**Definition of Statistics:** The science that deals with the collection, classification, analysis, and interpretation of numerical facts or data, where data are collected according to some pre-determined objective. Statistics is especially useful in drawing general conclusions about the population characteristics based on sample observations or population observations.

**Terms related to Statistics:**

**Population:** Population consists of all individuals or items or units which are under investigation in a statistical study. The size of the population is denoted by **N**.

**Sample:**Sample is a representative part of the population units from which information are to be collected.The size of the sample is denoted by **n** (≤ N).

Population and Sample

For **example** we are interested to study the average number of signals sent from a station in different days of a year. There are two ways to measure average: one way to collect the information from the station for every day of the year. This process of collection of data is known as census and using the census data we can calculate the average signal sent per day. Alternatively, instead of recording the information for every day we can record the information for some randomly selected days . This process of collection of data is known as sample survey and using the sample data we can calculate the average signal sent per day.

.**Variable:** The characteristic which varies from one unit to another is called a variable.

**Types of variables:**

**Qualitative Variable:**

The variable which cannot be measured by numerical figure is called a qualitative variable or categorical variable.

**Example:** Type of computer, Color of computer, Type of computer centre, Type of electronic devices, Type of bits etc. are qualitative variables.

**Quantitative Variable:**

The variable which is measured by numerical value is called a quantitative variable.

**Example:** Age, Weight, Height, time, speed, production etc. are quantitative variables.

**Discrete Variable:**

A quantitative variable which takes only integral values is called a discrete variable. It usually ranges from 0 to .

**Example:** Number of computers in laboratories, number of students in each section of a statistics course , number of e.mails received in different days by an organization,etc.

**Continuous Variable:**

The variable which takes integral as well as fractional values is called a continuous variable. It usually ranges from - to.

**Example:** Age of human beings, the life of a battery, speeds of automobile etc.

We can summarize types of variables in the following diagram:

Classification of variables

**Data :**The information collected from population units or from sample units are known as data. The data are of two types, namely ( a ) primary data , and ( b ) secondary data.

**Primary Data :** The data which are collected by investigating population units or sample units are known as primary data. For example, census data are primary data.

**Secondary Data :** The data which are collected from official records or from published works are known as secondary data. For example, census report is a type of secondary data.

**Sources of Statistical Data:** The sources of statistical data are

( a ) Census : The complete investigation of population units is known as census.

( b ) Sample Survey : Investigation of sample units is known as sample survey.

**Sample questions**

1. Mention the importance of statistical data and write down the sources of statistical data.
2. Distinguish between census and sample survey.
3. What are the differences between primary and secondary data?
4. What are the different types of variable? Write down some examples of quantitative and qualitative variable; discrete and continuous variable.
5. What are the different types of data? Write down some examples of different types of data.

**Array:** An arrangement of observations either in ascending order or in descending order is called an array. For example, let us consider the following observations of bits produced by an electronic device in different attempts:

Observation ( x ) :12,19,16,10,20,where number of observations is N=5.

The observations in ascending order are 10,12,16,19,20

The observations in descending order are210,19,16,12,10.

**Arithmetic mean:**

**For ungrouped data set:**

Let are n variates, then, the arithmetic mean is defined by

**For Frequency distribution:**

Let X1, X2,........ ,Xk are k values with frequencies, then, the arithmetic mean is defined by



Where, X= mid value , *f* = frequency , N= total frequency = ∑f

**Geometric mean:**

It is usually calculated if data are given in rates or ratios. The geometric mean of a set of n values of a variable is the nth root of their product.

**For ungrouped data:**

Let a variable x assumes n non-zero and positive values. Then its geometric mean is defined by:

**/ GM = ()1/n**

**For a frequency distribution:**

Let are n variates with frequencies, then, the geometric mean is defined by

**/ GM =)1/n []**

**Harmonic Mean:** It is used if data are given in rates or ratios.

**For ungrouped data:**

Let a variable x assumes n non-zero and positive values. Then its harmonic mean is defined by:  **/**

**For a frequency distribution:**

Let are n non-zero and positive values with frequencies , Then the harmonic mean is defined by:  **[ n = ∑f ]**

**For example:**

Let a data set X: 1,3,9 then

AM = =

GM = = 3

HM = = 2.08

**Here, AM > GM > HM. If data set is as X: 4, 4, 4, and 5; then, AM=GM=HM=4**

**Example :** The above data represent the number of computer centres in different localities. Calculate arithmetic mean of coputers per locality.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Class interval of computer centres | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | Total |
| No. of localities | 3 | 5 | 9 | 3 | 2 | 22 |

**Solution:** The calculation is shown in the table below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Class interval | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | Total |
| No. of localities (*fi*) | 3 | 5 | 9 | 3 | 2 | 22 |
| Mid value (*Xi*) | 15 | 25 | 35 | 45 | 55 |  |
|  | 45 | 125 | 315 | 135 | 110 | 730 |

**Example:** Find the geometric mean of the speed ( minute /customer) to serve in a mobile operator’s office in different days, where the speeds are ( *x*) :15, 12, 13, 19, 10.  
**Solution:** The calculation is shown in the table below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| x | 15 | 12 | 13 | 19 | 10 | Total |
| log x | 1.1761 | 1.0792 | 1.1139 | 1.2722 | 1 | 5.648 |

**Example:** The following data represent the distribution of time ( minutes) needed to develop computer programs for solving some mathematical problems:  
Calculate geometric mean of the distribution.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Class interval of time | 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | Total |
| No. of programs | 4 | 8 | 10 | 6 | 7 | 35 |

**Solution:** The calculation is shown in the table below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Mid-value (x) | log x | f | f log x |
| 0 - 10 | 5 | 0.6990 | 4 | 2.7960 |
| 10 - 20 | 15 | 1.1761 | 8 | 9.4088 |
| 20 - 30 | 25 | 1.3969 | 10 | 13.9790 |
| 30 - 40 | 35 | 1.5441 | 6 | 9.2646 |
| 40 - 50 | 45 | 1.6532 | 7 | 11.5724 |
| Total |  | | 35 | 47.0208 |

