SYA 211 (PROBABILITY II) TIMETABLE

DAY	VENUE	TIME
THURSDAYS	SLT	8:00am – 10:00am
FRYDAYS	SLT	12:00pm – 1:00pm

COURSE OUTLINE

- Further Permutation and Combination
- Probability Laws
- Conditional Probability
- Independence of Event
- Baye's Theorem
- Probability Distribution of Discrete and Continous Random Variables
- Expectation and Moment of Random Variable
- Chebyshev's Inequalities
- Joint, Marginal and Conditional Distribution and Their Moments
- Limiting Distribution and Their Moment

Further Permutaton And Combination

In many experiments, listing of outcomes may be difficult or time wasting. A more sophisticated way of counting such large number of outcomes falls within the range of a branch of mathematics or compinatoria analysis.

Compinatorial analysis, among other things is the method of counting the number of ways in which an action or sequence of actions can be performed. Interest in listin outcomes is always to know the number of the sample space of an event. There are basically two techniques for counting:

Permutation

Combination

Permutation

It is a special case of basic principle of counting and this occurs when we make successive choices from a single set of n objects where the first choice is made in n ways, and for each of the n ways, there are n-1 ways, for the second,

n-2 ways, for the third choice...
It also means the arrangement of objects in a given order.

N- Permutation

Is the arrangement of n distinct objects in a given order, taking one of them at a time. The number of permutation of n distinct objects is taken as n!
(It means the arrangement of n objects in a row / line). It is given by:

$$^nP_n = \frac{n!}{(n-n)!}$$

R- Permutation

~~The the~~ arrangement of objects in a given order taking r at a time from a set of n $(r \leq n)$ is called R permutation.
The number of R permutation is always denoted by:

$$^nP_r = \frac{n!}{(n-r)!}$$

Permutation

The number of different arrangements of n different objects, is equal to the product $n(n-1)(n-2)(n-3)\cdots \times 4 \times 3 \times 2 \times 1$. This product can be written as $n!$ for short and is read n-factorial.
The product of all consecutive integers starting from 1 to n is denoted by $n!$
$\therefore n! = (n-1)(n-2)\ldots \times 3 \times 2 \times 1$.
If an operation can be performed in p-ways and another operation can be performed in q-ways, then the number of different ways of performing the two operations one after the other is $p \times q$ ways. The arrangement of objects taking into account different orders of arrangement is called PERMUTATION.
The permutation of n objects taking r at a time is denoted ^nP_r where

$$^nP_r = \frac{n!}{(n-r)!}$$

NOTE: $0! = 1! = 1$

The permutation of n objects, taking r

at a time, is denoted by ^nP_r, $P(n,r)$, $P_{n,r}$ or ^nP_r. In this unit, we shall use the last one, ^nP_r.

Examples

1 Find the number of permutation of the letters of the word BACKGROUND.

Solution

There are ten different letters in the word. Hence, the number of permutation of the letters of the word $= 10!$
$= 3628800$.

2 Find the number of ways the 1st, 2nd and 3rd positions can be taken by 10 candidates in a scholastic aptitude test assuming that there must be no tie.

Solution

$$^{10}P_3 = \frac{10!}{(10-3)!} = \frac{10!}{7!}$$

$$= 10 \times 9 \times 8 = 720 \text{ ways}$$

Cyclic Permutation

The number of ways of arranging n-objects round a circular ring which can be turned over is:

$$\frac{1 \times (n-1)!}{2}$$

Examples

1 In how many ways can 9 beads of different colours be threaded in a circular ring?

Solution

$$\frac{1 \times (9-1)!}{2} = \frac{8!}{2} = \frac{40320}{2} = 20160 \text{ ways.}$$

2 In how many ways can 7 members of a Board of Directors of a company be seated round a circular table?

Solution

$1 \times (7-1)! = 1 \times 6! = 720$ ways.

Permutation Involving Indistinguishable Objects

The number of ways of permuting n-objects taking n at a time with n_1 objects alike, n_2 objects alike ... n_j objects alike and $n_1 + n_2 = n_3 + \ldots + n_j = n$ is:

$$\frac{n!}{n_1!\, n_2!\, n_3!}$$

Examples

1 Find the number of ways of arranging the letters of the word!

a ABAKALIKI

Solution

There are 3A, 2K, 2I, 1B, 1L, making a total of 9 letters. ∴ total number of permutation of the letter of the word is

$$\frac{9!}{3!2!2!} = 9 \times 8 \times 7 \times 6 \times 5$$
$$= 15120 \text{ ways}$$

b MATHS

Solution

It contains 5 element without repitition

$= 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$ ways

c ELEMENTS

Solution

It contains 8 elements with 3E

$$\frac{8!}{3!} = 8 \times 7 \times 6 \times 5 \times 4$$
$$= 6,720 \text{ ways.}$$

2 In how many different ways can eight balls of identical size be arranged in a line if:

a they are all of the same colour,

b 5 are red and 3 are blue,

c 3 are red, 2 are blue and 3 are white.

Solution

a $8! = 8\times7\times6\times5\times4\times3\times2\times1 = 40,320$ ways.

b $\frac{8!}{5!3!} = \frac{8\times7\times6}{3\times2\times1} = 8\times7$
$= 56$ ways.

c $\frac{8!}{3!2!3!} = \frac{8\times7\times6\times5\times4}{3!\,2!} = 8\times7\times5\times2$
$= 560$ ways.

3 ~~It~~ How many three-digit numbers can be formed from 8 digits 1, 2, 3, ... 8, if:

a repitition are not allowed

b repitition are allowed

Solution

Since repitition is not allowed,

the first digit can take any of 8 digits.

the second digit can take any of the rest 7 digits.

the third digit can take any of the remaining 6 digits.

$\therefore$ the number of arrangement is:

$= 8\times7\times6$

$= 336$ ways.

b Without repitition, the first digit can take any of the 8 digits, also the second and the third digit.

$\therefore$ the number of ways $= 8 \times 8 \times 8 = 512$ ways

Miscellaneous Examples

1 If ${}^nP_n = 720$, find n

Solution

$${}^nP_n = 720 \rightarrow \frac{n!}{(n-n)!} = 720 \rightarrow n! = 720$$

$$720 = 1 \times 2 \times 3 \times 4 \times 5 \times 6 = n!$$

$$n! = 6! \; ; \; n = 6$$

2 If ${}^nP_4 = 20\,{}^nP_2$, find n

Solution

$${}^nP_4 = \frac{n!}{(n-4)!} \qquad {}^nP_2 = \frac{n!}{(n-2)!}$$

$$\therefore \frac{n!}{(n-4)!} = 20 \times \frac{n!}{(n-2)!}$$

If ${}^{12}P_r = 1320$, find r

Solution

${}^{12}P_r = 1320$

$1320 = 12 \times 11 \times 10$

$\therefore (12-9)! = r!$

$3! = r!$

$r = 3.$

If ${}^{n}P^{2} = 132$, find n

solution

${}^{n}P_2 = 132 = 12 \times 11$

${}^{n}P_2 = 12 \times 11 \times \frac{10!}{10!}$

$r = 2$

$(n - 10) = 2$

$n = 10 + 2 \quad n = 12$

In how many ways can a lady with 8 books arrange 4 of them in her shelf?

Solution

She has to choose 4 books out of 8 books. she choses the first in 8 ways the second in 7 different ways, the third in 6 different ways and the fourth in 5 different ways.

$\therefore$ total number of ways $= 8 \times 7 \times 6 \times 5$

$= 1680$ ways.

6 How many numbers consisting of 2 digits can be formed with the help of the digits 1, 2, 3, 4 such that:

a each number contains 1

b no number contains 1.

Solution

a When a particular thing is to be included in each arrangement, you have:

$r \times {}^{n-1}P_{r-1}$ ways.

