① Mean (μ) whose position can be located anywhere in the x-axis ~~and~~

② The standard deviation σ which determines spread of its bell shaped curve.

Area property —



if $x=\mu-\sigma$

$$z = \frac{\mu-\sigma-\mu}{\sigma} = -1$$

if $x=\mu+\sigma$

$$z = \frac{\mu+\sigma-\mu}{\sigma} = 1$$

$\cancel{\int_{-\infty}^{\infty} \phi(z)\,dz = 1}$

$\cancel{2\int_{0}^{\infty} \phi(z)\,dz = 1}$

$\cancel{\int_{0}^{\infty} \phi(z}$

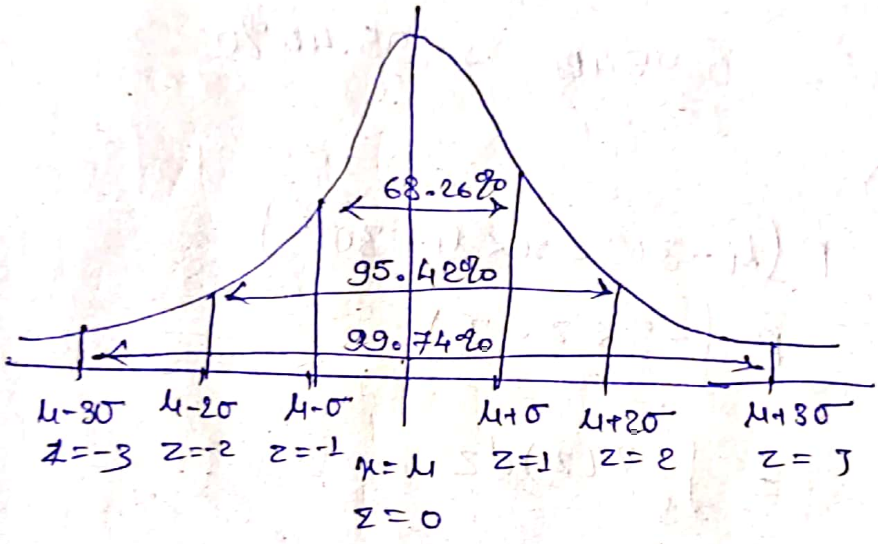
Scanned by CamScanner

Area property —

$$\int_{-\infty}^{\infty} \phi(z)\,dz = 1$$

$$2\int_{0}^{\infty} \phi(z)\,dz = 1$$

$$\int_{0}^{\infty} \phi(z)\,dz = 1/2$$



Example —

Area ① $P(\mu-\sigma < x < \mu+\sigma)$

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

$$= \int_{-1}^{1.0} \phi(z)\,dz$$

$$= 2\int_{0}^{1.0} \phi(z)\,dz$$

$$= 2(0.3413) \text{ or } 0.6826$$

$$= 68.26\%$$

Area ② $P(\mu-2\sigma < x < \mu+2\sigma)$

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

$= \int_{-2}^{2} \phi(z)\,dz$

$= 2\int_{0}^{2} \phi(z)\,dz$

$= 2(0.4772)$

$= 0.9544 \Rightarrow 95.44\%$

Area ③

$P(\mu-3\sigma < x < \mu+3\sigma)$

$P(-3 < z < 3)$

$= \int_{-3}^{3} \phi(z)\,dz$

$= 2\int_{0}^{3} \phi(z)\,dz$

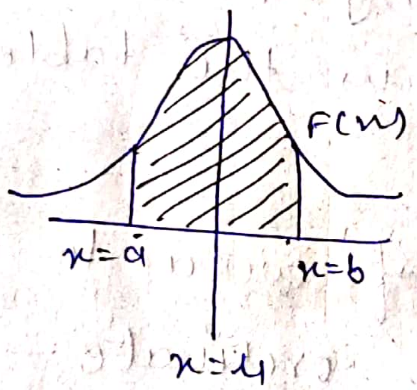
$= 2 \times 0.4987$

$= 0.9974 \Rightarrow 99.74\%$

How to compute area under normal probability curve –

Mathematically the area bounded by the curve $F(x)$ and the ordinates $x=a$, $x=b$ is given by the definate integral

$$A = \int_a^b F(x)\,dx$$



① Computation of the area to the right of the ordinate at $x=a$.
$P(x>a)$

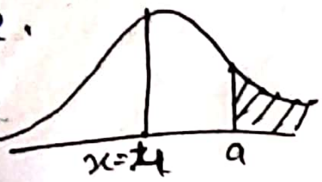
<u>Case 1:</u> $a>\mu$

$$X=a, \quad z=\frac{x-\mu}{\sigma} \Rightarrow \frac{a-\mu}{\sigma} = \cancel{z}\ z_1$$

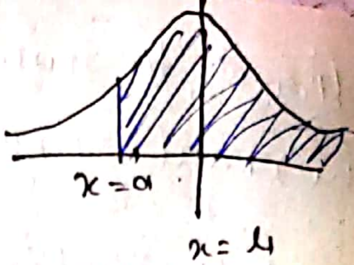
$$P(X>a) = P(z>z_1)$$

$$= \cancel{P(z}\ 0.5 - P(0<z<z_1)$$

The probablity $P(0<z<z_1)$ can be readed from the table.



Case 2 : $a < \mu$



$$z = \frac{a-\mu}{\sigma} = -z_1$$

$$p(x > a) = p(z > -z_1)$$

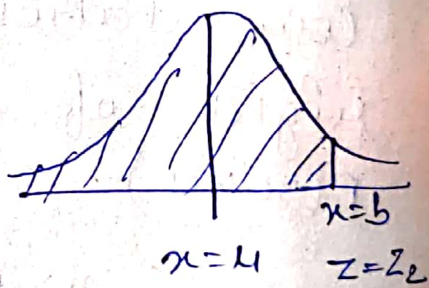
$$= p(-z_1 < z < 0) + 0.5$$

$$= 0.5 + p(0 < z < z_1)$$

the probability $p(0 < z < z_1)$ can be obtained from the table.

(2) Computation of area to the left of the ordinate –

Case (1) – $b > \mu$



$$z = \frac{x-\mu}{\sigma}$$

$$z_2 = \frac{b-\mu}{\sigma}$$

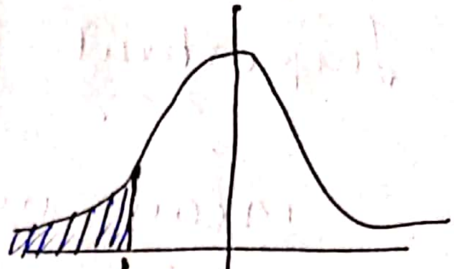
$$p(x < b) = p(z < z_2)$$

$$= 0.5 + p(0 < z < z_2)$$

The prob. $p(0 < z < z_2)$ can be obtained from the table.

Case ② : $b < \mu$

$$z = \frac{x-\mu}{\sigma}$$



$p(x<b) = P(z<-z_2)$

$= P(z > z_2)$ { Symmetry }

$= 0.5 - P(0<z<z_2)$

Q The random variable x is normal distributed with mean $\mu = 9$ and sd = 3 find the prob.

(A) ① $x \geq 15$

$x \leq 15$

$p(0 \leq x \leq 9)$

(B) Find x^* $P(x > x^*) = 0.16$

solⁿ after 3 page

Important property –

Mean, median mode of normal distribution have the same value.

Mode of the distr is that value of x for which prob. density function is maximum.

Median – the median M of the distr is defined by –

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

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

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

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

$$= \int_{M}^{\mu} f(x) + \frac{1}{2} = \frac{1}{2}$$

$$= 0 + \frac{1}{2} = \frac{1}{2} \qquad \left\{\because \int_{M}^{\mu} f(x)\,dx = 0\right.$$

$$\boxed{\mu = M}$$

The points of inflaction are obtained by $f''(x) = 0$.

Points of inflaction are ~~$\mu+\sigma$~~ $(\mu \pm \sigma)$

The moment generating function of normal distribution

The moment generating function is given by —

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Let, $\therefore z = \frac{x-\mu}{\sigma}$, $dx = \sigma dz$

$\therefore x = \mu + \sigma z$

$$\Rightarrow \frac{1}{\cancel{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu+\sigma x)} \cdot e^{-z^2/2} \cdot \cancel{\sigma}\, dz$$

$$\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\mu} \cdot e^{t\sigma z} \cdot e^{-z^2/2} dz$$

$$\Rightarrow \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{z^2}{2} - t\sigma z\right)} \cdot dz$$

$$\Rightarrow \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2t\sigma z + t^2\sigma^2 - t^2\sigma^2)} dz$$

$$\left\{ \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \right.$$

$$\Rightarrow \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z - t\sigma)^2} \cdot e^{\frac{t^2\sigma^2}{2}} dz$$

$$= \frac{e^{t\mu + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{z - t\sigma}{\sqrt{2}}\right)^2} dz$$

$$\because \text{put } y = \frac{z - t\sigma}{\sqrt{2}} \Rightarrow dz = \sqrt{2}\, dy$$

$$\Rightarrow \frac{e^{\mu t + \frac{\sigma^2 t^2}{2}}}{\cancel{\sqrt{2}}\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} \cancel{\sqrt{2}}\, dy$$

$$\boxed{M_x(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}}$$

$$M_x(t) = 1 + \frac{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}{1!} + \frac{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)^2}{2!} + \cdots$$

$$M_x(t) = 1 + t\left(\mu + \frac{\sigma^2 t}{2}\right) + \frac{t^2}{2}\left(\mu + \frac{\sigma^2 t}{2}\right)^2 + \cdots$$

$$\boxed{\mu_1' = \mu + \frac{\sigma^2 t}{2}}$$

$$\boxed{\mu_2' = \left(\mu + \frac{\sigma^2 t}{2}\right)^2}$$

Moment generating function for standard ~~ard~~ normal variate —

$$Z \sim N(0,1)$$

$$M_Z(t) = E(e^{tz})$$

$$= \int_{-\infty}^{\infty} e^{tz} \phi(z) dz$$

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

$$= e^{t^2/2}$$

Theorem —

proof at back. Let x_i, $i = 1, 2 \ldots n$ be an independent normal variates with mean μ_i and variance σ_i^2 i.e.

$$X_i \sim N(\mu_i, \sigma_i^2) \text{ then}$$

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{n=1}^{n} a_i^2 . \sigma_i^2\right)$$

this is called additive property of the normal distribution.