**Example:** The following is the distribution of consumption of electricity (MW/locality ) in different days :

Calculate harmonic mean of the diatribution.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 | 70-80 | 80-90 | 90-100 |
| Frequency | 3 | 12 | 8 | 7 | 15 | 26 | 5 | 3 | 1 |

**Solution:** The calculation is shown in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
| Class | Frequency (f) | Mid value (x) | f / x |
| 10-20 | 3 | 15 | 0.200 |
| 20-30 | 12 | 25 | 0.480 |
| 30-40 | 8 | 35 | 0.229 |
| 40-50 | 7 | 45 | 0.156 |
| 50-60 | 15 | 55 | 0.273 |
| 60-70 | 26 | 65 | 0.400 |
| 70-80 | 5 | 75 | 0.067 |
| 80-90 | 3 | 85 | 0.035 |
| 90-100 | 1 | 95 | 0.011 |
| Total | n = 80 |  | 1.849 |

**Exercise:**

1. A fellow travels from city A to city B. For the first 10 miles, he drove at the constant speed of 20 miles per hour. Then he (instantaneously) increased his speed and, for the next 10 miles, kept it at 30 miles per hour. Find the average speed of the movement.
2. A person runs 10 programs at a speed of 400 MB 20 program at a speed of 550 MB. Find the average speed of run per program.

**Sample Questions**

1. Find arithmetic mean, geometric mean, harmonic mean of the observations (x): 10, 18, 5, 7, 12, and 5.

3. A train moves 1st 80 km at a speed of 75 km/h, 2nd 70 km at a speed of 85 km/h, 3rd 85 km at a speed of 66 km/h and 4th 55 km at a speed of 50 km/h. Find the average speed throughout the journey.

**Standard deviation:** The standard deviation is defined as the positive square root of the mean of the square deviations taken from arithmetic mean of the data.

**For ungrouped data:**

**For grouped data:**

**Coefficient of Variation:** The most important of all the relative measure of dispersion is the Coefficient of Variation (CV), defined as: CV = .Thus CV is the value of SD when is assumed equal to 100. It is a pure number and the unit of observations is not mentioned with its value. It is written in percentage form like 20% or 25%. When its value is 20%, it means that the observations vary, on an average, 20% with respect to mean.

**Example:** Calculate the coefficient of standard deviation and coefficient of variation for the observations, X: 2, 4, 8, 6, 10, and 12.

**Solution:**

|  |  |
| --- | --- |
| x |  |
| 2 | (2−7)2=25 |
| 4 | (4−7)2=9 |
| 8 | (8−7)2=1 |
| 6 | (6−7)2=1 |
| 10 | (10−7)2=9 |
| 12 | (12−7)2=25 |
| ∑x = 42 | = 70 |

S2 = S =

Coefficient of Variation : %

**Example:** Calculate coefficient of variation from the following distribution of Missed calls recorded in different mobile phone sets:

|  |  |
| --- | --- |
| Marks | No. of Sets |
| 1−3 | 40 |
| 3−5 | 30 |
| 5−7 | 20 |
| 7−9 | 10 |

**Solution:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Marks | f | x | Fx | ()2 | f ()2 |
| 1−3 | 40 | 2 | 80 | 4 | 160 |
| 3−5 | 30 | 4 | 120 | 0 | 0 |
| 5−7 | 20 | 6 | 120 | 4 | 80 |
| 7−9 | 10 | 8 | 80 | 16 | 160 |
| Total | 100 |  | 400 |  | 400 |

S2 = S =

Coefficient of Variation: %

**Covariance:** In probability theory and statistics, covariance is a measure of how much two random variables change together. It is defined as:

Cov (x, y) =

where, (x1,y1) , (x2,y2) ,…,(xN,yN) are the N pairs of values of two variables X & Y.

**Example:** The table below provides the information on power of signal (xi) and the number stations from which signals are sent (yi) :

|  |  |
| --- | --- |
| (xi) | (yi) |
| 2.1 | 8 |
| 2.5 | 12 |
| 4.0 | 14 |
| 3.6 | 10 |

Calculate the value of covariance of x and y.

**Solution:**

Cov (x, y) =

= = 1.53

**Sample questions**

1. Calculate mean square deviation from mean (variance) and C.V. of the observations (x): 2, 8, 7, 0, 5, - 9, and 12.
2. The distribution of diameter ( in mm ) of a hole drilled in a sheet metal component is shown below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class interval of diameter | 10-12 | 12-14 | 14-16 | 16-18 | 18-20 |
| Number of holes (f) | 8 | 16 | 12 | 10 | 4 |

* Calculate standard deviation of diameter.
* Calculate C. V. of the distribution of diameter.

**Basics concepts of Probability:**

* **Experiment:** It is an act that can be repeated under similar conditions.
* **Outcomes:** The results of an experiment are known as outcomes.

**Example:** If signals are sent from a station, then the signal may be faded ( F ) or not faded ( N ) are the outcomes .

* **Random experiment:** It is an experiment whose outcomes cannot be predicted with certainty in advance, and these outcomes depend on chance.

**Example:** If two bits are produced by an electronic device , the produced bits may be fine ( 1 ) or not fine ( 0 ). The results of the experiment can be shown as follows. S :{ I I, 1 0 , 0 1, 1 1 }. Here there are 4 outcomes and these outcomes

are mutually exclusive and exhaustive. The outcomes form a probability space called sample space.

* **Exhaustive outcomes:** All possible outcomes of a random experiment are exhaustive outcomes.