$r = 2$, $n = 4$

$2 \times {}^{4-1}P_{2-1} = 2 \times {}^{3}P_{1}$

$= 2 \times \frac{3!}{(3-1)!} \rightarrow 2 \times \frac{3!}{2!} = 3! = 6$ numbers

b When a particular thing is always excluded, you have:

${}^{n-1}P_{r}$

then ${}^{4-1}P_{2} = {}^{3}P_{2} = \frac{3!}{(3-2)!} = \frac{3!}{1!}$

$= 6$ numbers

7 How many numbers greater than 600 can be formed from the digits 3, 4, 5 and 6, if no digit may be used twice?

Solution

The numbers ~~of~~ greater than 600 include all four digits and few three digit numbers.

The number of four digits that can be formed is $4 \times 3 \times 2 \times 1 = 24$.

For the three digit number, the first digit must be 6 and any other two digits can follow

1	3	2

$= 1 \times 3 \times 2 = 6$

Hence, we have in all $24 + 6 = 30$.

8 How many four digit numbers greater than 6000 can be formed using the digits 1, 2, 4, 5, 6, 8 if:

a no repitition allowed

b repitition allowed

Solution

a The first digit must be from either 6 or 8. This means there are only two ways of choosing the first digit. If no repitition, there are 5 digits left. So that we have:

2	5	4	3

= 120 numbers

b If repitition is allowed, after the first digit, all six digits given can be chosen for the second, third and fourth digits. Hence, we have:

2	6	6	6

= 432 numbers

Exercises

1 Evaluate:

a $^{6}P_4$ b) $^{13}P_1$ c) $^{5}P_5$ d) $^{9}P_1$ e) $^{10}P_3$ f) $^{8}P_1$

g $^{12}C_9$ h) $^{15}C_6$

2 Find the value of n if $^{n+1}P_3 = {^{n}P_4}$

3 Find the number of permutations of the letters in the following words:

a LAGOS b) ABUJA c) PHYSICS d) PEPPER

e PASSES f) FAILURES g) STATISTICS

h SUCCESSES i) CALCULUS j) HIPPOPOTAMUS.

4 In how many ways can 8 people be seated in a bench if:

a 8 seats are available

b only 5 seats are available

5 How many four digit numbers can be formed with the 9 digits 1, 2, 3, . . . 9, if:

repititions are allowed?

repititions are not allowed?

a) How many different ways can 5 ~~people~~ persons be seated in a bench?

In how many ways can the offices of Chairman, Vice - chairman, Secretary and Treasurer be filled from a committee of seven?

Solutions

- Uncertainty - refers to the outcome of some process of change
- Random Experiment - An experiment in which the possible outcomes cannot be possibly predicted.
- Trial - This is just one act performed.
- Outcome - One of the possible result that can happen in a trial or an experiment. It is the result when the experiment is performed once.

* An experiment can be deterministic or random.

- Deterministic Experiment - The experiment where the experiment is not subject to chance. It is an experiment of bias.
- Sample Space - The set of all possible outcomes of some given random experiment and each outcome can be called a sample point.
- Event - It is a collection of one or more of the outcomes of an experiment which has certain qualities in common.

* A simple event can not be broken down but a

compound event can be broken down into two or more sub-events.

- **Mutually Exhaustive Events** - if one of them must necessarily occur when a random experiment is performed.
- **Mutually Exclusive Events** - if they can not occur together. $P(A \cap B) = \phi$
- **Mutually Inclusive Events** - if they can occur together.

Two events A and B are said to be:

$S = \{x \mid x \in N \text{ and } x < 8\}$
$A = \{1, 2, 3\}$
$B = \{4, 5, 6\}$
$C = \{n \mid x$ is an even no. and divisible 3
$D = \{$

- Set A and B are not mutually exclusive because $P(A \cap B) = \phi$
- Set C and D are mutually inclusive
- **Prior Events** - are events that occur before another event.
- **Post Prior Events** - are events that occur after one event.

- **Independent Events** - Two events A and B are said to be independent if the event A has no event on the probability of the outcome of B.
- **Experiment** - Is a test or an act carried out to test the validity of an hypothesis.

Concept Of Probability

Three Basic Approach To Defining Probability

- Classical Approach - This assumes that the elementary outcomes of a conceived event/experiment are equally likely. It defines the probability $1/n$ for each elementary event;

$$a_i, i = 1, 2, \ldots, n,$$

where the total number of possible outcome of an experiment is n. Thus, if A is an event such that number of A is r and the total number of possible outcomes of the experiment $n(S)$ is n, then the probability of the occurence of A called the success is $P(A) = \frac{n(A)}{n(S)} = \frac{r}{n}$

The probability of the non occurrence of A called the failure is $P(A') = \frac{n-r}{n} = \cancel{\frac{1-r}{n}}$

$$= 1 - \frac{r}{n} = 1 - P(A)$$

$$\boxed{P(A) + P(A') = 1}$$

* When two die are thrown, what is the probability of getting the sum of 9?

Solution

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$\Pr(\text{sum of } 9) = \frac{4}{36} = \frac{1}{9}$$

- **Relative Frequency Approach** – A disadvantage in the classical approach is in the use of words "equally likely". This phrase is vague "equally probable" or "equiprobable" is synonymous to the phrase "equally likely". The empirical probability of an event is taken as the relative frequency of occurrence of the event when the number of observation is very large. i.e, if the expirement is repeated indefinitely the relative frequency will approach the theoretical probability.

If an experiment is repeated n times out of which an event A occurs m times, then the relative frequency of A is $\frac{m}{n}$, the probability of event A is given as $P(A) = \lim_{m \to \infty} \frac{m}{n}$

Example

a) Suppose a coin is tossed 1000 times and head turns up 520 times. The relative frequency of the coin is:

$$\frac{520}{1000} = 0.520$$

If a further total of 498 heads is obtained in another 1,000 tosses, the total number of heads at the end of the ~~other~~ second 1,000 tosses will be 1,018 out of 2,000 tosses

$$(520 + 498) = 1{,}018 \text{ heads.}$$

The relative frequency of heads then is:

$$\frac{1{,}018}{2{,}000} = 0.509$$

If we continue tossing the coin indefinitely we will eventually approach a value known as the probability of obtaining a head in a single toss of a coin

- Axiomatic Approach - The major disadvantage of the R.F approach is the non existence of an actual limiting number in some cases when viewed theoretically; How large will n be?
A thousand?
Ten thousand?
One million?
The modern probability tries to make up for this inadequacy by developing an axiomatic approach in which probability is regarded as an undefined object'

Definition

Let X be the sample space associated with an experiment "$\mathcal{E}$" with event A is associated a real number and is designated by P(A) is called the probability of A,

if $P(A)$ satisfies the following axioms:
~~for every event,~~
$0 \le P(A) \le 1$ for every event A.
$P(S) = 1$
If A and B are mutually exclusive,
$P(A \cup B) = P(A) + P(B)$
for pairwise mutually exclusive events,
$A_1, A_2, \dots A_n$
$\cancel{P(A_1 + A_2 + \dots}$ $P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) =$
$P(A_1) + P(A_2) + \dots + P(A_n) + \dots = \sum_{i=1}^{n} = P(A_i)$
we note ~~denotes~~ that; ~~for~~
for every finite n, $P(A_1 \cup A_2 \cup A \dots \cup A_n)$
$= \sum_{i=1}^{n} P(A_i)$

Consequences Of The Axion

1 If A and A' are complementary events in S, then;

$$P(A') = 1 - P(A).$$

Proof!