Scanned by CamScanner

Particular cases —

case (1) $a_1 = 1 = a_2$, $a_3 = a_4 = \dots a_n = 0$

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

case (2) $a_1 = 1$, $a_2 = -1$, $a_3 = a_4 = \dots a_n = 0$

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

case (3) $a_1 = a_2 = \dots\dots = a_n = \frac{1}{n}$

$$\frac{1}{n}\sum_{i=1}^{n} x_i \sim N\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i \cdot \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right)$$

$$\bar{x} \sim N\left(\frac{1}{n}\sum_{i=1}^{n}\mu_i \cdot \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2\right)$$

If x_i is normal independent variate with mean μ and variance σ^2, then the mean $\bar{x}$ is also a normal variate with

$$\bar{x} \sim N(\mu, \sigma^2)$$

Sol $\mu = 9, \sigma = 3$

(1) $P(X \geq 15)$

$$Z = \frac{X-\mu}{\sigma} = \frac{15-9}{3} = 2$$

$$P(Z \geq 2) = 0.5 - P(Z < 2)$$

$$= 0.5 - P(0 < Z \leq 2)$$

$$= 0.5 -$$

(2) $P(X \leq 15) \Rightarrow 1 - P(X \geq 15)$

$$1 - P(x \geq 15)$$

(3) $P(0 \leq X \leq 9)$

$$P(-3 \leq Z \leq 0)$$

$$= P(0 < Z \leq 3)$$

(3) $P(X > x^*) = 0.16$

$$P(Z > Z_1) = 0.16$$

$$0.5 - P(Z \leq Z_1) = 0.16$$

P.T.O

$$0.5 - 0.16 = P(Z \leq Z_1)$$

$$P(Z < Z_1) = 0.34$$

$$Z_1 = 1$$

$$\frac{x - 9}{8} = 1,$$

$$\boxed{x = 1}$$

Q Assume the mean height of the soldiers to be 68.22 inches with variance of 10.8 inches how many soldiers in regiment of 1000 would you expect to be –

Ans

1. Over 6ft tall $\Rightarrow$ ~~158~~ 125
2. Below 5 ft. $\Rightarrow$ ≈ 250

Q The average test marks in a particular class is 79 the sd is 5 if the marks are distr normally how many students in a class of 200 did not receive the marks b/w 75 and 8[illegible]

1. $P(75 < x < 82)$
2. $1 - P(75 < x < 82)$

Scanned by CamScanner

Theorem proof

$\sum_{i=1}^{n}$

It is given that x and y are independent normal variate with mean 2 and 5 and variance 4 and 9 find the value of k such that.

$$p(x+y \le k) = P(3x - y \ge 3k)$$

$$X \sim N(2,4) \qquad P(U \le k) = P(V \ge 3k)$$

$$Y \sim N(5,9)$$

let, $U = X+Y$, $V = 3X - Y$

$$U = X+Y \sim N(7,13)$$

$$V = 3X - Y \sim N[3(2) - 1(5), 3^2(4) + 1(9)]$$

$$U \sim N(7,13)$$

$$V \sim N(1,45)$$

var$(3x-4)$

$9\,\text{var}(x)$

$+ \text{var}(y)$

$36 + 9$

$$Z_1 = \frac{U-7}{\sqrt{13}}, \qquad Z_2 = \frac{V-1}{\sqrt{45}}$$

$\because U = k$ $\qquad$ $\because V = 3k$

$$Z_1 = \frac{k-7}{\sqrt{13}} \qquad Z_2 = \frac{3k-1}{\sqrt{45}}$$

Now, $Z_2 = -Z_1$

$$\frac{3k-1}{\sqrt{45}} = -\left(\frac{k-7}{\sqrt{13}}\right)$$

$$k = 6.28$$

Scanned by CamScanner

Skewness –

Skewness is reffers to assymetry of a distribution it distribution with ~~a~~ asymmetric ~~tate~~ tail extending out to the right is called +vely skewed. or skewed to the right. while distribution with an assymetric tail extending out to the left is reffered as -vely skewed or skewed to the left. Skewness can range ~~for~~ from $-\infty$ to $+\infty$

Carl Pearson's first suggested measuring the skewness given by –

$$\boxed{S_K = \frac{\mu - \text{mode}}{\sigma}}$$

Scanned by CamScanner

Population modes are not well estimated from sample mode Therefore -

$$\boxed{S_K = \frac{3(\text{Mean} - \text{Median})}{\sigma}}$$

The co-efficient of skewness is given by —

$$\boxed{\beta_1 = \frac{\mu_3^2}{\mu_2^3}}$$

$\beta_1 > 0$ +ve skew , $\beta_1 < 0$ −ve skew

$\mu_2 = E(x-\mu)^2 = E(x^2 - 2\mu x + \mu^2)$ $\because E(x) = \mu$

$= E(x^2) - 2\mu.\mu + \mu^2$

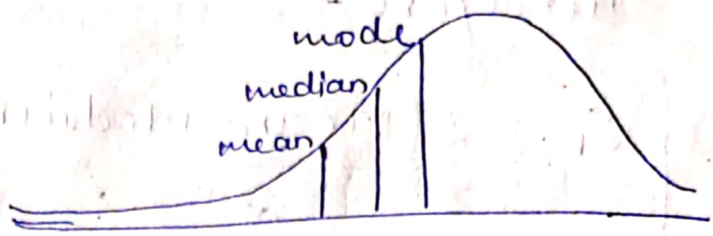
$= E(x^2) - \mu^2$

$$\boxed{\mu_2 = \mu_2' - \mu^2}$$

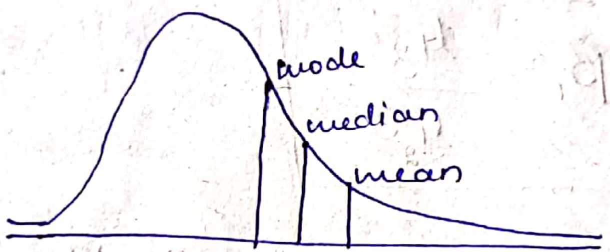
$\mu_3 = E(x-\mu)^3 = E(x^3 - 3\mu E(x^2) + 3\mu^2.\mu - \mu^3)$

$$\boxed{\mu_3 = E(x^3) - 3\mu E(x^2) + 2\mu^3}$$

Mode > Median > Mean



Negetive Skew.

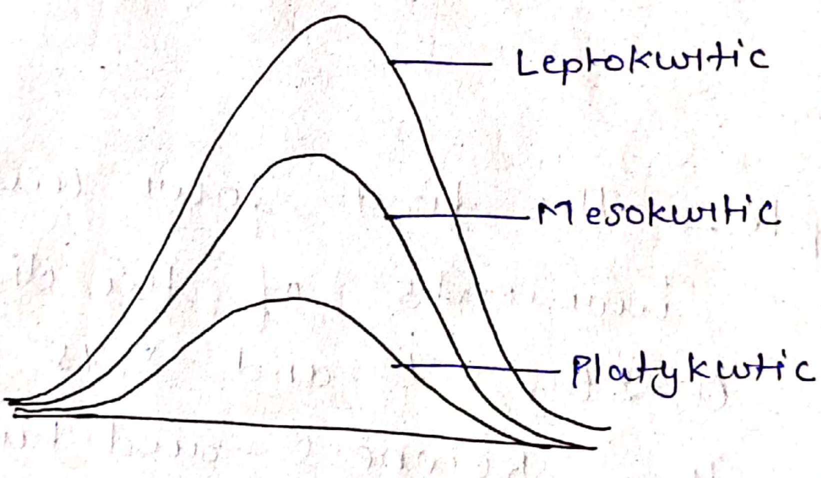


Mode < Median < Mean

Positive Skew

Kurtosis —

Kurtosis is measure of flatness or peakness of a distribution the normal curve or bell shaped curve is called mesokurtic. The curve which is more flat than the normal curve is called platykurtic. The curve which is more peak than the normal curve is called leptokurtic.



The coefficient of kurtosis is given by —

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

If $\gamma_2 = 0$ or $\beta_2 = 3$ the curve is called mesokurtic.

Ex – The Normal dist.

If $\gamma_2 > 0$ or $\beta_2 > 3$ then curve is calle Leptokurtic.

If $\gamma_2 < 0$ or $\beta_2 < 3$ the curve is called Platykurtic.

Example – The first four central moments of the distributio are 0, 2.5, 0.7 and 8.75 test the skewness and kurtosis of the distribution.

$\mu_1 = 0$, $\mu_2 = 2.5$, $\mu_3 = 0.7$, $\mu_4 = 8.75$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(0.7)^2}{(2.5)^3} =$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{8.75}{(2.5)^2} = 1.4$$

Q. Find $\mu_1, \mu_2, \mu_3, \mu_4$ for the uniform distribution and calculate the coefficient of skewness and kurtosis.