**Example:** In the in the sample space ( S ) given above the outcomes 1 1, 1 0 , 0 1, and 0 0 are exhaustive outcomes.

* **Mutually exclusive outcomes:** If two or more outcomes cannot occur together, then they are called mutually exclusive outcomes.

**Example:** In the example given above when 1 1 occurs 1 0 can occur together. These two outcomes are mutually exclusive outcomes.

* **Equally likely cases:** If all the exhaustive outcomes of a random experiment have equal chance to occur, then the cases are called equally likely cases.

**Example:** Let an unbiased coin is tossed once, where the equally likely outcomes are denoted by n. Sample space is S = {H, T}. If the probability of showing head is

P (H) =1/2 and also for tail is P (T) =1/2, then we can say that all 2 outcomes in the sample space are equally likely.

* **Not Equally likely cases:** If different outcomes have different chances of occurrence, then they are called not equally likely cases.

**Example:** Let a biased coin is tossed once such that H and T occur in the ratio H:T:: 2:1 , where the sample space is S = {H, T}. Here probability of showing head is P (H) =2/3 and for tail it is P (T) =1/3, then we can say that 2 outcomes in the sample space are not equally likely.

* **Sample point:** Any of the possible outcomes in a sample space is known as sample point.

**Example:** In tossing a coin, Head (H) is a sample point and Tail (T) is another sample point.

* **Sample space:** Aset or collection of all possible outcomes of a random experiment is known as sample space. It is denoted by S.

**Example:** In tossing a coin, the possible outcomes are Head (H) and Tail (T). So, the sample space for this random experiment is S = {H, T}.

* **Event:** Any statement regarding one or more of the outcomes of a sample space recorded from a random experiment is known as event. It is denoted by A/B/C/…

**Example:** I f two computers, one Samsung ( S ) and one ACER ( A ) are allotted to two persons to work at different time periods, then the probable arrangements of the computers will be S : { S S, S A, A S, A A }. At this stage one may be interested to calculate the chance of getting the Samsung and ACER computers by them , then the probable results in favour of the statement are A : { S A , A S }. Here statement A is an event. Again, let us consider the occurrence of getting no Samsung computer, then we can write this statement as B : { A A }. This statement B is also an event.

* **Favorable outcomes:** Number of outcomes in favor of an event is known as favorable outcomes. It is denoted by m (≤ n).

**Example:** In the previous example, for event A, m=2 and for B m = 1.

* **Mutually Exclusive Events:** Iftwo or more events have no common outcome(s), they are called mutually exclusive events.

**Example:** In the above case, event A and B are mutually exclusive events as they have no common outcome i.e. A∩B=.

* **Probability:** If a random experiment shows n exhaustive, mutually exclusive and equally likely outcomes and if m (≤ n) outcomes are in favor of an event A, then the probability of an event A is measured by :

P (A) = = ; where, 0 ≤ P (A) = ≤1

* **Complementary Event:** If there are n equally likely outcomes in a sample space and if an event A is defined with m (≤ n) outcomes, then with the remaining (n-m) outcomes another event can be defined. This latter event is known as complementary event. It is denoted by , where

P () = = 1- = 1- P (A)

=> P (A) = 1- P ()

P (A) + P () = 1.

* **Independent Events:** Two events A and B are said to be independent, if and only if P(AB) = P(A) P(B).
* **Mutually Independent Events :** Three events A,B and C are said to be independent, if and only if,

**P ( AB) = P( A) P(B), P ( AC) = P( A) P( C), P( BC) = P ( B ) P(C)**

**P ( A B C ) = P( A ) P( B) P( C ).**

**Example:** Let, 2 balls are randomly selected one by one with replacement from a box of 3 red and 2 black balls, then the selection of second red/black ball is not affected by the selection of first red/black ball. Here, second selection is independent of first selection. Again, if 2 balls are randomly selected one by one without replacement from a box of 3 red and 2 black balls, then the selection of second red/black ball is affected by the selection of first red/black ball selection. Here, second selection is not independent of first selection.

Problem : Two bits are produced one by one using an electronic device where device is such that it produces fine ( F ) bit 50% times and noisy ( N ) bits 50% times. Find the probability that (i ) both bits are fine, ( ii ) one bit is fine, (iii) at least one bit is fine.

Solution : The sample space is

S : { FF, FN, NF, NN }; n =4

1. Let A : both are fine bits, A { FF }, m=1, P( A ) = m/n = ¼ .
2. Let B : one is fine bit and one is noisy bit, B{ FN, NF }, m= 2, P(B)=m/n= 2/4.
3. Let C :at least one is fine bit, C{ FF,FN,NF,}, m=3, P(C) = m/n=3/8.

**Problem:** Inan office there are 5 computers identified by serial number 1,2,3,4 and 5. Two computers are a)selected for two person to work at different working hours,b)selected for two persons to work at the same time. Find the probability that **(i)** both of them get the same number of computers, **(ii)** sum of the numbers of the computers are 8 or first selected computer bears the number 3 **(iii)** sum of the numbers of the computers are 9 or first selected computer bears the number 2, **(iv)** sum of the numbers of the computers are 6 under the condition that second selected computer bears the number 4.

c)Two computers are selected at random. Find the probability that i)the numbers of both computers are odd.ii)sum of the numbers of computer is even.

**Solution: a)**The sample space of the experiment is

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 11 | 21 | 31 | 41 | 51 |
| 12 | 22 | 32 | 42 | 52 |
| 13 | 23 | 33 | 43 | 53 |
| 14 | 24 | 34 | 44 | 54 |
| 15 | 25 | 35 | 45 | 55 |

We have n=25

1. Let, **A** be the event that both computers are of same number.

Favorable cases to A, m=6; where, A = {11, 22, 33, 44, 55}

P (**A**) = = =

1. Let, **B** be the event that sum of the numbers of the computers is 8

and **C** be the event that first selected computer bears the number 3.