$A \cup A' = S$

$P(A \cup A') = P(S) = 1$

$P(A) + P(A') = 1$, once A' are mutually exclusive events. Hence;

$$P(A') = 1 - P(A).$$

2 Given that $\phi \subset S$, then $P(\phi) = 0$, for any same space S;

$$S \cup \phi = S.$$

Proof;

$P(S \cup \phi) = P(S) = 1$

$P(S) + P(\phi) = 1$

Since S and ϕ are mutually exclusive. Hence,

$$1 + P(\phi) = 1$$

$$\text{Since } P(\phi) = 0$$

Addition Rule

If A and B are any two events in a sample space S, then;

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

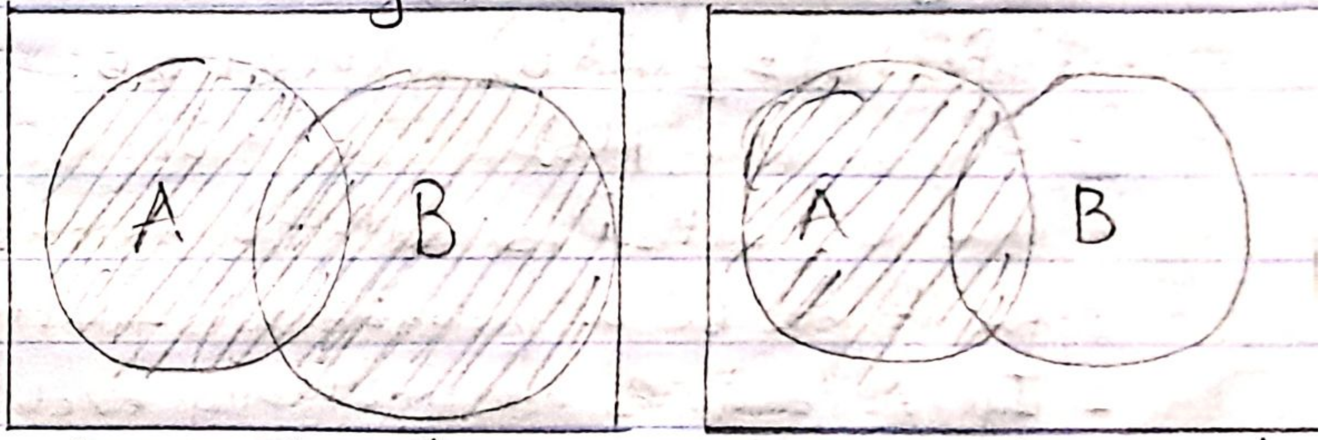
Proof:

The approach will be to express the event $A \cup B$ and A as unions of mutually exclusive events. From set property, we can show that;

$$A \cup B = (A \cap B') \cup B$$

$$A = (A \cap B) \cup (A \cap B')$$

These identities will be illustrated using a venn diagram.



$A \cup B = (A \cap B') \cup B$ $\qquad$ $A = (A \cap B) \cup (A \cap B')$

It also follows that event $(A \cap B)$ and B are mutually exclusive events since;

$$(A \cap B') \cap B = \phi,$$

so that the third Axioms implies;

$$p(A \cup B) = p(A \cap B') + p(B).$$

Similarly, $A \cap B$ & $A \cap B'$ are mutually exclusive so that;

$$p(A) = p(A \cap B) + p(A \cap B'),$$

the theorem follows from these equations;

$$p(A \cup B) = p(A \cap B') + p(B)$$
$$= p(A) - p(A \cap B) + p(B)$$
$$= p(A) + p(B) - p(A \cap B).$$

Conditional Probability

Two events A and B defined on a finite 1
sample space, the C.P of A with respect to B is defined as;

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0.$$ 2

The CP. satisfies all the axioms required for it to be a probability function.

1 If A and B are mutually exclusive events i.e they cannot occur together i.e

$$P(A \cap B) = \phi.$$

Hence, $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = 0$

Note:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 0$$

If A and B are two independent events, then;

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

$$\therefore P(A|B) = \frac{P(A) \cdot P(B)}{P(B)} = P(A).$$

There are two ways of computing C.P,

$P(A|B) =$

- It can be computed directly by conside-ring the probability of B with respect to the reduced sample space A.

Using the definition given above where;

$P(A \cap B)$ and $P(A)$,

are computed with respect to the orginal sample space.

Generally, $P(A_1 \cap A_2 \cap \dots \cap A_k) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1, A_2) \cdots P(A_k|A_1, A_2 \dots, A_{k-1})$.

Assignment

1 A bag containing 10 red, 7 blue and 8 green balls. Three balls are drawn in succession;

i with replacement

ii without replacement;

a what is the probability that the first, the second and the third balls are red, blue and green respectively.

b all three are red

c they are of different colours

d the third is green given that the first and the second are red.

2 A department of statistics has 120 calculators, some of these calculators are programmeable (P) while others are scientific (S). Also some of the calculators are new (N) while others are used (U) given the table below gives the number of calculator in each category and MB7 picks the calculator at random and discovers that it is programmeable. What is the probability that it is one of the new calculators?

	P	S	Total
M	60	15	75
U	20	25	45
Total	80	40	130

Solution

Baye's Theorem

Let $\mathcal{E}$ be a random experiment $B_1, B_2, \ldots B_K$ are partitions of a sample space S associated with $\mathcal{E}$ such that

i $B_i \cap B_j = \phi$ for all $i = j$

ii $B_1 \cup B_2 \cup \ldots \cup B_K = S$

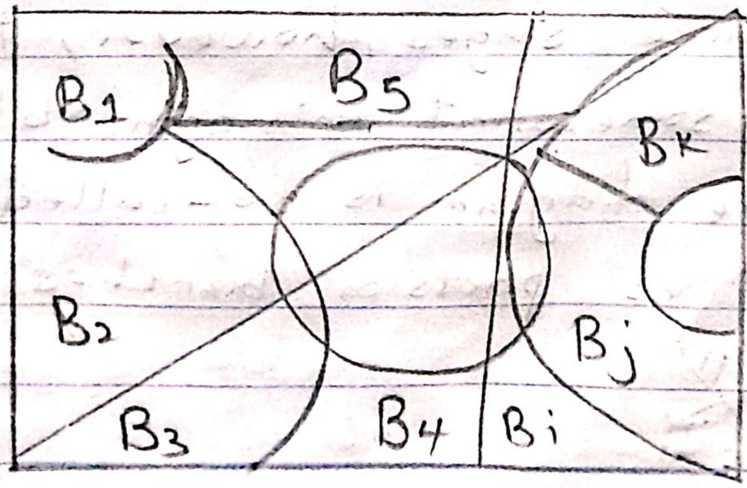
iii $P(B_i) > 0$ for all i

Let A be an arbitrary event that occurred when experiment is performed.

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

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

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



The composition of event A can be written as';

$p(A) = (A \cap B_1 \cup A \cap B_2 \cup \cdots \cup A \cap B_k)$
with some of the $A \cap B_j$ may be empty.
The total probability of;

$$p(A) = p(A \cap B_1) + p(A \cap B_2) + \cdots + p(A \cap B_k)$$

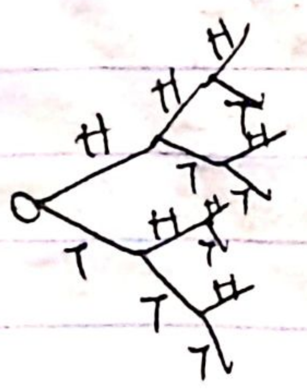
$$p(A) = p(A|B_1) \cdot p(B_1) + p(A|B_2) \cdot p(B_2) + \cdots + p(A|B_k) \cdot p(B_k)$$

$$= \sum_{i=1}^{k} p(A|B_i) \cdot p(B_i)$$

Tree Diagrams

The problem of enumerating or listing a sample point corresponding to various events in the sample space, S, of an experiment, E, can be simplified by the use of Tree Diagram. When the experiment has more than three stages, however, the three may become somewhat unimagineable.

Tree diagram is so-called because of the various parts or branches that appear.