The first four moments of the distribution above the mean=4 of a random variable are
-1.5, 17, -30, 108

find $\mu_1', \mu_2', \mu_3', \mu_4'$

$\Rightarrow \mu_1' = E(x-\mu) = -1.5$

$= E(x) - \mu = -1.5$

$= E(x) = 4 - 1.5$

$\boxed{\mu_1' = 2.5}$

$\Rightarrow \mu_2' = E(x-\mu)^2 = 17$

$= E(x^2 - 2\mu x + \mu^2) = 17$

$= E(x^2) - 2(4)(2.5) + 16 = 17$

$\mu_2' =$

$\Rightarrow \mu_3' = E(x-\mu)^3 = -30$

$E[x^3 - 3x^2\mu + 3x\mu^2 - \mu^3]$

$\mu_3' =$

$$\mu_4' = E(x-\mu)^4$$

$$= E\left[x^4 - 4x^3\mu + 6x^2\mu^2 - 4x\mu^3 + \mu^4\right]$$

Chebyshev's Theorem

If 'μ', 'σ' or the mean and standard deviation of a random variable X then, for any +ve constant K, the probability is $\left(1-\frac{1}{K^2}\right)$ that of X. will take on a value within 'K' standard deviations of the mean i.e.

$$\boxed{P(|X-\mu| < K\sigma) \geq 1-\frac{1}{K^2}} \quad , \sigma \neq 0$$

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

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

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

$\because \int_a^b f(x)dx$ means $a < x < b$

similarly

$x < \mu - K\sigma \Rightarrow x-\mu < -K\sigma$

$-(x-\mu) > K\sigma \Rightarrow (x-\mu)^2 > K^2\sigma^2$

Similiarly

$x > \mu + K\sigma$

$(x-\mu)^2 > K^2\sigma^2$

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

$$\cancel{\sigma^2} \geq k^2 \cancel{\sigma^2}\left[\int_{-\infty}^{\mu-k\sigma} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} f(x)\,dx\right]$$

$$\frac{1}{k^2} \geq \int_{-\infty}^{\mu-k\sigma} f(x)\,dx + \int_{\mu+k\sigma}^{\infty} f(x)\,dx$$

$$\frac{1}{k^2} \geq P(x \leq \mu - k\sigma \text{ or } x > \mu + k\sigma)$$

$$P(x-\mu \leq -k\sigma \text{ or } x-\mu > k\sigma) \leq \frac{1}{k^2}$$

$$\boxed{P(|x-\mu| > k\sigma) \leq 1/k^2}$$

$$\cancel{1 - P(|x-\mu| \geq k\sigma) \leq \frac{1}{k^2}}$$

$$\boxed{1 - P(|x-\mu| > k\sigma) \leq 1 - \frac{1}{k^2}}$$

Note → The probability assign to the values of X outside the interval $[\mu - k\sigma, \mu + k\sigma]$ is atmost $\frac{1}{k^2}$.

(2) The probabilities assign to the values of X within a distance of $k\sigma$ of the mean atleast $1 - 1/k^2$.

Q. Given that, $\mu = 0$, $\sigma = 1$ atleast how much of the prob. of X lies between two units of the mean.

$\mu = 0$, $\sigma = 1$, $k = 2$

$\because$ $P(|x - \mu| < k\sigma) \geq 1 - 1/k^2$ ——— (1)

$P(|x - 0| < 2(1)) \geq 1 - 1/4$

$P(|x| < 2) \geq \frac{3}{4}$

Ans atleast $\frac{3}{4}$ of the total prob. should lie within two units of distance from mean.

Q. what is the minimum value of - $P(-3 \leq X \leq 3)$ at $\mu = 0$, $\sigma = 1$

then, From eqn - (1)

$P(|X| < 3) \geq 1 - \frac{1}{9}$

$P(|X| < 3) \geq \frac{8}{9}$

Minimum value of the prob. is $\frac{8}{9}$.

Scanned by CamScanner

Note – In Chebyshev's inequality k must be greater than 1

Q What value of k guarantees that – $P(|X-7| \leq k) \geq 0.99$

for $\mu = 7$, $\sigma = 2$

On comparing

$$P\left(|x-7| < \frac{k}{2} \cdot 2\right) \geq 1 - \frac{1}{(k/2)^2}$$

$$\geq 1 - \frac{4}{k^2}$$

Then,
$$1 - \frac{4}{k^2} = 0.99$$

$$1 - 0.99 = \frac{4}{k^2}$$

$$k = 20$$

Q A random variable X has a density function $f(x) = e^{-x}$, $x \geq 0$

Show that Chebyshev's inequality give – $P(|x-1| > 2) < \frac{1}{4}$ show that the actual probability is e^{-3}.

Joint Probability Distribution -

If X and Y are two random variable the probability distribution that defines they simultaneous behaviour is called joint probability distr.

* Ex - of discrete random variable

If X and Y are descrete random variables this distribution can be described with joint probability mass function.

Ex - Year in a college versus the no. of creadits taken

* Example for Continuous random var -

If X and Y are continuous random variable this distribution can be described as joint prob. density functio.

Ex - (1) Dosage of a drug versus blood ~~count~~ compound measure
(2) Time when bus driver pics you a versus quantity of caffine in bus drives syst..

The joint probability mass function is denoted by —

$$P_{xy} = P(X=x \text{ and } Y=y) = P(X=x, Y=y)$$

which means the distribution of probability is specify by listing the probabilities associated with all possible pairs of the values of X and Y.

Calculation of probabilities from a descrete joint prob. distr —

x \ y	0	1
0	0.1	0.2
1	0.4	0.2
2	0.1	0

Q Find the probability distribution –

① $P(X+Y \geq 1)$

② $P(X=x)$

Sol

① $P(X+Y \geq 1) = P(1,1) + P(2,0) + P(2,1)$

{Joint Prob distr}

$= 0.2 + 0.1 + 0$

$= 0.3$

② $P(X=x) \Rightarrow P(X=0) = P(0,0) + P(0,1) = 0.3$

$P(X=1) = P(1,0) + P(1,1) = 0.6$

$P(X=2) = P(2,0) + P(2,1) = 0.1$

Marginal probability distribution –

If we are given a joint probability distribution for X and Y we can obtain the individual prob. distribution for X and for Y these ~~pro~~ individual prob. are called marginal prob. distribution.

Probability function (X, Y) —

If X, Y is a two dimentional random variable such that prob. of $P(X = x_i, Y = y_j) = P_{ij}$

then P_{ij} is called probability mass function of (X, Y) ordered pair. provided that the following conditions are satisfied —

① $P_{ij} \geq 0 \quad \forall \, i j$

② $\sum_j \sum_i P_{ij} = 1$

Joint prob. density function —

If X, Y is a two-dimentional random variable such that $P(X = x, Y = y) = f(x, y)$ is called joint prob. density function of (X, Y) provided $f(x, y)$ satisfy following conditions —

① $f(x, y) \geq 0 \quad \forall (x, y) \in R^2$

② $\iint_R f(x, y)\, dx\, dy = 1$

Cumulative distribution function —

If (X, Y) is two diamentional random variable ~~then~~ (descrete or continuous) then —

$$F(x,y) = P(X \le x, Y \le y)$$

is called cumulative distribution of (X, Y).

For descrete case —

$$F(x,y) = P(X \le x, Y \le y) = \sum_i \sum_j P_{ij}.$$

For continuous case —

$$F(x,y) = P(X \le x, Y \le y) = \int_{\infty}^{y} \int_{-\infty}^{x} f(x,y)\,dx\,dy.$$

Marginal probability distribution ~

$$P(X = x_i) = \sum_j P_{ij}$$

In the continuous case —

Marginal density funᶜ of X is defined by —

$$\boxed{f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy}$$

Marginal density funct^n of Y is defined by —

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

Note — Probability of —

$$p(a \leq X \leq b) = p(a \leq x \leq b, -\infty < Y < \infty)$$

$$p(c \leq Y \leq d) = p(-\infty < x < \infty, c \leq Y \leq d)$$

Scanned by CamScanner

Conditional probability distribution –

Given that –

$$\boxed{P\left(X = x_i / Y = y_i\right) = \frac{P(X = x_i, Y = y_i)}{P(Y = y_i)}}$$

or $P\left(Y = y_i / X = x_i\right)$

This is called the conditional probability of X.

Independent random variable –

"Happening of the one event doesent effect the happening of second event."

If (X, Y) is 2-D discrete random variable such that –

$$\boxed{P\left(X = x_i / Y = y_j\right) = P(X = x_i)}$$

Similary, If (X, Y) 2-D continuous random variable such that

$$\boxed{f(x, y) = f_X(x) f_Y(y)}$$

then (X, Y) are said to be indepentent random variable.

Q Three balls are drawn at random without replacement from a box containing two white, three red. four black balls if X denotes the number of white balls drawn and Y denotes the number of Red balls drawn. Find the joint probability distribution of (X, Y)

X = white = 0, 1, 2
Y = Red = 0, 1, 2, 3

Red

x \ y	0	1	2	3
0	1/21	3/56	1/7	1/84
1	1/7	2/7	1/14	0
2	1/21	1/28	0	0

→ more than three balls

$$P(X=0, Y=0) = \frac{4C_3}{9C_3} = 1/21$$

$$P(X=0, Y=1) = \frac{4C_2\, 3C_1}{9C_3} = 3/56$$

$$P(X=0, Y=2) = \frac{4C_1\, 3C_2}{9C_3} = 1/7$$

$$P(X=0, Y=3) = \frac{3C_3}{9C_3} = 1/84$$

$$P(X=1, Y=0) = \frac{2C_1 * 4C_2}{9C_3} = 1/7$$

Q Find the joint probability distr function of (x,y) given below –

		y_1	y_2	y_3	y_4	y_5	y_6	
	$x \backslash y$	1	2	3	4	5	6	Σsum
x_0	0	0	0	1/32	2/32	2/32	3/32	= 8/32
x_1	1	1/16	1/16	1/8	1/8	1/8	1/8	= 5/8
x_2	2	1/32	1/32	1/64	1/64	0	2/64	= 4/32

$= \frac{8+20+4}{32} = 1$

① $P(X \leq 1) = 7/8$

② $P(Y \leq 3) = \frac{23}{64}$

③ $P(X \leq 1, Y \leq 3) = \frac{9}{32}$

④ $P(X \leq 1 / Y \leq 3) = 18/23$

⑤ $P(Y \leq 3 / x \leq 1)$

⑥ $P(X + Y \leq 4)$

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

$$= \sum_{j=1}^{6} P(X=0, Y=y_j) + \sum_{j=1}^{6} P(X=1, y=j)$$

$$= P(x=0, y=1) + P(x=0, y=2) \cdots P(x=0, y=6)$$

$$= 0 + 0 + 1/32 + 2/32 + 2/32 + 3/32 = 7/8$$

② $P(Y \leq 3) = P(Y=1) + P(Y=2) + P(Y=3)$

$$= P(X=0, Y=1) + P(X=1, Y=1) + P(X=2, Y=1)$$

$$= \frac{23}{64}$$

(3) $p(X \le 1, Y \le 3) = p(X=0, Y \le 3) + p(X=1, Y \le 3)$

(4) $$\frac{p(X \le 1, Y \le 3)}{p(Y \le 3)} = \frac{9/\cancel{32}}{23/\cancel{64}_{\,2}} = \frac{9}{1} \times \frac{2}{23} \quad \frac{18}{23}$$

(6) $$p(X+Y \le 4) = \sum_{j=1}^{4} p(X=0, Y=y_j) + \sum_{j=1}^{3} p(X=1, Y=y_j) + \sum_{j=1}^{2} p(X=2, Y=y_j)$$