* Favorable cases to B, m=3; where, B = {53,44,35}

Favorable cases to C, m=5; where, C= {31, 32, 33, 34, 35,}

Favorable cases to BC, m=1; where, BC = {35}

P(B or C) = P(BC) = P(B) + P(C) P(BC) = + =

1. Let, **D** be the event that sum of the numbers of the computers 9

and **E** be the event that first selected computer bears the number 2.

* Favorable cases to D, m=2; where, D = {45,54}
* Favorable cases to E, m=5; where, E = {21, 22, 23, 24, 25,}
* Favorable cases to DE, m=0; so, D & E are mutually exclusive events.

P(D or E) = P(DE) = P(D) + P(E) P(DE) = + =

1. Let, F be the event that sum of the numbers of the computers 6 and G be the event that first selected computer bears the number 4.

Favorable cases to F, m=5; where, F = {15, 33, 24, 51, 42}

Favorable cases to G, m=5; where, G = {41, 42, 43, 44, 45}

Favorable cases to FG, m=1; where, FG = {42}

P (F/G) = = = given P(G)0

P (F/G) is called **Conditional Probability** of F under the condition that G occurs first.

b)the sample space is

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 21 | 31 | 41 | 51 |
| 12 |  | 32 | 42 | 52 |
| 13 | 23 |  | 43 | 53 |
| 14 | 24 | 34 |  | 54 |
| 15 | 25 | 35 | 45 |  |

**S=**

N=20

i) A is impossible event, P(A)=0

ii)B={53,35,}, C={31,32,34,35}

P(B or C) = P(BC) = P(B) + P(C) P(BC) = + = =1/4

iii)D={45,54},E={21,23,24,25}

P(D or E) = P(DE) = P(D) + P(E) P(DE) = + = .3

iv)F={15,24,51,42},G={41,42,43,45}

Favorable cases to FG, m=1; where, FG = {42}

P (F/G) = = = given P(G)0

c)The sample space is, S={12,13,14,15,23,24,25,34,25,45}

Two computers can be taken from 5 in 5C2=10ways=n

i)let A: number of both computers are odd, A={13,15,35},m=3

P(A)=m/n=3/10=.3

ii) Let B: sum of number of computers is even,B={13,15,24,35},m=4

P(B)=m/n=4/10=.4

**Problem:** In a mobile operator’s office there are 5 electronic engineers and 6 computers scientists.A committee of 4 is to be formed to perform a special duty. Find the probability that the committee will consist of i)all electronic engineers

ii) 2 electronic engineers and 2 computer scientists.

Solution: From 5+6=11 a committee of 4 can be found in 11C4ways=n=462

i)Let A:all 4 are electronic engineers.A can occur in 5C4ways=m=5

P(A)=m/n=5/11C5=5/462

ii)Let B: 2 are electronic engineers and 2 are computer scientists.

B can occur in 5C2×6C2=150ways=m

P(B)=m/n=150/462

**Problem:** In a packet there are 7 Samsung and 5 Nokia mobile phone sets. Two sets are drawn one after another **a**)at random **(b)** one by one with replacement, **(c)** one by one without replacement. Find the probability that **(i)** both are Samsung sets **(ii)** one set is Samsung and another one is Nokia.

**Solution:**

a)there are 7+5=12 sets.Two sets can be selected in 12C2=66ways=n

i)Let A: both are Samsung sets.A can occur in 7C2=21ways=m

P(A)=m/n=21/66=7/22

ii)Let B: one set is Samsung and one is Nokia.

B can occur in 7C1×5C1=35ways=m

P(B)=m/n=35/66

**b**) The sample space to draw 2 sets one after another is S = {SS, SN, NS, NN},

Where S= Samsung, N=Nokia

i)P(A) = P(NN) = =

ii) P(B) = P(NS) + P(SN) = + =

**c**)

i) P(A) = P(NN) = =

ii) P(B) = P(NS) + P(SN) = + =

**Problem:** Two students A and B have started a game to win a prize. He will win the prize that will be able to get head first if an unbiased coin is tossed once. Find the probability of winning the prize by **(i)** A, **(ii)** B, if A starts the game.

**Solution:** The sample space of the experiment is

S: H, TH, TTH, TTTH, TTTTH, TTTTTH, ……

A B A B A B

Here, the sample points H, TTH, TTTTH, …. are favor of winning by A.

P(A) = P(H) + P(TTH) + P(TTTTH) + …. = + + +… = =

P(B) = 1 P(A) = 1 =

**Problem:** In a box there are 30 bulbs. The bulbs are identified by identity number 1 to 30. One bulb is selected at random. Find the probability that the number of selected bulb has the identity number **(i)** either multiple of 3 or 5, **(ii)** either multiple of 5 or 7, **(iii)** even under the condition that it is multiple of 3.

**Solution:** One bulb is selected randomly from 30 bulbs in n = = 30 ways.

1. Let, **A**: multiple of 3, m=10; where, A={3, 6, 9, 12, 15, 18, 21, 24, 27, 30}

**B**: multiple of 5, m=6; where, B = {5, 10, 15, 20, 25, 30}

For AB, m=2; where, AB = {15, 30}

P(AB)= P(A) + P(B) P(AB) = + = =

1. Let, B: multiple of 5, m=6; where, B = {5, 10, 15, 20, 25, 30}

C: multiple of 7, m=4; where, C = {7, 14, 21, 28}

For BC, m=0; where, BC =

P(BC) = P(B) + P(C) = + = =

P(BC) is called Additional Rule of Probability for two mutually exclusive events.

1. Let, D: even no., m=15; where, D={2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30}

A: multiple of 3, m=10; where, A= {3, 6, 9, 12, 15, 18, 21, 24, 27, 30}

For DA, m=5; where, DA = {6, 12, 18, 24, 30}

P (D/A) = = =

**Problem:** In a bag there are 20 cards bearing numbers 1 to 20. Three cards are taken at random and these are arranged in ascending order. Find the probability that the card in second position bears the number 12.