~~Five~~ Example
values

1 5 defective bags got mixed up with 5 non defective ones in a container. 3 bags are randomly selected in succession without replacement;

i what is the probability that the three are defective valves bags

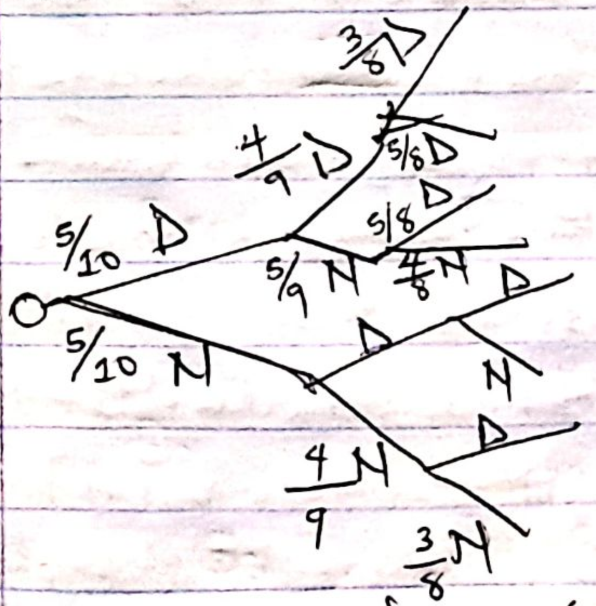
ii only one of the valve bag is defective

iii only two of the valves bags are defective

iv none is defective.

Solution

Let D denote the event the defective valve is choosing and N, non defective valve.



I Pr(defective) = 5/10

Pr (non defective) = 5/10.

i Pr(three are defective valves)

$$= \frac{5}{10} \times \frac{4}{9} \times \frac{3}{8}$$

$$= \frac{1}{12} = 0.0833$$

ii Pr(only one of the valves is defective)

$$= \left(\frac{5}{10} \times \frac{5}{9} \times \frac{4}{8}\right) + \left(\frac{5}{10} \times \frac{5}{9} \times \frac{4}{8}\right) + \left(\frac{5}{10} \times \frac{4}{9} \times \frac{5}{8}\right)$$

$$= \frac{5}{12} = 0.4167$$

iii Pr(only two of the valves are defective)

$$= \left(\frac{5}{10} \times \frac{4}{9} \times \frac{5}{8}\right) + \left(\frac{5}{10} \times \frac{5}{9} \times \frac{4}{8}\right) + \left(\frac{5}{10} \times \frac{5}{9} \times \frac{4}{8}\right)$$

$$= \frac{5}{12} = 0.4167$$

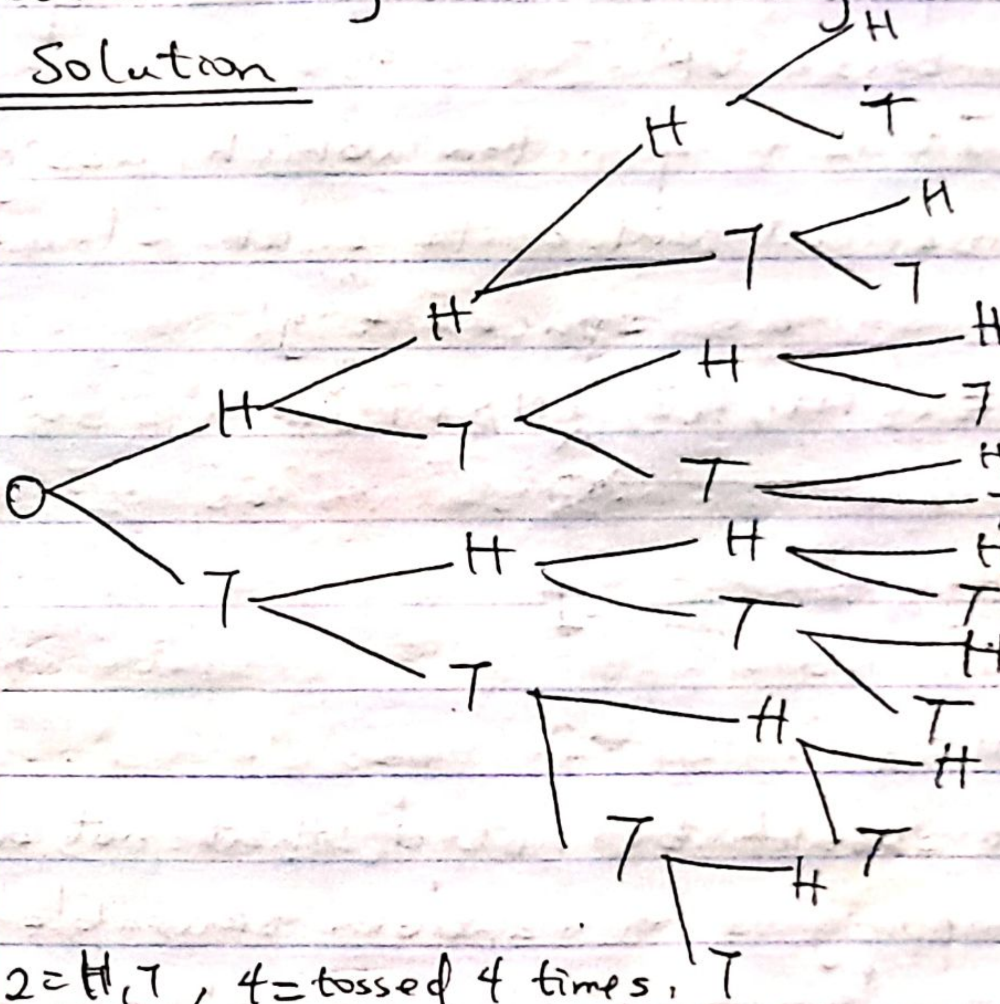
iv pr(none is defective)

$$= \frac{5}{10} \times \frac{4}{9} \times \frac{3}{8}$$

$$= \frac{1}{12} = 0.0833.$$

2 List out the sample space of 4 tossed coins using the tree diagram.

Solution



2 = H, T , 4 = tossed 4 times,

$2^4 = 16$ sample space

HHHH	HTTT	TTTH
HHHT	THHH	TTTT
HHTH	THHT	
HHTT	THTH	
HTHH	THTT	
HTHT	TTHH	
HTTH	TTHT	

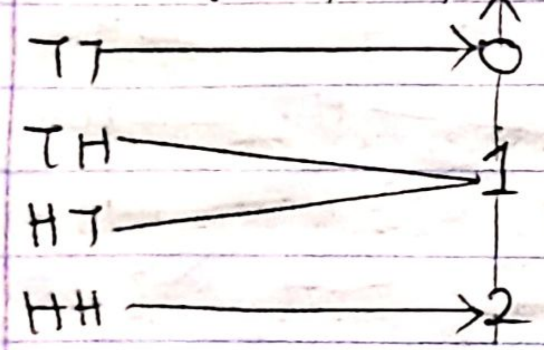
01/08
2024

Random Variables

It is a function that maps a sample space unto a set of real numbers. OR

It is a function which assigns a numerical value to each elements in the sample space. e.g!

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



- It should be noted that not every mapping is a random variable.

A random variable are always represented with the upper case letter, while the value assigned to them are represented with lower case letter.

The Tabular Representation Of Random Variable

Sample Points	HH	HT	TH	TT
X	2	1	1	0

In the table above, the event:

$\{s : X(s) = 1\}$ is simply the set $\{HT, TH\}$.

The notation $\{s : X(s) = 1\}$ is often shortened to $\{X = 1\}$.

In general, $\{X = x\}$ will be used to represent the event A_x. Similarly, $\{s : X(s) < 1\}$ is the same as $\{X < 1\}$,

And finally, the probability of two events: $\{X = 1\}$ and $\{X < 1\}$ are usually written as;

$P(X = 1)$ is simply $p(1)$ and $p(X < 1)$.