Q The joint probability –

$$p(x,y) = k(2x+3y)$$

$x = 0, 1, 2$
$y = 1, 2, 3$

Find all the marginal and conditional prob. distr also find all the prob. distr of x+y.

x \ y	1	2	3	Σ
0	3k	6k	9k	18k
1	5k	8k	11k	24k
2	7k	10k	13k	30k
				72k

then $72k = 1$

$$\boxed{k = \frac{1}{72}}$$

(A) All marginal prob –

$P(X=0)$ | $P(Y=1)$
$P(X=1)$ | $P(Y=2)$
$P(X=2)$ | $P(Y=3)$

(B) Conditional prob

① $P(X=x_i/Y=1)$

② $P(X=x_i/Y=2)$

③ $P(X=x_i/Y=3)$

Joint probabilities density function of 2-D random variable given by—

$$f(x,y) = xy^2 + \frac{x^2}{8} \qquad \begin{matrix} 0 \le x \le 2 \\ 0 \le y \le 1 \end{matrix}$$

Compute —

1. $P(X>1)$
2. $P(Y<1/2)$
3. $P(X>1 / Y<1/2)$
4. $P(Y<1/2 / X>1)$
5. $P(X<Y)$
6. $P(X+Y\le 1)$

① $P(X>1) = \int_{y=0}^{1}\int_{1}^{2} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_{y=0}^{1} \left[\frac{x^2y^2}{2} + \frac{x^3}{24}\right]_1^2 dy$$

$$= \int_{y=0}^{1} \left[\frac{4y^2}{2} + \frac{8}{24} - \frac{y^2}{2} - \frac{1}{24}\right] dy$$

$$= \int_{y=0}^{1} \frac{3y^2}{2} - \frac{7}{24}\, dy$$

$$= \left[\frac{y^3}{2} - \frac{7y}{24}\right]_0^1$$

$$= \left[\frac{1}{2} - \frac{7}{24}\right]$$

$$= 5$$

(2) $$P(Y < 1/2) = \int_0^2 \int_0^{1/2} \left(xy^2 + \frac{x^2}{8}\right) dy\, dx$$

(3) $$\frac{P(X > 1, Y < 1/2)}{P(Y < 1/2)} = \frac{\int_1^2 \int_0^{1/2} \left(xy^2 + \frac{x^2}{8}\right) dy\, dx}{P(Y < 1/2)}$$

(5) $$P(X < Y) = \int_0^1 \int_0^y \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$$

(6) $$P(X + Y \leq 1) = \int_{x=0}^{1} \int_{y=0}^{1-x} \left(xy^2 + \frac{x^2}{8}\right) dy\, dx$$

Q The joint prob. distr function of random variable (X, Y) given by —

$$f(x,y) = kxy\, e^{-(x^2+y^2)} \quad ; x>0,\ y>0$$

find the value of k and prove that X and Y are independent.

Sol

Find k,

$$\int_0^\infty \int_0^\infty k\left[xy\, e^{-(x^2+y^2)}\right] dy\, dx = 1$$

(

$F_X(x) =$

$F_Y(x)$

Scanned by CamScanner

Functions of random variables—

~~Here~~ consider the problem for finding the prob. functions or density function of two or more random variables.

Given a set of random variable $x_1, x_2 \ldots x_n$ and their Joint density will be ~~given~~ in finding the density of

$$y = U(x_1, x_2 \ldots x_n)$$

Find the density of Y by using distribution method—

Suppose $x_1, x_2 \ldots x_n$ are continuous random variable with a given probability density function.

Let $y = U(x_1, x_2 \ldots x_n)$

then, cumulative density func^n is given by,

$$F(y) = F(Y \leq y) = P[U(x_1, x_2 \ldots x_n) \leq y]$$

diff^r w.r.t y we can obtain prob. density function given by

$$f(y) = \frac{dF}{dy}$$

Q If the p.d.f of -

$$f(x) = \begin{cases} 6x(1-x) & 0 < x < 1 \\ 0 & ew \end{cases}$$

$Y = x^3$.

$F(Y) = F(Y \le y)$
$= F(x^3 \le y)$
$= F(x \le y^{1/3})$

$F(Y) = F(x \le y^{1/3})$

$$= \int_0^{y^{1/3}} 6x(1-x)\,dx \Rightarrow \left[3x^2 - 2x^3\right]_0^{y^{1/3}}$$

$F(y) = 3y^{2/3} - 2y$

$$f(y) = \frac{dF(y)}{dy}$$

$$= \frac{d}{dy}\left[3y^{2/3} - 2y\right]$$

$=$

Q The joint density function of X_1 and X_2 is given by:

$$f(x) = \begin{cases} 6e^{-3x_1-2x_2} & \text{for } x_1, x_2 > 0 \\ 0 & \text{e/w} \end{cases}$$

Find the density function $Y = X_1 + X_2$

$$F(y) = F(Y \leq y)$$

$$= F(x_1 + x_2 \leq y)$$

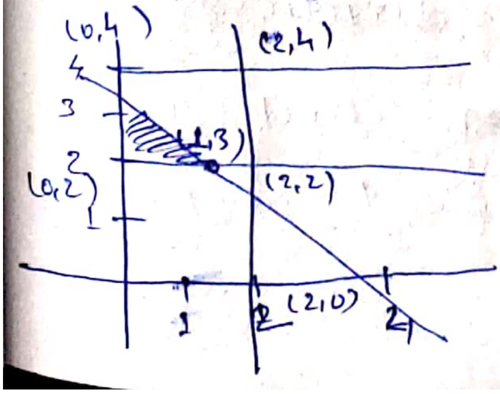
$$= \int_0^y \int_0^{y-x_1} 6e^{-3x_1-2x_2}\, dx_2\, dx_1$$

$$\left\{ \because x_1 + x_2 \leq y \atop x_2 = y - x_1 \right\}$$

Q $$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y) & 0<x<2 \\ & 2<y<4 \\ 0 & \text{e/w} \end{cases}$$

$$P(X+Y<3)$$

$$P(X+Y<3) = \frac{1}{8}\int_0^1 \int_{y=2}^{3-x} (6-x-y)\,dy\,dx$$



$x+y=3$

$y=2$

$x=1$

Testing of Hypothesis —

~~# Random Samples~~ —

The set of a hypothesis which assumes that there is no significance difference b/w the sample static and the corresponding population parameter or b/w two sample statistics.

Such a hypothesis of no difference is called a NULL hypothesis denoted by H_0.

(A hypothesis that is different (or complementary to) the NULL hypothesis is called Alternative hypothesis and denoted by H_a or H_1

A procedure for deciding whether to accept or check a NULL hypothesis is called testing of hypothesis.

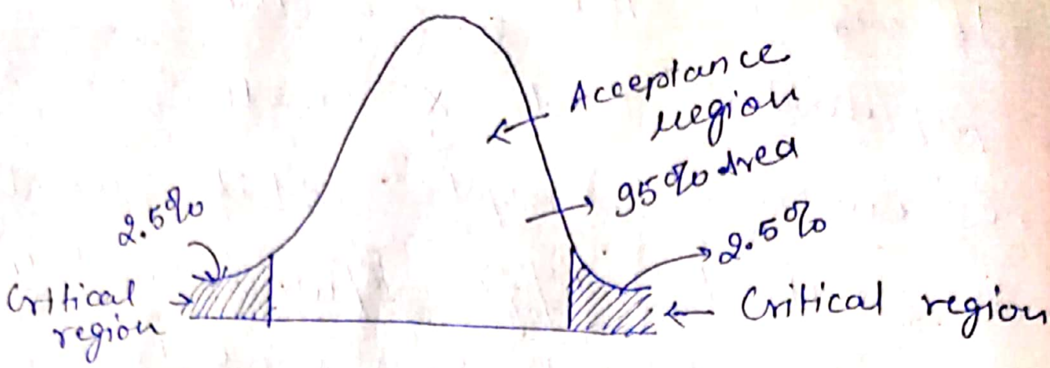
Scanned by CamScanner

If μ_0 is the parameter of the population and $\bar{x}$ is the corresponding sample static usually there will be diffr. b/w $\bar{x}$ and μ since $\bar{x}$ is based on sample observation.

Such differences is caused due to the sampling fluctuations is called insignificant difference. The procedure of testing whether the diffrence b/w μ and $\bar{x}$ is significant or not is called test of significance.

Critical region and level of significance –

If we are prepared to accept the difference b/w the sample static and the corresponding parameter i.e diff^r is significant when the sample static lie in a certain region of interval then that region is called the critical region or region of rejection.



The region complementary to the Critical region is called the acceptance region in the case of the large samples. The sampli distribution of many statics tend to become normal distribution.

If t is a static in a large sample then t follows the normal distribution with mean E(t) of (the population parameter) and the standard deviation = SE(t) {standard error of the test static of sample}

Hence,

$$Z = \frac{t - E(t)}{SE(t)} \sim N(0,1)$$

Z is called standard normal variate (test statics)

The probability that a random value of the statistics lies in the critical region is called the level of significance. and LOS is expressed in terms of the percentages. In terms of mathematical notation.

$$P\left(\left|\frac{t-E(t)}{S.E(t)}\right| < 1.96\right) = 0.95$$

$$P(|Z| > 1.96) = 0.05 \quad \{\text{Rejection region}\}$$

Errors in testing of Hypothesis —

There are two types of error in testing of Hypothesis.

Type-1 Rejecting the NULL Hypothesis when the NULL Hypothesis is true

It is also called as consumers risk.

Type-2 Accepting the NULL Hypothesis when H_0 is False. This is also called as producers risk.

⇒ Prob. of Type ① error —

$$= P\left(\frac{\text{Rejecting } H_0}{H_0 \text{ is true}}\right) = \alpha$$

→ Prob. of Type ② error —

$$= P\left(\frac{\text{Accepting } H_0}{H_0 \text{ is false}}\right) = \beta$$

<u>Note</u> — The prob. α is committing the type ① error is also called the level of significance.

One tailed test and two tailed test.

~~If μ is a population parameter~~

~~and~~

Let θ_0 is population parameter and t is the corresponding sample static and set of the NULL hypothesis is —

$$H_0 : t = \theta_0$$

eq^n ① — $H_1 : t \neq \theta_0 ; t > \theta_0, t < \theta_0$

② — $H_1 : t > \theta_0$

③ — $H_2 : t < \theta_0$

⇒ Alternative Hypothesis H_1 given in eqⁿ (1) is called two tailed test or two sided alternative hypothesis test.