**Solution:** Three cards are taken randomly from 20 cards in n = = 1140 ways.

Let, **A** be the event that there is one card numbered below 12 and one card numbered above 12. There are 11 cards bearing number 1 to 11(12) and there are 8 cards bearing number 13 to 20(12). Number of favorable cases to A are m = = 88 ways. P(**A**) = = =

**Problem:** In an area there are 4 centers to send e-mails. Four persons have gone there to send mails. Find the probability that **(i)** all 4 persons have entered in the same center, **(ii)** four persons have entered in 4 different centers.

**Solution:** Four persons can be gone 4 centers to send e-mails in n = = 256 ways.

1. Let, **A** be the event that 4 persons have entered in the same center.

This can be done in m = 4 ways. P(A) = = =

1. Let, **B** be the event that 4 persons have entered in 4 different centers.

This can be done in m = 4! = 24 ways. P(A) = = =

**Problem:** From a pack of well shuffled 52 cards 2 cards are taken at random. Find the probability that **(i)** all are aces, **(ii)** all are kings, **(iii)** all are spades, **(iv)** one is spade and one is club, **(v)** all cards are of same color, **(vi)** all cards are of same number.

**Solution:** Two cards are taken randomly from 52 cards in n = = 1326 ways.

1. Let, A: two cards are aces. There are 4 aces. Where, m= = 6 ways.

P(A) = = =

1. Let, B: two cards are kings. There are 4 kings. Where, m= = 6 ways
2. P(B) = = =
3. Let, C: two cards are spades. There are 13 spades. Where, m= = 78 ways.

P(C) = = =

1. Let, D: one is spade and one is club. There are 13 spades and 13 clubs.

Where, m= = 169 ways. P(D) = = =

1. Let, E: two cards are of same color. There are 26 cards of black color and 26 cards of red color. Two black cards are drawn in = 325 ways. Similarly, two red cards are drawn in 325 ways. Where, m= 325+325 = 650 ways. P(E) = = =
2. Let, F: two cards are of same number. There are 4 varieties card. Each variety has 13 cards. Two cards of any one number from 4 varieties can be drawn in = 6 ways. Since, there are 13 number of one variety, two cards of same number can be drawn in m=136 = 78ways.

P(F) = = =

**Problem:** Eighty five per cent e-mails sent from a cyber cafe reach to the destination properly. Once 3 mails are checked randomly, Find the probability that **(i)** all 3 reach properly, **(ii)** two reach properly, **(iii)** at least one reaches properly, **(iv)** at best two reach properly.

**Solution:** Let, R = reach properly the e-mail and N = not reach properly the e-mail.

Sent of 3 e-mails can occur in n = = 8 ways. The sample space is-

S = { RRR, RRN, RNR, RNN, NRR, NRN, NNR, NNN }

Given, P(R) = 0.85 and P(N) = 10.85 = 0.15. So, R and N are not equally.

1. Let, **A**: all 3 reach properly; where, A = {RRR}

P(**A**) = P(RRR) = 0.850.850.85 = 0.614125

1. Let, **B**: two reach properly; where, B = {RRN, RNR, NRR}

P(**B**) = P(RRN)+ P(RNR)+ P(NRR)

= (0.850.850.15)+(0.850.150.85)+(0.150.850.85) = 0.325125

1. Let, C: at least one reaches properly;

where, C = {RRR, RRN, RNR, RNN, NRR, NRN, NNR} = {NNN}

P(C) =1- P () = 1- P(NNN) = 1- (0.150.150.15) = 0.996625

1. Let, D: at best two reach properly;

where, D = {RRN, RNR, RNN, NRR, NRN, NNR, NNN} = {RRR}

P(D) =1- P () = 1- P(RRR) = 1- (0.850.850.85) = 0.385875

**Problem:** In an office there are 50 computers, out of which 20 are ACER and 30 are Samsung. The computers are investigated and found that 15 ACER and 25 Samsung computers are good. One computer is selected at random. Find the probability that the selected computer is **(i)** Either Samsung or good compter. **(ii)** ACER given that it is good,

**Solution:** Let, A: Samsung computer, : ACER computer

G: Good, : Not Good

|  |  |  |
| --- | --- | --- |
| G  A | G | Total |
| A | 25 5  15 5 | 30  20 |
| Total | 40 10 | 50 |

i)P(G) = P(A) + P(G)P(A∩G) = + =

ii)P(A̅/G) = = =

b)ten computers are selected at random. Find the probability that

i)all ten are Samsung computers,ii)6 are Samsung and 4 are ACER computers

solution: From 50 computers10 are selected in 50C10 ways=n

i)Let A: all are Samsung computers. A can occur in 30C10=m

P(A)=m/n=30C10/50C10

ii)Let B:6 are Samsung and 4 are ACER computers

B can occur in 30C6×20C4ways=m,P(B)=m/n=(30C6×20C4)/50C10

**Ex:** In a communication system signals are sent, where some of the signals are faded (F) and some reached to the destination properly (P). If power of the signals are not appropriate , then 50% signals are faded. But if signals are sent with sufficient power, still 10% signals are faded. An inspection team observed 3 consecutive signals. Find the probability that i)no signal is faded. Ii) on signal is faded. Iii) at least one signal is faded. Iv) at best one signal is faded.