In general, for a sequence of X tossing experiments, we have 2^n points in the original sample space, the event space would be defined by $(n + 1)$ points.

There are two types of Random Variable namely;

i Discrete Random Variable

ii Continuous Random Variable.

Probability Distribution

The concept of probability distribution is a very crucial word in probability theory because if we know the probability distribution of a random variable, we'll be able to solve many problems.

i Discrete Random Variable

It can take only finite or countably finite number.

Let X be a Random Variable, if the number of possible values of X (i.e the range space R_x) is finite or countably finite, We call X a discrete Random Variable, i.e the possible values of X may be listed as $x_1, x_2, x_3, \ldots$ In finite case, the list terminates and in the countably infinite case, the list continues indefinitely.

Example

1 Suppose the interest is on the number of particles edited by a radioactive force. Counters observes the emission of these particles during a specific period of time. Here the random variable ~~will be~~ of interest will be X : the number of particles observed. What are the possible values of X.

i

ii

Solution

The possible values of X

$R_x = \{0, 1, 2, 3, 4, \dots\}$

Definition

Let X be a Discrete Random Variable and the Range space of X R_x consists of at most countably infinite number of values $\{x_1, x_2, \cdots x_n\}$ which each possible outcomes x_i will associate a number.

$$p(x_i) = P[X = x_i]$$

called the probability of x_i satisfying the following conditions;

$p(x_i)$ is non-negative i.e $p(x_i) \geq 0$ for all $x_i \in R_x$

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

in the above definition, we see that if the discrete random variable X may assume only a finite number of integer values; $x_1, x_2, \cdots x_k$.

Also, if X assumes a countably infinite number of values, then it is not possible to have all outcome equally likely, it will then become impossible to satisfy the condition;

$$\sum_{i=1}^{\infty} P(x_i) = 1$$

Examples

i An experiment that involves the tossing of a coin, taking count of the number of Heads that shows up.

ii An experiment that involves the tossing of a die, where even number is taken account of;

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

$\mathcal{E} = \{2, 4, 6\}$

Continuous Random Variable

Assuming that the range of X; R_x is made up of very large finite number of values say all values of x; $0 \leq x \leq 1$ of the form $0, 0.01, 0.02, \ldots, 0.98, 0.99, 1.0$ with each of the values associated a non-negative number.

$p(x_i) = p[X = x_i]$ $i = 1, 2, 3, \ldots$ with $\sum P(x_i) = 1$.

It is mathematically easy to idealise

the above probabilistic distribution of x, by supposing that x can assume all possible values $0 \leq x \leq 1$. Since the possible values of x are uncountable, we cannot really speak of the highest of x and hence $p(x_i)$ becomes meaningless. What we then do is to replace the function p defined only for x_1, x_2, by a function f defined for all x; $0 \leq x \leq 1$.

Then we have $f(x) \geq 0$ for all x and the $\int f(x)\,dx = 1$ ~~for~~ :

* for discrete random variable, we sum up.
* for continuous random variable, we integrate

A random variable X is said to be continuous if it takes on an uncountable number of values, that are not all integer values.

OR

if the set of its possible values are contained in an interval such as ;

[a, b]

(a, b).

Example

i An experiment that involves the determination of the life length of a bulb X can assume any value between the interval $[0, \infty]$ or $[a, b]$.

ii Consider the height of trees in a particular location. If H denotes the height of any of the trees, H will assume value between an interval $H = \{h \mid h > 0\}$ i.e $H = \{h \mid 0 < h < \infty\}$.

Probability Mass Function

Let X be a discrete random variable which can assume possible values; $x_1, x_2, x_3, \ldots$ with association probability; $p(x_1), p(x_2), p(x_3), \ldots$ Then the set of other pairs called the probability function (or point probability function or discrete probability density function or probability mass function) of X if;

i $p(x_i) \geq 0$ for all x_i

ii $\sum_{i=1}^{\infty} p(x_i) = \sum_{x_i \in R_x} p(x_i) = 1$

Example

i Let X be the number of heads $R_x = \{0, 1, 2, 3\}$ in tossing three coins together or one coin thrice, then: ~~X(ttt)~~

~~x=0~~ $X(ttt) = 0$ $p(x=0)$

$X(tth) = X(tht) = X(htt) = 1$.

$\therefore p(x=1) = p\{tth, tht, htt\} = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$.

$p(x=2)$ and $p(x=3) = p(hhh) = \frac{1}{8}$.

Then it can be summarized as follow:

x	0	1	2	3
$P(X=x)$	$1/8$	$3/8$	$3/8$	$1/8$

the above is the probability mass function of the discrete random variable X. Observe that;

i $p(x_i) > 0$ for all $x_i + R_x$ and

ii $\sum_{x=1}^{3} p(x_i) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$

Hence, $P[X=x]$ as defined is the legitimate probability mass function of X.

CS 108

2024

Probability Density Function

The probability distribution of a continuous random variable. Let X be a continuous random variable, a function f is called the P.D.F provided:

i $f(x) > 0$ for all x

ii for any a, b with $-\infty < a < b < \infty$, the pdf lies between a and b i.e

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

Remark

1 Essentially x is a CRV if X may assume all values in some interval c, d where c & d may be $-\infty$ and $+\infty$ respectively.

2 $P[a < x < b]$ represents the area under the graph of $f(x)$ of x between a and b.

3 For any specified values of X, say $X = c$ we have $P[X = c] = 0$, since $P[X = c] = \int f(x)\,dx = 0$.

In the light of the diagram above, we see that $P[a \leq x \leq b] = P[a < x < b]$.

Expectation Function

The expected value of a DRV having a PDF $f(x)$ is given by:

$E(x) = \sum_{i=1}^{\infty} x\,p(x)$, where the sum is over all values of x of which $p(x) \geq 0$.

Correspondingly, if X is CRV and $f(x)$ is a value of probability of its random variable, then;

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

Properties Of Expectations Of A DRV.

If x is a DRV, and $f(x)$ is the value of continuous pdf at (x), the expected value of a random variable $g(x)$;

$$Eg(x) = \int g(x) f(x)\,dx.$$

Correspondingly, if x is a DRV and $p(x)$ is the value of its PMF at x the expected value $E(x)$ is given by

$$E(x) = \Sigma\, g(x)\,p(x).$$

i $E(ax) = aE(x)$

ii $E(x+a) = E(x) + a$ where a is a constant and it is a DRV.

* Variance of a constant $= 0$

Expected value of a constant = the constant.

Theorem

For any given random variable:

1 $E(ax+b) = aE(x)+b$

2 $V(ax+b) = a^2 v(x)$

Proving The Theorem

1 $E(ax+b) = aE(x)+b$

$= E[(ax+b)p(x)] = E axp(x) + E bp(x)$

$= a E xp(x) + b E p(x)$

$= aE(x) + b(x)(1)$

$= aE(x) + b$

2 $V(ax+b) = E[(ax+b) - E(ax+b)]^2$

$= E[(ax+b-(aEx+b)]^2$

$= E[ax - aE(x)]^2$

$= E[a^2(x - E(x)]^2$

$= a^2 v(x)$

* $= a^2 E(x - E(x)]^2$

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

$M = E(x)$

$v(x) = E[x - E(x)]^2$.

Example

1 A continous random variable x have a density function given by

$$f(x) = \begin{cases} cx & 0<x<1 \\ 0 & \end{cases}$$

i Obtain the value of constant c of which $f(x)$ is a PDF.

ii Obtain the mathematical representation of y a function of x

$y=3x+1$ is a function of x.

Solution

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

$$\int_0^1 cx\,dx = 1$$

$$c\int_0^1 x\,dx = 1$$

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

$$c\left[\frac{1^2}{2} - \frac{0^2}{2}\right] = 1$$

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

$\frac{c}{2} = 1$, $c=2$.