⇒ eqⁿ (2) and (3) are called one tailed test or one side alternative hypothesis test.

⇒ eqⁿ - (2) is also called right tailed alternative hypothesis test.

⇒ eqⁿ - (3) is also called left tailed alternative hypothesis test.

The application of the one tailed and two tailed test depends upon the nature of the problem of the alternative hypothesis.

Z_α = Level of significance (LOS)

Nature of test	1% (0.01)	2% (0.02)	5% (0.05)	10% (0.1)
Two-tailed	$\lvert Z_\alpha\rvert = 2.58$	$\lvert Z_\alpha\rvert = 2.33$	$\lvert Z_\alpha\rvert = 1.96$	$\lvert Z_\alpha\rvert = 1.645$
Right tailed	$Z_\alpha = 2.33$	$Z_\alpha = 2.055$	$Z_\alpha = 1.645$	$Z_\alpha = 1.28$
Left tailed	$Z_\alpha = -2.33$	-2.055	-1.645	-1.28

Procedure for testing of Hypothesis—

Step ① Define the NULL Hypothesis H_0

Step ② Define the Alternative Hypothesis H_1

Step ③ Fix a suitable level of significance (LOS) α which depends on the particular problem, decide the test to be used. ~~which means~~

Step ④ Compute the test static,

$$z = \frac{t - E(t)}{SE(t)}$$

Step ⑤ Compare the values of $|z|$ and Z_α. If $|z| < Z_\alpha$, accept H_0 reject H_1 at (LOS) at α

Test of Significance for large Samples —

Assumptions —

1. The sampling distribution of a statistics is approx. normal irrespective of whether the dist^r is normal or not.

2. Sample statistics are sufficiently closed to the corresponding population parameter and hence may be used to calculate the standard error of the sampling distribution.

Test ① – Test of Significance of the difference b/w sample proportion and Population proportion ~

$$\boxed{Z_{\text{Statistics}} = \frac{p - P}{\sqrt{\frac{PQ}{n}}}}$$

∵ P = population proportion

p = sample proportion

~~where~~ If P is not known assume that p is nearly equal to P, Hence Z statistics is —

$$\boxed{Z_{\text{Statistics}} = \frac{p - P}{\sqrt{\frac{pq}{n}}}} = 0$$

Scanned by CamScanner

Test (2) - Test of significance of the difference b/w two sample proportion —

$$\boxed{Z_{statistics} = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

Suppose x follows the Bernoulli distr, small p_1 and p_2 are proportions of success in two large samples of size n_1 and n_2.

Test (3) - Test of significance of the sample mean $(\bar{x})$ and population mean (μ) -

$$\boxed{Z_{statistics} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}}$$

where σ is standard devit, n is sample size

Type (4) Test of significance of differens b/w means of two samples —

$$\boxed{Z_{statistics} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}}$$

Suppose $\sigma_1 = \sigma_2 = \sigma$ (both the samples are drawn from same population) then

$$\boxed{Z_{statistics} = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}$$

Type ⑤ Test of significance of the diffr b/w sample standard deviation and population standard deviation.

$$\boxed{Z_{statistics} = \frac{S - \sigma}{\frac{\sigma}{\sqrt{2n}}}}$$

where n is the sample size

Test ⑥ – Test of significance of the diffr b/w standard deviations of the two large samples.

$$\boxed{Z_{statistics} = \frac{S_1 - S_2}{\sigma\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}}}$$

Suppose σ (SD) is not known then it can be approximated by –

$$\boxed{\sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2}}}$$

Q In a large lot of electric buld the mean life and the standard deviation of the bulb are 360 hr and 90 hr respectively. If sample of 625 bulbs is choosen it is found that the mean life and the SD of the bulb in the sample are 350 hr and 90 hr respectively. Can we conclude that the sample is drawn from the given population. Test at 5% level of significance.

Sol

$\mu = 360, \ \sigma = 90$

$n = 625$

$x = 350, \ S = 90$

Assume $\bar{x} = \mu \rightarrow H_0$

$\bar{x} \neq \mu \rightarrow H_a$

(two tailed test)

Table $\rightarrow \alpha = 5\%$, $Z_\alpha = 1.96$

$$Z_{statistic} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{350 - 360}{\frac{90}{\sqrt{625}}}$$

$= -2.77$

$|Z| > Z_\alpha$

∴ Reject the null Hypothesis H_0 at 5% level of significance which means the sample is not drawn from the given population.

Sizes and means of two independent random samples are 400, 225, 3.5, 3.0 respectively. Can we conclude that the samples are drawn from same population with SD 1.5

Sol

$$n_1 = 400, \quad n_2 = 225$$

$$\bar{x}_1 = 3.5, \quad \bar{x}_2 = 3.0$$

$$\sigma = 1.5$$

$H_0 \rightarrow \bar{x}_1 = \bar{x}_2$ (Two failed test)

$H_a \rightarrow \bar{x} \neq \bar{x}_2$

$$Z_\alpha = \frac{x_1 - x_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} =$$

For any α Reject the NULL Hy. H_0 i.e. sample is not drawn from same population.

Q The number of student in a class is 100 the average marks scored by 64 boys is 66 with S.D 10 while the average marks scored by 36 girls is 70 with S.D 8 test at 1% level of significance whether girls are performed better than boys.

$n_1 = 64 \quad \bar{x}_1 = 66 \quad \sigma_1 = 10 = S_1$

$n_2 = 36 \quad \bar{x}_2 = 70 \quad \sigma_2 = 8 = S_2$

$H_0 \rightarrow \bar{x}_1 = \bar{x}_2$

$H_1 \rightarrow x_2 > x_1$

$x_1 < x_2$

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \cancel{\frac{-4}{\sqrt{0.96+}}} = \frac{-4}{\sqrt{1.51+0.92}}$$

$$= -2.56$$

$2.18 > -2.33$

reject

Q A manufacturer of an electric item finds the standard deviation of the life of the item to be 60 hrs. The manufacturer wants to adopt a new process for producing the same item which improves the life of the item. A random sample of 200 item produced by the new process is selected and it was found that the SD σ is 52hrs. Should the manufacturer adopt the new process test of 2% LOS.

$\sigma = 6$ hrs, $n = 200$, $S = 52$ hrs.

$H_0 \rightarrow S = \sigma$

$H_1 \rightarrow S < \sigma$ (Right hand test)

$\alpha = 2\% = 0.02$, $Z_2 = 2.055$

$$Z_{stat} = \frac{S.\sigma}{\frac{\sigma}{\sqrt{2n}}} = \frac{52-60}{\frac{60}{\sqrt{400}}} = -2.66$$

Then, reject H_0 and accept H_1 at 2% LOS. So he adopt the new process.

Scanned by CamScanner

Q Two independent random samples of sizes 1200, 800 are drawn the S.D of the samples are found to be 36 and 37 respe. find the possibility that the two samples are drawn ~~for~~ from population of same SD test at 5% LOS.

$x_1 = 1200, \quad x_2 = 800$

$S_1 = 36, \quad S_2 = 37$

$H_0 \rightarrow \sigma_1 = \sigma_2$

$H_1 \rightarrow \sigma_1 \neq \sigma_2$

LOS, $\alpha = 5\%$, $Z_\alpha = 1.90$

$$Z_{stats} = \frac{S_1 - S_2}{\sigma\sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}}$$

$$\sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2}}$$

$\sigma = 36.40$

$$Z = \frac{-1}{36.40\sqrt{}} = -0.85$$

Samples are drawn from population with same SD, Accept H_0 at 5% LOS.

Q The mean breaking stregth of the cable supplied by the manufacturer is 1800 with SD of 100 by new technique in the manufacturing process it is claimed that the breaking strength of cable is increased to test this claim a sample of 50 cables is tested and it is found that the mean breaking strength is 1850 can we support the claim at 1% LOS.

$\mu_1 = 1800, \quad \sigma_1 = 100$

$\overline{x_0} = 1850.$

$H_0: \mu = x$

Q In a large city A, 20% of random samples of 900 school ~~body~~ boys had a slight physical defect in another large city B, 18.5% of a random sample of 1600 school boys had the same defect & diff^r b/w the proportion test at 5% LOS.

$A = 20\%$, 900

$B = 18.5\%$, 1600

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} , \quad Q = 1 - P$$

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

Scanned by CamScanner

Student t-test —

Testing of significance for small cycles —

1. Student's t-test
2. F-test

These two test are applicable for the sample size $n \leq 30$.

Definition — A t-test is a form of the statistical hypothesis test based on students t-statistics and t-distribution to find the p-value. which can be used to accept or reject the NULL hypothesis, t-test analyzes the mean of the two samples test are greatly diffr with each other.

Assumptions —

1. the population is infinite and normal.
2. the population is unknown and estimated from the sample.

③ the mean must be known.

④ The sample observations are independent and random.

⑤ The sample size is small.

⑥ H_0 must be one sided or two sided.

Mean and standard deviation of the two samples are used to make the comparission b/w them such that —

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$\bar{x}_1, \bar{x}_2$ are means of the samples.

⇒ The t-distribution is given by —

$$F(t) = \frac{1}{\sqrt{\nu}\, \beta\left(\frac{1}{2}, \frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

where ν is called the degrees of freedom of the t-distributions and it is denoted by —

$$t \sim t_\nu$$

The mean of t-distribution is zero and the variance of the t-dist^r is

$$\sigma^2 = \frac{\nu}{\nu-2}, \quad \nu>2,$$ the variance goes to 1 as ν goes to infinity.

The $f(t)$ is symmetric about $t=0$.

Applications of t-distribution.

1. To test the significance of the difference b/w the mean of a small sample and the mean of population.
2. To test the significance of the diff^r of means of the two small random samples.
3. To test the significance of diff^r of the correlation in a small sample.

Degrees of freedom —

The no. of degrees of freedom of a statistics is the no. of independent variates using to computed this statistics.

Note : 1) $t(1\% \text{ LOS}) = t(0.01)$ for a single tail test

$= t(2\% = 0.02)$ for two tailed test.

2) $t(0.1)$ for two tailed test =

$t(0.05)$ for single tailed test.