**Solution:** The sample space of the inspection is

**S= { PPP,PPF,PFF,PFP,FFP,FPP,FPF,FFF}**

First case P(F)=P(P)=1/2

So, Equally likely cases are, n=8

i)Let A: no signal is faded, A={PPP},m=18

P(A)=m/n=1/8

ii)Let B:one signal is faded, B={PPF,PFP,FPP},m=3

P(B)=m/n=3/8

iii)Let C: At least one signal is faded, C={PPF,PFF,PFP,FFP,FPP,FPF,FFF}, m=7

P(C)=m/n=7/8

[ Alternative: here, C=Ā , P(C)= P(Ā)=1-P(A)=1-(1/8)=7/8]

iv)Let D: at best one signal is faded., D={PPP,PPF,PFP,FPP}, m=4

P(D)=m/n=4/8=1/2

Second case, Given P(P)=0.9, P(F)=0.1

i)P(A)= P(PPP)= 0.9×0.9×0.9=.729

ii)P(B)= P(PPF)+P(PFP)+P(FPP)=.9\*.9\*.1+.9\*.1\*.9+.1\*.9\*.9=.273

iii)As C=Ā, P(C)=P(Ā)=1-P(A)=1-.729=.271

iv)P(D)=P(PPP)+P(PPF)+P(PFP)+P(FPP)

=.9\*.9\*.9+.9\*.9\*.1+.9\*.1\*.9+.1\*.9\*.9=.972

Ex: Three consecutive phone calls are monitored. The calls may be voice call(V) or data call(D).Voice call means someone is speaking and data call means it carries a signal. Find the probability that there will be i) 3 voice calls ii)no voice calls, iii) one voice call ,iv)at least one voice call,v)at best one voice call, vi)second call is a voice call.

Solution: The sample space of the experiment is,

**S={VVV,VVD,VDD,VDV,DDV,DVV,DVD,DDD}**

If it is assumed that 50% calls are voice call,then there are n=8 equally likely outcomes.

i)Let A= there are 3 voice calls, A={VVV},m=1

P(A)=m/n=1/8

ii)Let B: there is no voice call, B={DDD},m=1

P(B)=m/n=1/8

iii)Let C: One voice call, C={VDD,DDV,DVD},m=3

P(C)=m/n=3/8

iv)Let D: there will be at least one voice call,

D={VVV,VVD,VDD,VDV,DDV,DVV,DVD} m=7

P(D)= m/n =7/8 [ Alternative: it is seen that D=B̅=1-P(B)=1-(1/8)=7/8

v)Let E: at best one voice call, E={VDD,DDV,DVD,DDD},m=4

P(E)=m/n=1/2

vi)Let F: Second call is a voice call, F={VVV,VVD,DVV,DVD},m=4

P(F)=m/n=1/2.

**Problem:** In a packet,there are 70% computer science and 30% electrical engineering books.20%books are of local writers. one book is selected at random.Find the probability that the selected one is i)either computer science book or book of local writers,ii)electrical engineering book under the condition that it is not of local writer.

Solution:Let A: Computer science book, A̅: electrical engineering book

B: book of local writer,B̅:not of local writer

|  |  |  |  |
| --- | --- | --- | --- |
| B  A | B | B̅ | Total |
| A | .14 | .56 | .7 |
| A̅ | .06 | .24 | .3 |
| total | .2 | .8 | 1 |

A & B are independent events

P(AB)=P(A)P(B)=.7×.2=.14

i)P(B)=P(A)+P(B)-P(AB)=.7+.2-.14=.76

ii)P(A̅/B̅)==.24/.80=.30

**Bayes’ theorem:** Let, S is the sample space having n equally likely outcomes. With some of outcomes let us define an event E. With some of outcomes of E we can define, separate mutually exclusive events H1, H2, …….. Hk . Then-



………………….

**Hk**

**H2**

**H1**

**S**

………………….

**Hk**E

**H1**E

**H2**E

**E**

We have, E = E + E + …….. + E

P(E) = P(E) + P(E) + ……. + P(E)

Again, P(E/) = ; i = 1, 2, ……., k P(E) = P(E/)

Now, Bayes’ theorem states that, P(/E) = ; i = 1, 2, ……., k

**Problem:** In a box there are 70% mathematics books and 30% electrical engineering books. Among mathematics books 40% are foreign books and among electrical engineering books 50% are foreign books. A foreign book is selected. What is the probability that the selected one is an electrical engineering book?