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{, otherwise} \end{cases}$$

ii $E(Y) = \int y f(y) dy$

$E(3x+1) = E(3x) + E(1) = 3E(x) + 1$

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

$$= 3\int_{-\infty}^{\infty} x \cdot 2x dx + 1$$

$$= 6\int_{-\infty}^{\infty} x^2 dx + 1$$

$$= 6 \times \frac{x^3}{3}\Big|_0^1 + 1$$

$$= 2x^3\Big|_0^1 + 1$$

$$= 2[1^3 - 0^3] + 1$$

$$= 2 \times 1 + 1$$

$$= 3$$

$\therefore E(3x+1) = 3.$

Remark

Some other propbities of mathematical expectations are

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

Find k and $p(0.5 \leq x \leq 1)$.

i Observe that some properties are

$$E\left(\sum_{i=1}^{K} \lambda_i X_i\right) = \sum_{i=1}^{K} \lambda_i E(X_i)$$

ii Note that $E(X,Y) \neq E(X) \cdot E(Y)$ but if X and Y are independent variables, then $E(X,Y) = E(X) \cdot E(Y)$

iii Note also that if $E(X,Y) = E(X) \cdot E(Y)$, this does not imply that X and Y are independent.

iv In general, $E(X/Y) \neq \frac{E(X)}{E(Y)}$, but if X and Y are independent, then $E\left(\frac{X}{Y}\right) = \frac{E(X)}{EY)}$

Expected Values Of X^2 Of A Random Variable X

Let X be a discrete random variable, with possible values; $x_1, x_2, \ldots, x_n, \ldots$ The mathematical expression of X^2, $E(X^2)$ is given as $E(X^2) = \sum_{i=1}^{n} x^2 p(x)$ (Discrete)

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx \text{ (Continuous).}$$

Variance Of A Random Variable

Let X be a random variable (discrete or continuous), the variance of X denoted by $Var(X)$ or σx^2 is defined as;

$$Var(X) = E[X - E(X)]^2.$$

The standard deviation of X is the possible squareroot of variance of X.

Theorem

If X is a random variable, (discrete or contins the Variance of X is defined as;

$$V(X) = E[X - E(X)]^2 = E(X^2) - [E(X)]^2$$

Proof:

Noting that if $p(x^2)$ exists, then $p(x)$ exists

$$\sigma x^2 = E[X - E(X)]^2 = E(X^2 - 2xE(X) + (E(X))^2]$$

$$= E(X^2) - 2E(X).E(X) + (E(X))^2$$

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

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

Properties Of The Variance Of A Random Variable

1 If c is a constant, $V(c) = 0$

2 If c is a constant, $V(cx) = c^2 Var(x)$

3 If $x_1, x_2, \ldots, x_n$ are n independent variables and let $y = x_1 + x_2 + x_3 + \ldots + x_n$,

$$Var(y) = Var(x_1 + x_2 + \ldots + x_n)$$
$$= Var(x_1) + Var(x_2) + \ldots + Var(x_n)$$

Discrete Probability Distributions

1 Bernoulli Distribution

2 Binomial Distribution

3 Geometric Distribution

4 Poisson Distribution

5 Negative Binomial Distribution

6 Hyper Geometric Distribution.

1 Bernoulli Distribution – If an experiment has two possible outcomes success (p) and failure (q) and their probability are respectively θ and $1-\theta$. Then the number of successes or failure (0 or 1) follows the bernoulli distribution. [$n = 1$ (Bernoulli) $n > 1$ (Binomial) n trials]

– A random variable x has a bernoulli

distribution iff its probability distribution is given by:

$$f(x;\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1$$

108 2 __Binomial Distribution__ - Is one of the most
.024 widely used discrete probability distribution. It is used to find the probability of obtaining x successes in n trial of an experiment.

__E.g:__ Toss a coin 100 times and looking for the probability of getting 20 heads. Then the binomial distribution is the best.

A random variable X has a binomial distribution and it is refered to as a binomial random distribution iff the P.D.F is given as $x = 0, 1, \cdots, n$, where n is the total no. of trials, p is the probability of success, q is the probability of failure, s is the probability of successes and $n-s$ is the probability of failures

__Properties Of A Binomial Distribution__

i The experiment must consist of n independent number of trials.

ii Each trials has only two possible outcomes (p&q)

iii The probability of success (p) must remain constant from trial to trial.

iv The repeated trials must be independent

v The random variable is the total number of success.

Structural Properties Of A Binomial Distribution

1 The mean of a binomial distribution is given as;

$M = E(x) = np$

2 Variance

$Var(x) = \sigma^2 = npq$

3 Standard Deviation

$\sigma = S.D(x) = \sqrt{npq}$

4 Moment Coefficient of Skewness

$$\alpha_3 = \frac{q-p}{\sqrt{npq}}$$

5 Moment Coefficient of Kurtosis

$$\alpha_4 = \frac{3+(1-6pq)}{\sqrt{npq}}$$

Example: Supposed that a large number of fusses contains 10% defective. If 4 fusses are randomly sampled from a lot;

i find the probability that exactly 1 fuse is defective

ii what is the probability that at least one fuse in the sample four are defective.

Solution

$n = 4$

$p = 10\% = 0.1$

$q = 0.9$

i $P(x = 1)$

$\binom{n}{x} p^{x} (1-p)^{n-x}$

$\binom{4}{1} 0.1^{1} (1-0.1)^{4-1}$

$= \binom{4}{1} 0.1 (0.9)^{3}$

$= 0.4 \times 0.729$

$= 0.2916$

ii $P(x \geq 1)$

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

$= \binom{4}{0} 0.1^{0} (1-0.1)^{4-0}$

$= 1 \cdot 1 \cdot (0.9)^{4}$

$= 1 \cdot 0.6561$

$= 0.6561$

$P(x \geq 1)$

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

$= 1 - 0.6561$

$= 0.3439$

2024

1 If five fair dice are thrown together once. find the probability of obtaining;
i Number 6. ii) Four 6 iii) at least one 6.

2 The probability that a patient recovers from a rare blood disease is 0.2 if 10 people are known to have contacted the disease and once recovery or not, does not depend on any other. What is the probability that;
i at least 6 survived
ii from 6 - 8 survived

3 A lot containing 80 items is known to contain 6 defective items. From the lot, 5 items are randomly selected, one after the other with replacement. Let D be the number of defective found. Obtain the Probability ~~random~~ distribution of the random distribution D.

3 Geometric Distribution

Suppose the number N of independent trials of bernoulli experiment is not fixed, let the Random Variable of interest X, be the number of trials on which the first success occur. The first N − 1 trial must be the failure. If θ is the probability of success in any trial and probability of failure is $1-\theta$, the probability distribution of X is given by;

$$P_x(n) = P(X=n) = \begin{cases} (1-\theta)^{n-1} & n = 1, 2, 3, \dots \\ 0 & , \text{otherwise} \end{cases}$$

When a RV has a geometric distribution with parameter θ;

$$X \sim G(\theta)$$

* Poisson Distribution (Properties)

1 The average number of success, θ, occurring in a given number of time or specified region, etc, is known.

2 That the probability that a single success will occur during a short time interval or in a short region is proportional to the length of the interval or the size of

the region and does not depend on the number of success occurring outside the region

3 The probability that more than 1 success will occur in such a short time or falling in such small region as negligible.

4 The Poisson process is generally assumed in computing the prob. of RV X, that represents the number of object (or event) distributed over a fixed time interval or region.