Test 1:- Test of significance for T-distribution of difference b/w mean of small sample or mean of population

Assumption: The random sample drawn from normal population the standard deviation (σ) of parent population is not known.

$$t_{statistic} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

t doesn't follow the normal distribution $N(0, 1)$

but t follow a T distribution with ν degree of freedom denoted by $t \sim t_{\nu}$

$t \sim t_{n-1}$

$\bar{x} = \frac{1}{n}\Sigma x_i$, $s^2 = \frac{1}{n}\Sigma_i (x_i - \bar{x})^2$

Test 2 :- Test of significance of difference b/w the means of two small random samples from the same normal population.

Assumption : Parent population from which samples are drawn is normally distributed. The population variance are equal but unknown

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad \text{where} \quad \sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2}$$

with degree of freedom $t \sim t_{n_1+n_2-2}$

or $t \sim t_v$, $v = n_1+n_2-2$

If sample sizes are equal say,

$n_1 = n_2 = n$

then

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{2/n}} \quad \sigma^2 = \frac{n(s_1^2+s_2^2)}{2n}$$

$$= \frac{s_1^2+s_2^2}{2}$$

Q. The annual rainfall at a certain place is normally distributed with the mean. 45cm the rainfall during last 5 years are 48 cm, 42 cm, 40 cm, 44 cm, 43 cm. Can we conclude that the average rainfall during last 5 years is less than normal rainfall. Test at 5% level of significance

$\Rightarrow$ $\bar{x} = 43.4$ cm

$$s^2 = \frac{1}{n}\Sigma(x_i-\bar{x})^2 = \frac{21.16 + 1.96 + 11.56 + 0.36 + 0.16}{5}$$

$n = 5$

$\nu = 4$ $\{n-1 = 5-1 = 4\}$

t_4 (5% LOS) for single tail test =

t_4 (0.1 LOS) for two tailed test

$= 2.132$

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}} = \frac{43.4-45}{\left(\frac{2.65}{2}\right)} = -1.13$$

$|t| < t_4$

So, ~~appe~~ accept the null hypothesis

$\therefore$ Rainfall is same as last 5 years

Q. A random sample of 17 values from a normal population as a mean of 105 cm and the sum of squares of deviation from the mean is 1225 cm is the assumption of mean of 110 cm for the normal population reasonable.
Test under 5% and 1% level of significance.

$\Rightarrow$ $n = 17$ $\quad \nu = 16$

$\Sigma(x-\bar{x})^2 = 1225$

$s^2 = 72.05$

$s = 8.48$

~~H~~ $H_0 : \mu = 110$ cm

$H_a : \mu \neq 110$ cm

Two ~~tail~~ tail test.

$t_{16}(5\% \text{ LOS}) = 2.120$

$t_{16}(1\% \text{ LOS}) = 2.92$

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} = \frac{105-110 \times 4}{8.48} = \frac{(-5) \times 4}{8.48}$$

$$= \frac{-20}{8.48}$$

$$= -2.358$$

$|t| > t_{16}$ at 5% LOS

$|t| < t_{16}$ at 1% LOS

Q Following two independent random samples of sizes 10 and 8 are given

sample 1	sample 2
165	160
162	164
164	163
169	168
170	171
166	162
165	166
163	163
168	
171	

is the difference b/w sample mean is significant at 5% level of significance

$\Rightarrow \quad H_0 : \bar{x}_1 = \bar{x}_2 \qquad H_a : \bar{x}_1 \neq \bar{x}_2$

$\bar{x}_1 = 166.3$

$\bar{x}_2 =$

$s_1^2 =$

$s_2^2 =$

$t \sim t_v = t_{n_1+n_2-2}$

$t \sim t_{16} \; (5\%, .05)$.

Ans: accept the H_0 hypothesis.

F- Test is described with two degree of freedom ν_1, ν_2 and denoted by $F \sim (\nu_1, \nu_2)$

F-test depends on two degree of freedom ν_1 and ν_2

It always a single tail test.

Application of F-test

1) we use F test wethear the difference b/w the population variance are significant or not

To test the significance.

estimate the population variant σ_1^2, σ_2^2 Based on small variance S_1^2, S_2^2 respectively

$$\sigma_1^2 = \frac{n_1 s_1^2}{(n_1-1)} \qquad \nu_1 = n_1 - 1$$

$$\sigma_2^2 = \frac{n_2 \cdot s_2^2}{n_2-1} \qquad \nu_2 = n_2 - 1$$

f statistics is defined by

$$F = \frac{\sigma_1^2}{\sigma_2^2}$$

with ν_1 and ν_2 degrees of freedom

Q. Two random samples of sizes 9 & 7 gave the sum of squares of deviation from their means is 175 and 95 respectively can they be regarded as drawn from the Normal population with same variance

$n_1 = 9 \quad \Sigma(x_1-\bar{x}_1)^2 = 175 = n_1 s_1^2$

$n_2 = 7 \quad \Sigma(x-\bar{x}_1)^2 = 95 = n_2 s_2^2$

$\sigma_1^2 = \frac{175}{8}$, $\sigma_2^2 = \frac{95}{6}$

$= 21.875 \qquad = 15.833$

$F = \frac{\sigma_1^2}{\sigma_2^2} = 1.38$

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_a : \sigma_2^2 < \sigma_1^2$

$F(\nu_1, \nu_2) = F(8, 6) = 4.15$ (5% LOS)

$\Rightarrow F(8, 6) = 8.10$ (1% LOS)

$F < 4.15$

we should accept H_0
Sample can be drawn from same Normal population

Q Two random samples of sizes $n_1 = 9$, $n_2 = 6$ give the following value of variable

sample 1	sample 2
15	8
22	12
28	9
26	16
18	15
17	10
29	
21	
24.	

Test the difference of the estimate of population variance at 5% level of significance.

$\bar{x}_1 = 22.22$ $\qquad$ $\bar{x}_2 = 11.66$

$\sigma_1^2 = 52.12 + 0.048 + 33.40 + 14.28 + \cancel{14.28}^{17.80} + 27.24$
$+ 45.96 + 1.48 + 3.16$

Remark : 1. To test if two small random samples have been drawn from the same normal population.

we must apply both f-test and t-test ie test equality of population variance if equality of population variance are shown (accept H_0)

then apply t-test significance of difference two small sample means.

Q The values of two random sample are given here

Sample 1	Sample 2
15	35
25	31
16	25
20	38
22	26
24	29
21	32
17	34
19	33
23	21
	29
	31

Can we conclude that two sample are drawn from same population test at 5% LOS.

$\Rightarrow$ $\nu_1 = n_1 - 1 = 10 - 1 = 9$

$\nu_2 = n_2 - 1 = 12 - 1 = 11$

$\bar{x}_1 = 20.2$ $\quad s_1^2 = 10.56$

$\bar{x}_2 = 30.83$ $\quad s_2^2 = 13.641$

$\sigma_1^2 = 11.73$

$\sigma_2^2 = 14.88$

$F_{(0.05)}(9, 11) = 2.86$

$H_0 = \sigma_1^2 = \sigma_2^2$

$H_a = \sigma_2^2 > \sigma_1^2$

B.S Grewal Engineering Mathematics

$F = \dfrac{\sigma_2^2}{\sigma_1^2} = 1.26$

$|F| < F_{(0.05)}(9, 11)$

accept the null hypothesis

i.e sample are drawn from same population

$H_0 : \bar{x}_1 = \bar{x}_2$

$H_a : \bar{x}_1 \neq \bar{x}_2$

$t \sim t_0 = t_{n_1+n_2-2} = t_{20}$.

$t_{20} = 2.086$

$t_{stat} = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $\qquad \sigma^2 = \dfrac{n_1 S_1^2 + n_2 S_2^2}{n_1+n_2-2}$

$= \dfrac{20.2 - 30.83}{3.66\sqrt{0.1+0.048}}$ $\qquad = \dfrac{10 \times 10.56 + 12 \times 13.64}{22-2}$

$= \dfrac{-10.6}{3.66 \times 0.428}$ $\qquad = \dfrac{269.29}{20} = \cancel{12.24}\ 13.46$

$= -6.76$ $\qquad = 3.66$

$|t| > t_{20}(0.05)$

reject H_0 so there is difference b/w sample.
Mean of two population sample and hence the population defer significantly
So, sample are not drawn from same population

Q. Two random sample gave following data.

Sample size	Mean	Variance
1	8	9.6	1.2
2	11	16.5	2.5.

Scanned by CamScanner

can we conclude that sample are drawn from same population.
LOS = 0.05.

⇒

χ^2-test {"chai"-square test}

χ^2 distribution is denoted by χ^2-distribution

If a normal variable then $X \sim N(\mu, \sigma^2)$ then std normal variate

$$z = \frac{x-\mu}{\sigma} \sim N(0,1)$$

The square of normal variate

$z^2 = \left(\frac{x-\mu}{\sigma}\right)^2$ is called χ^2-variate with one degree of freedom.

Generating if x_i $i = 1, 2, \ldots n$ are n independent normal variate then $N(\mu_i, \sigma_i^2)$

$$\boxed{\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i-\mu_i}{\sigma_i}\right)^2}$$

is a χ^2-distribution with n-degree of freedom.

Application of χ^2-distribution

i) It is used to test the goodness of fit

for eg: suppose that we have fitted a poisson distribution and binomial distribution, we use χ^2-distribution to test wether this poisson and binomial distribution of given data is acceptable.

ii) It is used to test the independence of attribute of population.

χ^2-test for Goodness of fit :-

Let a distribution be given.

Suppose ~~tha~~ that O_i and E_i $i = 1, 2, 3 -- n$ are the ~~obj~~ observed and expected frequency of ith class with

$\Sigma O_i = \Sigma E_i$

The expected frequency is computed using the hypothesis assumed.

then.
$$\boxed{\chi^2 = \sum_{i=1}^{n} \left[\frac{(O_i - E_i)^2}{E_i} \right]} \text{ with } \sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i \quad \text{—— (1)}$$

follows χ^2-distribution with $(n-1)$ degree of freedom.

This formula doesn't contain any parameters hence it is called non-parametric test.