**Solution:** Let, E: Foreign book : Mathematics book : Electrical engineering book

Given, P() = 0.7, P(E/) = 0.4; P(E) = P(E/) = 0.70.4 = 0.28

P() = 0.3, P(E/) = 0.5; P(E) = P(E/) = 0.30.5 = 0.15

P(E) = P(E) + P(E) = 0.28 + 0.15 = 0.43

So, P(/E) = = =

**Sample questions**

1. Two digits are (a) randomly, (b) one by one with replacement, (c) one by one without replacement selected from the digits 1 , 2, 3, 4, 5 . Find the probability that (i) both the digits will be odd number, ( ii ) sum of the digits will be even.
2. A student becomes successful in 80% cases to write a program. One day he is asked to write 3 programs. Find the probability that (i) all 3 programs are written successfully, (ii) exactly two programs are written successfully, (iii) at best 2 programs are written successfully, (iv) no program is written successfully.
3. Out of 20 electrical installations 12 are installed by Company A and 8 are installed by Company B. Eight installations of A and 6 installations of B served well. One installation is chosen at random to observe its performance. Find the probability that the selected installation is- ( i ) of Company A under the condition that its performance is good, ( ii ) of Company B and its performance is not good, (iii ) performing well, (iv) either an installation of Company A or an installation of good service.
4. In an electrical installation there are 80% graduates who are of the field EEE and 20% are of other fields. Ten per cent of EEE and 20% of other fields are not satisfied with the authority. Once one unsatisfied graduate is identified. Find the probability that he is a graduate of EEE.
5. Signals are sent from Station - 1 and Station - 2. Fifty signals from Station- 1 and 30 signals from Station - 2 are sent. It is known that 20% sent from Station - 1 and 30% sent from Station - 2 do not reach properly. One day on random investigation it is found that one signal is not reached properly. Find the probability that the signal is sent from Station - 2.
6. In a class there are 32 male and 8 female students. Eight students are selected at random. Find the probability that (i ) all 8 are female students, ( ii ) all are male students, ( iii ) five male and 3 female students.
7. Two students A and B have started independently to develop a program to solve a mathematical problem. It is known that A becomes successful in 80% cases and B becomes successful in 60% cases. Find the probability that- ( i ) the program will be developed, ( ii ) A becomes successful under the condition that B fails, ( iii ) both of them fail.
8. In an office there are 15 Philips computers, out of which 5 are defective. Four computers are selected at random. Find the probability that, out of 4 computers, 2 are defective and 2 are good computers.
9. In a class there are 40 students. They are identified by the serial number 1 to 40. One student is selected at random. Find the probability that the identification number of the selected student is either multiple of 3 or multiple of 5.
10. From a pack of 52 cards one card is selected at random. Find the probability that the selected card is an ace under the condition that it is a spade.
11. In a mobile operator’s office 12 electrical engineers and 8 computer engineers are working. Among electrical engineers 5 are experienced and 7 are newly appointed. The corresponding figures among computer engineers are 3 and 5. One of the engineers is selected at random. Find the probability that the selected one is experienced under the condition that he is computer engineer.
12. Two unbiased dice are thrown once. Find the probability that- (i) both dice show same number, (ii) first die shows even number, (iii) both dice show even number, (iv) sum of the upper faces of the dice is 8 or more, (v) sum of the upper faces of the dice is above 10, (vi) sum of the upper faces of the dice is less than 7, (vii) second dice shows number 5 or more.
13. In a packet there are 6 books, three of which are on mathematics and 3 are on statistics. Two books are taken at random. Find the probability that- (i) the drawn books are on mathematics, (ii) the drawn books are on statistics, (iii) one of the drawn books is on mathematics and another one is on statistics.
14. An urn contains 6 red and 4 black balls. Three balls are taken at random from the urn. Find the probability that- (a) all three are red, (b) two balls are red, (c) one ball is red.
15. In a box there are 30 tickets numbered 1, 2, 3, 4, 5, ……… , 30. Five tickets are drawn at random from the box and these are arranged in ascending order. Find the probability that the ticket in third position bears the number 20.
16. In a university 70% students are from city centre and 30% are from outside city. Among the students of city centre 90% wear ties. The corresponding percentage among students outside city is 50%. Find the probability that a student wear a tie comes from out of the city.
17. A surgeon operates 70% male patients and 30% female patients. If in a day he operates 3 patients, what is the probability that- (i) 3 male patients are operated, (ii) at best 2 male patients are operated, (iii) no male patient is operated, (iii) at least one male patient is operated? If the male and female patients are equi-probable, find the probabilities of these events.

**Random variable:** Consider an experiment; where 2 coins are tossed. For this experiment, the sample space is, S= {HH, HT, TH, TT}.

Let, X be the no. of heads. Then, in case of TT, x=0; Here, P(x=0) =P (TT) =

for HT and TH, x=1; Here, P(x=1) =P (HT)+P (TH) = + =

for HH, x=2; Here, P(x=2) =P (HH) =

Here, X takes different values 0, 1 or 2 with associated probabilities. So, X is called random variable. We can write,

|  |  |
| --- | --- |
| X | 0 1 2 |
| p(x) |  |

This p(x) is called probability function for ***discrete random variable*** X. Here,

1. 0 p(x) 1, (ii) = 1

Here, {x, p(x)} is called probability distribution.

For ***continuous random variable*** X, the probability density function is given by f(x); where

1. 0 f(x) 1, (ii) dx = 1

As for example, f(x) = 2x; 0 x 1 . Here, {x, f(x)} is called probability distribution.

If X and Y are ***two discrete random variables*** and p(x,y) is a function of X and Y such that-

1. 0 p(x,y) 1, (ii) = 1

then, p(x,y) is called **joint probability function**.

If X and Y are ***two continuous random variables*** and f(x,y) is a function of X and Y such that-

1. 0 f(x,y) 1, (ii) = 1

then, f(x,y) is called **joint probability density function**.

***Mean and Variance of Random variable:***

If X is a random variable, then its Mean (Expectation) is given by-

= E(X) = , if X is discrete

= dx, if X is continuous.

In general, = E() = , if X is discrete

= dx, if X is continuous.

Variance of X is given by-

V(X) = E() [E(X) =

Where, = E() = , if X is discrete

= dx, if X is continuous.

***Properties of Mean (Expectation) and Variance:***

1. If A is any constant and X is a random variable, then

E(A) = A, V(A) = 0.

1. If A and B are any two constants and X is a random variable, then

E(AXB) = AE(X)B, V(AXB) = V(X).

1. If A and B are any two constants and X and Y are two random variables, then

E(AXBY) = AE(X)BE(Y),

V(AXBY) = V(X) + V(Y) 2ABCov(XY) [if X and Y are not independent]

V(AXBY) = V(X) + V(Y) [if X and Y are independent, then Cov(XY)=0]

1. If X and Y are two independent random variables, then

E(XY) = E(X) E(Y)

Covariance of X and Y:

Covariance of X and Y is given by-

Cov(XY) = E(XY) – E(X)E(Y)

If X and Y are independent random variables, then using the 4th property-

Cov(XY) = 0.

**Example.1:**

Given, the probability function of X as follows:

|  |  |
| --- | --- |
| X | 0 1 2 3 |
| p(x) | 0.2 0.3 0.2 0.3 |

Calculate (i) P(X2), (ii) P(X1), (iii) P(1X3), (iv) E(3X-2), (v) V(3X+4).