If the average occurrence of the event per time interval or region or space is θ, then the prob. that there will exactly X of successes or occurrences is given by;

$$P(X=x) = \begin{cases} \dfrac{e^{-\theta}\theta^{x}}{x!} & x = 1,2,3,\ldots \\ 0, & \text{otherwise} \end{cases}$$

where x is the required number of successe in the given time interval of space or region, and e is the Euler constant. When the RV X, has the poisson distribution with parameter θ (mean), we write;

$$X \sim P(\theta).$$

5 The mean of x $E(x) = \theta$, $Var(x) = \sigma_x^2 = \theta$.

Example

1 The number of telephone calls arriving at MTN switch board is known to be a Poisson process with 120 calls/h on the average. Find the;

i Prob. distribution of X

ii Number of calls that arrives in 1 minute period.

a prob. of no calls in one minute interval

iii prob. of between 1 and 3 calls, both inclusive, arriving in this 1 minute interval.

iv Prob. between 1 and 3 calls, both inclusive in exactly 2 of 3 1 minute interval

Solutions

The average occurrence of the event in a 1 minute interval

$$M_x = \theta = \frac{120}{60} = 2 \text{calls/minute.}$$

i from $p(X = x) = \frac{e^{-\theta}\theta^x}{x!}$

$$P(X=x) = \frac{e^{-2} \cdot 2^x}{x!}$$

ii $p(x=0) = \dfrac{e^{-2}2^{0}}{0!}$

$= 0.1353$

iii $p(1 \leq x \leq 3) = p(x=1) + p(x=2) + p(x=3)$

$p(x=1) = \dfrac{e^{-2}2^{1}}{1!} = 0.2707$

$p(x=2) = \dfrac{e^{-2}2^{2}}{2!} = 0.2707$

$p(x=3) = \dfrac{e^{-2}2^{3}}{3!} = 0.1804$

$P(1 \leq x \leq 3) = 0.2707 + 0.2707 + 0.1804$

$= 0.7218$

iv

Hyper Geometric Distribution

A DRV X, is said to have ~~HGD, if~~ hyper geometric distribution, if;

$$P(X=x) = \begin{cases} \dfrac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, & x=0,1,2, (m,n) \\ 0, & \text{otherwise} \end{cases}$$

where N: total number of items of which m are of the first kind and N−m are of the second type.

n: number of items selected randomly without replacement out of N.

P(X=x): the prob. that there will be exactly x items selected, where x has hyper geometric distribution, we write;

$$X \sim HY(x: n, N)$$

Continuous Distribution

i Exponential distribution
ii Uniform distribution
iii Beta distribution

Moments & Moments Generating Functions

1 The Raw Moment: is also known as moment about the origin and it is defined as the expectations of the powers

1st moment abt d origin $= (r = 1)$
2nd " " " " $= (r = 2)$
3rd " " " " $= (r = 3)$
⋮ ⋮ ⋮ ⋮ ⋮ ⋮

of the RV which has a distribution.

The R-th Raw Moment Of A Random Variable X Denoted By $M'_r = E(X^r)$

The rth raw moment of a RV, X is is denoted by $M'_r = E(X^r)$
If the expected value exists.

If X is a DRV with possible values; $x_1, x_2, \ldots, x_k, \ldots$, with associated probabilities; $P(x_1), P(x_2), \ldots, P(x_k), \ldots$, then;

$$E(x^r) = M'_r = \sum_{i=1}^{\infty} x_i^r \, p(x_i)$$

Provided the series converges.

If X is a CRV with PDF, then;

$$M'_r = E(x^r) = \int_{-\infty}^{\infty} x^r f(x)\,dx. \quad < \infty.$$

The Central Moment (Moment About The Mean)

Let X be a RV, the rth central moment of X about the mean M_x denoted by M_r is defined as;

$$M_r = E[X - M_x]^r$$

If X is a DRV, with possible values; $x_1, x_2, \ldots, x_k, \ldots$, with associated probabilities; $P(x_1), P(x_2), \ldots, P(x_k), \ldots$, the

$M_r = E[X - M_x]^r = \sum_{i=1}^{\infty}(x_i - M_x)p(x_i)$, provided the series converges.

However, if X is a CRV with PDF $f(x)$ $= M_r = E[X - M_x]^r = \int_{-\infty}^{\infty}(x_i - M_x)p(x_i)$

3 <u>Factorial Moment</u>: We observe that;

$\frac{x!}{(x-r)!} = x(x-1)(x-2)\ldots(x-r+1)$, denoting $\frac{x!}{(x-r)!}$ by $x^{(r)}$.

If X is a RV, the rth factorial about the origin of X denoted by;

$M'_{(r)} = E(x^r) = E(x(x-1)\ldots(x-r+1))$

If X is a RV, the rth moment about the mean M_x denoted by M_r is;

$$M_r = E[(x - M_x)^{(r)}]$$

<u>Remarks</u>

1 The first raw moment of a RV is its mean, $E(x)$ i.e $M'_1 = E(x) = M$

<u>Proof</u>

From $E(x^r) = M'_r = \sum_{i=1}^{\infty} x^r p(x)$

Since $r = 1$

$$E(x) = M'_1 = \sum_{i=1}^{\infty} x\,p(x) = M$$

Proof of 3rd & 4th central moment

2 The first central moment is 0.

$$E(x-\mu)^r = \sum_{i=1}^{\infty} (x-\mu)^r p(x)$$

Since $r = 1$

$$E(x-\mu) = \mu_1 = \sum_{i=1}^{\infty} (x-\mu)^1 p(x)$$

$$= \sum_{i=1}^{\infty} x p(x) - \mu \sum_{i=1}^{\infty} p(x)$$

$$= \mu - \mu(1)$$

$$= \mu - \mu = 0.$$

3 The second central moment is the variance of X.

$$E(x-\mu)^2 = \mu_2 = \sum_{i=1}^{\infty} (x-\mu)^2 p(x) = \sum_{i=1}^{\infty} (x^2 - 2\mu x + \mu^2)$$

$$= \sum_{i=1}^{\infty} x^2 p(x) - 2\mu \sum_{i=1}^{\infty} x p(x) + \mu^2 \sum_{i=1}^{\infty} p(x)$$

$$= \mu_2^1 - 2\mu \cdot \mu + \mu^2(1)$$

$$= \mu_2^1 - 2\mu^2 + \mu^2$$

$$= \mu_2^1 - \mu^2$$

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

Assignment

1. Proof the 3rd and 4th central moment
2. A CRV X, has a PDF given as;

$$f(x) = \begin{cases} \frac{1}{6}(2x+1), & 0 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

find the;

i. mean of X
ii. variance of X
iii. third raw moment
iv. third central moment.

Exercise

1. Given that X is a RV with PDF:

$f(x) = \frac{x}{3} = 0,1,2$, find the first two moment about the origin and use the result to find the 2nd and 3rd moment about the mean.

Solution

Moment About The Origin

i $M_1' = \sum_{i=1}^{\infty} x\,p(x) = E(x^1)$

$E(x^1) = \cancel{\int_0^1 \left(x \cdot x/3\right) dx} \int_0^2 \left(x \cdot x/3\right) dx$

$= \cancel{\int_0^1 \left(x^2/3\right) dx} \int_0^2 \left(x^2/3\right) dx$

$= \cancel{1/3 \int_0^1 x^2 dx} \; 1/3 \int_0^2 x^2 dx$

$= \cancel{\frac{1}{3} \frac{x^3}{3} \Big|_0^1} + \frac{1}{3} \frac{x^3}{3} \Big|_0^2$

$= \cancel{\frac{x^3}{9} \Big|_0^1} + \frac{x^3}{9} \Big|_0^2$

$= \cancel{\frac{1^3}{9} - \frac{0^3}{9}} + \frac{2^3}{9} - \frac{0^3}{9} = \cancel{\frac{1}{9} - 0} + \frac{8}{9} - 0$