Also frequency distribution doesn't follow the normal distribution

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i^2 + E_i^2 - 2O_iE_i)}{E_i}$$

$$= \sum \left[\frac{O_i^2}{E_i} + E_i - 2O_i \right]$$

$$= \sum \frac{O_i^2}{E_i} + \sum E_i - 2\sum O_i$$

$$= \sum_{i=1}^{n} \frac{O_i^2}{E_i} - \sum O_i$$

$$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{O_i^2}{E_i} - N} \quad \text{—— (2).}$$

6/Nov/18.

i) It is one tailed test

ii) formula (2) is particularly useful when expected frequencies E_is and consequently the deviations $(O_i - E_i)$ are in decimals.

Condition of for χ^2 test

1) No. of total frequencies should be resonably large. ie $N > 50$

2) Sample observations should be independent

3) No theoritical frequency is small $\Sigma O_i - \Sigma E_i$ {small is relative term}

4) If any frequency is small use the ~~techniq~~ technique of pooling. (which consist in adding the frequencies which is small with the preceding or succesive frequency so that resulting sum is large.) in this case degree of freedom also change.

- The following figure show the distribution of digit in number choosen at randomly from a Telephone Directory.

Digit	0	1	2	3	4	5
Observed Frequency	1026	1107	997	996	1075	933

Digit	6	7	8	9
Observed Frequency	1107	972	964	853

Test whether digits may be taken occur equally frequently in the directory. The table value of χ^2 9 degree of freedom at 5% level of significance is 16.92.

$\Rightarrow \quad \Sigma f = 10030$

Digits	Observed frequency O_i	Expected frequency $E_i \left(\frac{10030}{10}\right)$	$(O-E)$	$(O-E)^2$
	1026	1003	23	529
	1107	1003	104	10816
	997	1003	-6	36
	996	1003	-7	49
	1075	1003	72	5184
	933	1003.	-70	4900
	1107	1003	104	10816
	972	1003	-31	961
	964	1003	-39	1521
	853	1003.	-150	22500
	10030	10030.		57312

H_0: The digit 0, 1, 2, --- in number occur equally frequently

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E} = \frac{57312}{1003} = 57.14.$$

Here, $\chi^2 > \chi^2_9(0.05)$

Reject H_0 at 5% LOS which means the digit in no. doesnot occur equally frequently in directory

Q. Records taken of the nos of males and females births in 800 families having 4 children are given in table. Test whether the data are consistent with the hypothesis that binomial law holds and the chance of male birth = chance of female birth.

No. of births		O_i				
Male	Female	Frequency	E_i	$O-E$	$(O-E)^2$	$\frac{(O-E}{N}$
0	4	32	50	−18	324	64
1	3	178	200	−22	484	24
2	2	290	300	−10	100	0.3
3	1	236	200	36	1296	64
4	0	64.	50	14	196	39
		800	800		2400	

H_0: The distribution follows the binomial law.

$$p(x) = 800\ {}^4C_x \left(\tfrac{1}{2}\right)^x \left(\tfrac{1}{2}\right)^{4-x}.$$

$$= \frac{800}{16}\ {}^4C_x = 50\ {}^4C_x.$$

$\frac{4\times3}{2\times1}$

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i} = 19.63.$$

$x^2(0.05) = 9.488$

$x^2 > x^2(0.05)$

Reject. Ho. null hypothesis

Q. The following mistake per page observed in book given in table fit the poisson distribution and test the goodness of fit.

⇒

No. of mistake per page (x)	No. of pages (f)	fx	E_i	$(O-E)^2$
0	211	0	209.31	2.8561
1	90	90	92.1	
2	19	38	20.3	
3	5 (24)	15	3.0	
4	0	0	0.3 (23.6)	
	325			

$\lambda = \frac{\Sigma fx}{\Sigma f} = \frac{143}{325} = 0.44$.

$E_i = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{0.644\,\lambda^x}{x!} \times 325$

$x = 0, 1, 2, 3, 4$

Note:- Degree of ~~freedre~~ freedom is lost ~~becom~~ because the parameter λ is estimated from data.

and a degree of freedom is lost because it follows of 3'cell pooling which are small.

$\chi^2(0.05) = 3.841$

Q. The figures below are.

a) The frequency of distribution

b) The frequency of normal distribution having same mean is std. deviation ~~as in~~ a. and total frequency is same as in a.

a	b.
1	2
12	15
66	66
220	210
495	484
792	799
924	944
792	799
495	484
220	210

CO-RELATION

$$\text{var}(x+y) = E[(x+y) - (\mu_x + \mu_y)]^2$$
$$= E[(x-\mu_x) + (y-\mu_y)]^2$$
$$= E[(x-\mu_x)^2 + (y-\mu_y)^2 + 2(x-\mu_x)(y-\mu_y)]$$
$$= E(x-\mu_x)^2 + E(y-\mu_y)^2 + 2E[(x-\mu_x)(y-\mu_y)]$$

$$\boxed{\text{Var}(x+y) = \sigma_x^2 + \sigma_y^2 + 2\,\text{cov}(x,y)}$$

$$\text{cov}(x,y) = E[(x-\mu_x)(y-\mu_y)]$$
$$= E(XY) - \mu_x \mu_y$$

where, μ_x, μ_y are means of x, y respectively.

If x, y are independent then

$$\boxed{\text{cov}(x,y) = 0}$$

Co-relation: The magnitude of $\text{cov}(x,y)$ doesn't have much meaning without the knowledge of $\text{var}(x)$ and $\text{var}(y)$.

co-efficient of co-relation: It is a mathematical method for measuring intensity of linear relationship between two variables x, y. It is suggested by Karl Pearson. hence ~~its~~ it is called Karl Pearson coefficient of co-relation.

let, x, y be two random variable. Karl Pearson's co-efficient of co-relation is defined by

$$\boxed{\rho_{xy} = \frac{cov(x,y)}{\sqrt{var(x).var(y)}}} \qquad \begin{array}{l} var(x) \neq 0 \\ var(y) \neq 0 \end{array}$$

$$\boxed{\rho_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y}}$$

Note: If x, y are independent then $cov(x,y) = 0$, hence $\boxed{\rho_{xy} = 0}$

but the converse is not true.

Q. ~~Ded~~ let x, y be two independent variable with same variance $Var(x) \neq 0$. $var(y) \neq 0$

$U = x + y \quad V = x - y$

~~E(U)~~ $Cov(U,V) = E[(U - E(U))(V - E(V))]$

$E(U) = E(X) + E(Y), \; E(V) = E(X) - E(Y)$

$= E[[X + Y - (E(X) + E(Y))].[X - Y - (E(x) - E(y))]]$

$=$

$cov(x,y) = 0$.

Since $cov(x,y) = 0$.
but u, v, are not independent.

Note:—

i) The values of ρ lies b/w -1 and 1.

ii) if $|\rho| = 1$ then there exist a linear relation b/w x & y

iii) Alternative formula to calculate the co-efficient of co-relation

If $(x_1,y_1)\ (x_2,y_2)\ (x_3,y_3)$ — — (x_n,y_n) are n pair of observation in a bivariate distribution. then

$$cov(x,y) = \frac{1}{n}\Sigma(x-\bar{x})(y-\bar{y})$$

$$\sigma_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

$$\sigma_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{n}}$$

$$\boxed{\rho_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2}\sqrt{\Sigma(y-\bar{y})^2}}} \quad (i)$$

Alternative Formula:

$$\boxed{r_{xy} = \frac{\Sigma d_x d_y}{\sqrt{\Sigma d_x^2 \, \Sigma d_y^2}}} \quad , \quad d_x = x - \bar{x} \quad d_y = y - \bar{y} \qquad \text{(ii)}$$

Eq (1) is convinient to apply if means $\bar{x}$ & $\bar{y}$ come out to be integars

If $\bar{x}$ and $\bar{y}$ are fractional then eq (2) convinient to use.

Q.) Calculate Karl Pearson coefficient of corelation b/w expenditure on advertising and sales from data given below. (0.7804)

Advertising Expenses (in ₹)	Sales (in ~~lak~~ ₹)	d_x	d_y	d_x^2	d_y^2	$d_x d_y$
39	47	-26	-19	676	361	494
65	53	0	-13	0	169	0
62	58	-3	-8	9	64	24
90	86	25	20	625	400	500
82	62	17	-4	289	16	-68
75	68	10	2	100	4	20
25	60	-40 33	-6	1600 1089	36	240
98	91	-29	23	841	529	
36	51	13	-15	169	225	
78	84		18		324	
650	660	0	0	5398	2128	

$\bar{x} = 65 \quad \bar{y} = 66.$

Q From the following table we have calculated the coefficient of co-relation Given then $\bar{x} = 6$, $\bar{y} = 8$.

X	Y
6	9
2	11
10	?(5)
4	8
8	7

$\frac{35+x}{5} = 8$

40

Q. Numerical value of co-efficient of co-relation lies b/w -1 & +1.

i) If $\rho > 0$, then X, Y increase together or decrease together

ii) If $\rho = 1$, then there is perfect +ve co-relation b/w X & Y.

iii) If $\rho = -1$ then there is perfect -ve co-relation

iv) If $\rho < 0$ then variable X & Y move opposite directions

v) If $\rho = 0$ then variable are unco-related.

Q. x, y are two independent random variable with mean 5 and 10 and variances 4 & 9 respectively find the correlation b/w $3x+4y$ & $3x-4y$

$\Rightarrow$ $E(x) = 5$, $E(y) = 10$, $\sigma_x^2 = 4$, $\sigma_y^2 = 9$

$U = 3x+4y$ $\quad V = 3x-4y$, $\text{cov}(x,y) = 0$

$E(U)$ ~~$E(U,V)$~~ $=$

$E(U) = E(3x+4y) = 3\times5 + 4\times10$
$= 15 + 40 = 55$

$E(V) = E(3x-4y) = 3\times5 - 4\times10$
$= 15 - 40 = -25$

$\text{cov}(U,V) = E[(U - E(U)) \cdot (V - E(V))]$
$= E[[(3x+4y) - 3E(x) - 4E(y)]]$
$= [3x-4y - 3E(x) + 4E(y)]$
$= \_\_ \ ? \ \_\_$

$\text{var}(U) = E(U - E(U))^2$
$= E(U^2 + 2E(U)U + E(U)^2]$
$= E(U^2) - [E(U)]^2$

$\text{var}(U) = \text{var}(V) = 180$, $\text{cov}(U,V) = 0$

Regression

There are two types in regression analysis one is linear regression analysis in which relation is in form of $y = ax + b$ or $x = cy + d$ where, a, b, c, d are constant

other is multivariable regression in which lograthimic function. ~~it~~.