$= \cancel{\frac{1}{9}} + \frac{8}{9}$

ii $M_2' = \sum_{i=1}^{\infty} x^2 p(x) = E(x^2)$

$E(x^2) = \int_0^2 \left(x^2 \cdot x/3\right) dx$

$E(x^2) = \int_0^2 (x^3/3)\, dx$

$= \frac{1}{3}\int_0^2 \left(\frac{x^3}{1}\right) dx = \frac{1}{3} \cdot \frac{x^4}{4}\Big|_0^2$

$= \frac{x^4}{12}\Big|_0^2 = \frac{2^4}{12} - \frac{0^4}{12} = \frac{16}{12} - 0$

$= \frac{16}{12} = \frac{8}{6} = \frac{4}{3}$

iii $M'_3 = \sum_{i=1}^{\infty} x^3 p(x) = E(x^3)$

$E(x^3) = \int_0^2 (x^3 \cdot x/3)\, dx$

$= \frac{1}{3}\int_0^2 (x^4)\, dx = \frac{1}{3} \cdot \frac{x^5}{5}\Big|_0^2$

$= \frac{x^5}{15}\Big|_0^2 = \frac{2^5}{15} - \frac{0^5}{15} = \frac{32}{15}.$

iv $M'_4 = \sum_{i=1}^{\infty} x^4 p(x) = E(x^4)$

$E(x^4) = \int_0^2 (x^4 \cdot x/3)\, dx$

$= \frac{1}{3}\int_0^2 (x^5)\, dx = \frac{1}{3} \cdot \frac{x^6}{6}\Big|_0^2 = \frac{x^6}{18}\Big|_0^2$

$= \frac{2^6}{18} - \frac{0^6}{18} = \frac{64}{18} = \frac{32}{9}$

2: $M_2 = E[(X-\mu)^2]$

$E[(X-\mu)^2] = \sum_{i=1}^{\infty} (X-\mu)^2 p(x)$

$= \sum_{i=1}^{\infty} (X^2 - 2\mu X + \mu^2)\, p(x)$

$E[(x-\mu)^2] = \sum_{i=1}^{\infty} x^2 p(x) - 2\mu \sum_{i=1}^{\infty} x p(x) + \mu^2 \sum_{i=1}^{\infty} p(x)$

Recall: $\sum_{i=1}^{\infty} x^2 p(x) = 4/3$

$\sum_{i=1}^{\infty} x\, p(x) = 8/9$

$E[(x-\mu)^2] = 4/3 - 2\mu\, 8/9 + \mu^2 \sum_{i=1}^{\infty} p(x)$

$\sum_{i=1}^{\infty} p(x) = \int_0^2 (x/3)\,dx = 1/3 \int_0^2 x\,dx$

$= \frac{1}{3}\frac{x^2}{2}\Big|_0^2 = \frac{x^2}{6}\Big|_0^2 = \frac{2^2}{6} - \frac{0^2}{6} = \frac{4}{6}$

$= 4/6 = 2/3 \cdot \mu = 8/9$

$E[(x-\mu)^2] = 4/3 - 2\mu\, 8/9 + \mu^2\, 2/3 = 4/3 - 2(8/9)8/9 + (8/9)^2 \ldots$

$= 44/81.$

ii $E[(x-\mu)^3] = \mu_3$

$E[(x-\mu)^3] = \sum_{i=1}^{\infty} (x-\mu)^3 p(x)$

$= \sum_{i=1}^{\infty} (x^3 - 3\mu x^2 + 3\mu^2 x - \mu^3] p(x)$

$= \sum_{i=1}^{\infty} x^3 p(x) - 3\mu \sum_{i=1}^{\infty} x^2 p(x) + 3\mu^2 \sum_{i=1}^{\infty} x p(x) -$
$\mu^3 \sum_{i=1}^{\infty} p(x)$

Recall: $\sum_{i=1}^{\infty} x^3 p(x) = 32/15$

$\sum_{i=1}^{\infty} x^2 p(x) = 4/3$

$\sum_{i=1}^{\infty} x\, p(x) = 8/9$

$E[(x-\mu)^3] = 32/15 - 3\mu\, 4/3 + 3\mu^2\, 8/9 - \mu^3$

$\mu = 8/9$

$E[(x-\mu)^3] = 32/15 - 3(8/9)4/3 + 3(8/9)^2\, 8/9 - 8/9 =$

$= \frac{-248}{1215}$ *

Moment Generating Function

The moment generating function of a random variable, x is defined as;

$$M_x(\theta) = E(e^{\theta x})$$

Discrete Random Variable

$$M_x(\theta) = \sum_{i=1}^{n} e^{\theta x_i} f(x_i) = \sum e^{\theta x} f(x)$$

Continous Random Variable

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

Recall that;

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^r}{r!} + \dots$$

$$M_x(\theta) = E(e^{\theta x}) = E\left[1 + \theta x + \frac{(\theta x)^2}{2!} + \frac{(\theta x)^3}{3!} + \dots + \frac{(\theta x)^r}{r!} + \dots\right]$$

$$= 1 + \theta E(x) + \theta^2 E(x^2) + \theta^3 E(x^3) + \dots + \theta^r E(x^r) + \dots$$

$$= 1 + \theta \mu_1' + \frac{\theta^2}{2!}\mu_2' + \frac{\theta^3}{3!}\mu_3' + \dots + \frac{\theta^r}{r!}\mu_r' + \dots$$

where $\mu_1', \mu_2', \mu_3', \dots, \mu_r'$ represent the first, second, third till the rth moment about the origin, respectively and their coefficient

Example

The RV X, can assume value 2, -2 with probability 1/2 each. Find

i the moment generating function of the RV X.

ii the first four moment about the origin

Solution

$X_1 = 2$, $X_2 = -2$

$E(e^{tx}) = \sum_{i=1}^{2} e^{tx_i} f(x) =$

$= E[1 +$

Properties Of Moment Generating Function Using Differential Approach

$$M(\theta) = 1 + \theta \mu_1' + \frac{\theta^2}{2!}\mu_2' + \frac{\theta^3}{3!}\mu_3' + \frac{\theta^4}{4!}\mu_4' + \cdots$$

$$\therefore M'(\theta) = \left(\mu_1' + \frac{2\theta\mu_2'}{2!} + \frac{3\theta^2\mu_3'}{3!} + \frac{4\theta^3\mu_4'}{4!} + \cdots\right)$$

$$M'(\theta) = \mu_1' + \theta\mu_2' + \frac{3\theta^2\mu_3'}{3!} + \frac{4\theta^3\mu_4'}{4!} + \cdots$$

$$M'(0) = \mu_1'$$

$$M''(\theta) = \left(\mu_2' + \theta\mu_3' + \frac{12\theta\mu_4'}{4!} + \cdots\right)$$

$$M''(0) = \mu_2'$$

The connection between $M'(0)$ and $M''(0)$ is the variance;

$$Var(x) = E(x^2) - [E(x)]^2$$
$$= \mu_2' - (\mu_1')^2$$
$$= M''(0) - (M'(0))^2$$

$M^{(r)}(0) = \mu_r'$, where μ_r' is the rth moment about the origin.

Moment Generating Function Of Some Discrete Probability Density Function

Binomial Distribution

$$f(x) = \binom{n}{x} p^x q^{n-x}$$

$$m(\theta) = E(e^{\theta x}) = \sum_{x=0}^{n} e^{\theta x} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^{\theta})^x q^{n-x}$$

$$= (pe^{\theta} + q)^n$$ by binomial expansion

$$V(x) = m''(0) - (m'(0))^2$$

$$m'(\theta) = npe^{\theta}(pe^{\theta} + q)^{n-1}$$

$$m'(0) = np$$

$$m''(\theta) = npe^{\theta}(pe^{\theta} + q)^{n-2}(n-1)pe^{\theta} + npe^{\theta}(pe^{\theta} + q)^{n-1}$$

$$m''(0) = np(p+q)^{n-2}(n-1)p + np(p+q)^{n-1}$$

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

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