Regression is proposed by french Mathematician "Gauss".

- Linear Regression: Study of relation-ship between the variable (2).

If y is dependent variable and ~~x~~ is independent variable. the linear relationship suggest b/w the variable is called regression eq^n. of y on x.

Similarly, x is dependent

- - -

regression eq^n x on y

To estimate the value of y to correspo-nding value of x. regression line y on x is used.

eqⁿ of regression line of y on x is:

$$\boxed{y-\bar{y} = \rho \frac{\sigma_y}{\sigma_x}(x-\bar{x})}$$

where, ρ = co-efficient of co-relation

eqⁿ of regression line of x on y is:

$$\boxed{x-\bar{x} = \rho \frac{\sigma_x}{\sigma_y}(y-\bar{y})}$$

where, ρ = coefficient of co-relation

Note:-

i) The slope of $\rho \frac{\sigma_y}{\sigma_x}$ is called regression co-efficient denoted by b_{yx} similarly ~~of~~ slope in eqⁿ (ii) is called regression coefficient of x on y denoted by b_{xy}

$$b_{yx} = \rho \frac{\sigma_y}{\sigma_x} \qquad b_{xy} = \rho \frac{\sigma_x}{\sigma_y}$$

$$= \frac{cov(x,y)}{\sigma_x \sigma_y} \cdot \frac{\sigma_y}{\sigma_x}$$

$$= \frac{cov(x,y)}{\sigma_x^2}$$

$$= \frac{1/n \sum(x-\bar{x})(y-\bar{y})}{1/n \sum(x-\bar{x})^2}$$

$$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2} \qquad b_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(y-\bar{y})^2}$$

Scanned by CamScanner

• Properties of Regression coefficient.

i) $b_{xy}\, b_{yx} = \rho^2$

$$\rho = \pm\sqrt{b_{xy}\, b_{yx}}$$

correlation coefficient is geometric mean b/w two regression coefficient

ii) ρ is +ve when both regression co-efficient is +ve.

ρ is -ve when both regression coefficient is -ve.

iii) If one regression coefficient is greater than one then other must be less than one.

Example: The following table gives age (x) in yrs of the cars and annual maintainance cost (y) in (₹100 rupees). Estimate the maintainence cost for a 4 year old car after finding regression coefficient.

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1	15	−4	−4.8	16		19.2
3	18	−2	−1.8	4		3.6
5	21	0	1.2	0		0
7	23	2	3.2	4		6.4
9	22	4	2.2	16		8.8
25	99	0	0	40		38

$\bar{x} = 5$

$\bar{y} = 19.8$

$b_{yx} = \frac{\cancel{8}\times 38}{\cancel{408}} = \cancel{4.75}$ 0.95

$y - 19.8 = \cancel{4.75}\ 0.95\,(x-5)$

$y - 19.8 = \cancel{4.75x}\ 0.95x - \cancel{23.75}\ 4.75$

$\cancel{23.75 - 19.8 = 4.75x - y}$

$\cancel{3.95 = 4.75x - y}$

$\cancel{y = 4.75x - 3.95}$

$\cancel{y = 4.75\times 4 - 3.95}$

$y - 0.95x = 19.8 - 4.75$

$$\boxed{y - 0.95x = 15.05}$$

$y - 0.95\times 4 = 15.05$

$y = 15.05 + 3.8$

$y = 18.85$

Q. A panel of judge A, B graded seven debaters and independently awarded the following marks. and eight debater was awarded 36 marks a judge A while judge B was absent if judge B were present How many would you expect him to award to 8th debater assuming the same degree of relation exist in their judgement.

	A (x)	B (y)	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$\Sigma(x-\bar{x})(y-\bar{y})$
1	40	32				
2	34	39				
3	28	26				
4	30	30				
5	44	38				
6	38	34				
7	31	28				
8	36	?				
	242	227				

$\bar{x} = 34.57$

$\bar{y} = 32.42$

Scanned by CamScanner

. For two variable x and y eqn of regression lines are $9y - x - 288 = 0$ and $x - 4y + 38 = 0$ find

i) Mean of x & y

ii) Coefficient of co-relation (ρ)

iii) ratio of st. dev of y to that of x $\left(\frac{\sigma_y}{\sigma_x}\right)$

iv) The most probable value of y when $x = 145$

v) The most probable value of x when $y = 35$

i)

$$-\cancel{x} + 9y - 288 = 0$$
$$\cancel{x} - 4y + 38 = 0$$

$$5y - 250 = 0$$

$$y = \frac{250}{5}$$

$$\boxed{\bar{y} = 50}$$

$$\bar{x} - 200 + 38 = 0$$

$$\boxed{\bar{x} = 162}$$

ii) $\rho = \sqrt{b_{yx} b_{xy}}$

$b_{yx} = 1/9$

$b_{xy} = 4$

$\rho = \sqrt{4/9} = 2/3$

$9y - x - 288 = 0$

$y = \frac{x}{9} + \frac{288}{9}$

$x - 4y + 38 = 0$

$x = 4y - 38$

iii) $b_{yx} = \rho \frac{\sigma_y}{\sigma_x}$

$\frac{1}{\cancel{3}9} = \frac{2}{\cancel{3}} \frac{\sigma_y}{\sigma_x} \Rightarrow \frac{\sigma_y}{\sigma_x} = 1:6$

iv) $9y - x - 288 = 0$

$9y - 145 - 288 = 0 \Rightarrow 9y = 433$

$y = 433/9$

$\boxed{y = 48.11}$

v) $x - 4y + 38 = 0$

$x - 4 \times 35 + 38 = 0$

$x - 140 + 38 = 0$

$\boxed{x = 102}$

Q) The line of regression of y on x and x on y are $y = x + 5$, $16x - 9y = 5$ find $\sigma^2 x$ if $\sigma^2 y = 16$ also find cov(x, y).

$b_{yx} = 1$, $b_{xy} = \frac{9}{16}$

$\rho = \sqrt{1 \times 9/16}$

$\rho = 3/4$

$b_{yx} = \rho \frac{\sigma_y}{\sigma_x}$

$1 = \frac{3}{\cancel{4}} \cdot \frac{\cancel{4}}{\sigma_x}$

$\boxed{\sigma_x = 3}$

cov(x, y)

• Queuing ~~the~~ theory is mathematics of waiting lines.

Queu

~~Tomth~~.

15/11

Kendell's ~~for~~ Notation for Queuing Theory

Queueing process are specified in symbolic form as:

$$(a|b|c):(d|e)$$

so where the a = arrival distribution
b = service time distribution
c = no. of servers.
d = capacity of server
e = queue disciplines

for eg:- $(M|M|1):(\infty|FIFO)$

A queue with poission arrival has exponential ~~d~~ inter arrival time this is denoted by M which specify memory lit property of the exponential distribution

Notation:

L_s : average length of the system = average no. of customers in the system

L_q : average length of the queue

W_s : Average waiting time of a ~~cost~~ customer in the system

W_q : Average waiting time of a ~~cou~~ customer in the queue.

Model 1: Single server infinite capacity $(M|M|1):(\infty|FIFO)$.

This is a simple queue with poisson arrival & exponential service time single server infinite capacity and FIFO.

The arrival rate (λ) and service rate (μ) are constant to find the steady state probability.

$$P_n = \left(\frac{\lambda}{\mu}\right)^n P_0 \quad , \text{where, } P_0 = 1 - \frac{\lambda}{\mu}$$

$$\boxed{P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right).}$$

Scanned by CamScanner

Note :- $\rho = \frac{\lambda}{\mu}$ is called utilization factor and for the existance of steady state solution. ($\rho < 1$).

Characteristics of queue :- (John Little)

i) Average no. of customer in the system $L_s = \frac{\lambda}{\mu - \lambda} = \frac{\lambda/\mu}{1 - \lambda/\mu} = \frac{\rho}{1-\rho}$

ii) Average no. of customer in the queue $\Rightarrow L_q = L_s - \frac{\lambda}{\mu} = \frac{\lambda^2}{\mu(\mu - \lambda)}$

iii) Average no. of customer in non-empty queue.

$$L_w = \frac{\mu}{\mu - \lambda} = \frac{1}{1-\rho}$$

iv) Probability that no. of customer in system exceeds k.

$$P(N > K) = \left(\frac{\lambda}{\mu}\right)^{k+1}$$

v) Probability that the system is busy

$$= 1 - P_0 = 1 - \left(1 - \frac{\lambda}{\mu}\right) = \frac{\lambda}{\mu} = \rho.$$

vi) Probability that the system is idle or empty $= 1 - \lambda/\mu = 1 - \rho.$

7.) Average waiting time in queue.

$$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$$

8) Average waiting time in the system is $W_s = \frac{1}{\mu-\lambda}$

9) Average waiting time of a customer in ~~queq~~ queue if he has to wait

$$= \frac{1}{\mu-\lambda}.$$

Q1 Customer arrived at one man barber shop to a poisson process with mean inter arrival time 20 min customer spent an average of 15 minutes. in the barber ~~than~~ chair if an hour is used as unit of time than

i) What is the probability need not wait for a hair cut?

ii) What is the expected no. of customer in the barber shop and in queue?

iii) How much time can a customer ~~spent~~ to spent in barber shop?

iv) what is probability that 6 or more customer are waiting for service?

$\lambda = 3$ / hr.

$\mu = 4$ / hr

i) $P_0 = 1 - 3/4 = 0.25$ / hr.

ii) $L_s = \frac{\lambda}{\mu - \lambda} = \frac{3}{4-3} = 3$ / hr

$L_q = L_s - \frac{\lambda}{\mu} = 3 - \frac{3}{4} = 2.25$ / hr

iii) $W_s = \frac{1}{4-3} = 1$ / hr

iv) $P(N > 6) = \left(\frac{\lambda}{\mu}\right)^{6+1} = \left(\frac{3}{4}\right)^7$

Q2) Customer arrival a watch repair shop acc. to a poission process at rate(λ) 1 per / 10 min and service is exponential distribution with mean 8 min(μ). i) Find the average time that a customer spend in shop.

ii) Find the average no. of time a customer spends in a shop (W_s)

iii) what is the probability the shop is idle.

iv) Find average no. of customer in a queue