**Solution:** (i) P(X2) = P(X=0) + P(X=1) + P(X=2) or, P(X2) = 1- P(X=3)

**=** 0.2+0.3 +0.2 = 1- 0.3

**=** 0.7 = 0.7

1. P(X1) = P(X=2) + P(X=3)

= 0.2+0.3

= 0.5

1. P(1X3) = P(X=2) = 0.2
2. E(3X-2) = 3E(X) – 2 = (31.6) – 2 = 2.8

where, E(X) = = (00.2)+(10.3)+(20.2)+(30.3) = 1.6

1. V(3X+4) = V(X)+0 = 9V(X) = 91.24 = 11.16

where, V(X) = E() [E(X) = 3.8 – = 1.24

Here, E(X) = 1.6

and E() = = (0.2)+(0.3)+ (0.2)+ (0.3) = 3.8

**Example.2:**

Given, the probability density function of X as follows:

f(x) = ; 0x2

Find (i) P(X1), (ii) P(X0.5), (iii) P(1X2), (iv) E(2X+3), (v) V(3X-2).

**Solution:** (i) P(X1) = dx = dx = [ =

1. P(X0.5) = dx = dx = [ = [] = 0.9375
2. P(1X2) = dx = [ = [] =
3. E(2X+3) = 2E(X) + 3 = (2) + 3 = 5

Where, E(X) = dx = dx = dx = [ =

1. V(3X-2) = 9V(X) = (9) = 2

where, V(X) = E() [E(X) = 2 =

here, E(X2) = dx = dx = [ = 2

**Example.3:**

The joint probability density function of X and Y is as follows:

f(x, y) = 2x, 0<x<1, 0<y<1

Calculate- (i) marginal distribution of X and Y, (ii) P (X > ) , (iii) P ( < X < ) , (iv) P (Y > ) , (v) E(3X – 2), (vi) V(2X – 3), (vii) E (X – 2Y), (viii) V (2X – Y).

**Solution:** We have, Marginal distribution of X is,

g(x) = dy = dy = dy = 2x [y = 2x (10) = 2x.

Marginal distribution of Y is,

h(y) = dx = dx = [ = 1

Since, g(x) h(y) = 2x 1 = 2x = f(x, y)

So, X & Y are independent random variables; i.e. Cov (XY) = 0.

1. P (X > ) = dx = dx = [ = 1 =
2. P ( < X < = dx = [ = =
3. E(3X-2) = 3E(X) – 2 =( 2 ) 2

where, E(X) = dx = dx = dx = [ =

1. V(2X-3) = 22V(X) – 0 = 4V(X) = 4 =

where, V(X) = E(X2) – [E(X)]2 = ()2 =

here, E(X2) = dx = dx = dx = [ =

1. E(X–2Y) = E(X) – 2E(Y) = – (2 ) = –

where, E(Y) = dy = dy = [ =

1. V(2X-Y) = 22 V(X) + V(Y) – 2.2.1 Cov(XY) = 4V(X) + V(Y) [as Cov(XY) = 0]

V(Y) = E(Y2) – [E(Y)]2 = - (2 =

E(Y2) = dy = dy = [ =

V(2X-Y) = (4 ) + =

**Example.4:**

The joint probability function of two random variables X and Y is shown below:

|  |  |  |
| --- | --- | --- |
| X/Y | 0 1 2 | g(x) |
| 0  1  2  3 | 0.1 0.2  0.1 0.2 0.1  0.1 0.1  0.1 | 0.3  0.4  0.2  0.1 |
| h(y) | * 1. 0.5 0.2 | 1 |

Calculate- (i) P(X<2) , (ii) P(0<X<3) , (iii) P(Y≥1) , (iv) E(2X-Y).

**Solution:**

1. P(X<2) = 0.3 + 0.4 = 0.7
2. P(0<X<3) = 0.4 + 0.2 = 0.6
3. P(Y≥1) = 0.5 + 0.2 = 0.7
4. E(2X-Y) = 2E(X) – E(Y) = (2×1.1) – 0.9 = 1.3

Where, E(X) = g(x) = (0×0.3) + (1×0.4) + (2×0.2) + (3×0.1) = 1.1

and E(Y) = h(y) = (0×0.3) + (1×0.5) + (2× 0.2) = 0.9

* **Random variables (Sample questions related to this chapter):**

1. The probability function of a random variable of X as follows:

|  |  |
| --- | --- |
| X | 0 1 2 3 |
| p(x) | 0.3 0.2 0.3 0.2 |

Calculate (i) P(X3), (ii) P(X2), (iii) P(1X3), (iv) E(3X-1), (v) V(2X+4).

1. The joint probability function of two random variables X and Y is shown below:

|  |  |  |
| --- | --- | --- |
| X/Y | 0 1 2 3 | g(x) |
| 1  2  3 | 0.2 0.1 0.1 0.0  0.1 0.0 0.1 0.1  0.0 0.2 0.0 0.1 | 0.4  0.3  0.3 |
| h(y) | 0.3 0.3 0.2 0.2 | 1 |

Calculate- (i) P(X<1) , (ii) P(0<X<2) , (iii) P(Y≥2) , (iv) E(X-2Y).

1. Given, the probability density function of X as follows:

f(x) = 2x ; 0x1

Find (i) P(X1), (ii) P(X0.5), (iii) E(X+2), (iv) V(2X-1).

1. The joint probability density function of X and Y is as follows:

f(x, y) = 4xy, 0<x<1, 0<y<1

Calculate- (i) P (0< X < 0.5) , (ii) P (Y ), (iii) E(2X – 3), (iv) V(3X – 4), (v) E (X – 3Y), (vi) V (2X – Y).

**Ex:** The probability function of a continuous random variable x is given by

f(x)=x/8, 0<x<6

=0, otherwise.

Find the probability density function of y=x+2

Solution: We have F(y)=P(Y≤y)=P(x+2≤y)

=P(x≤y-2)

=

=[

=(y-2)2

So, f(y)== = , 2<x<8

=0, otherwise.

**Ex:** The probability function of a discrete function of a discrete random variable is given by

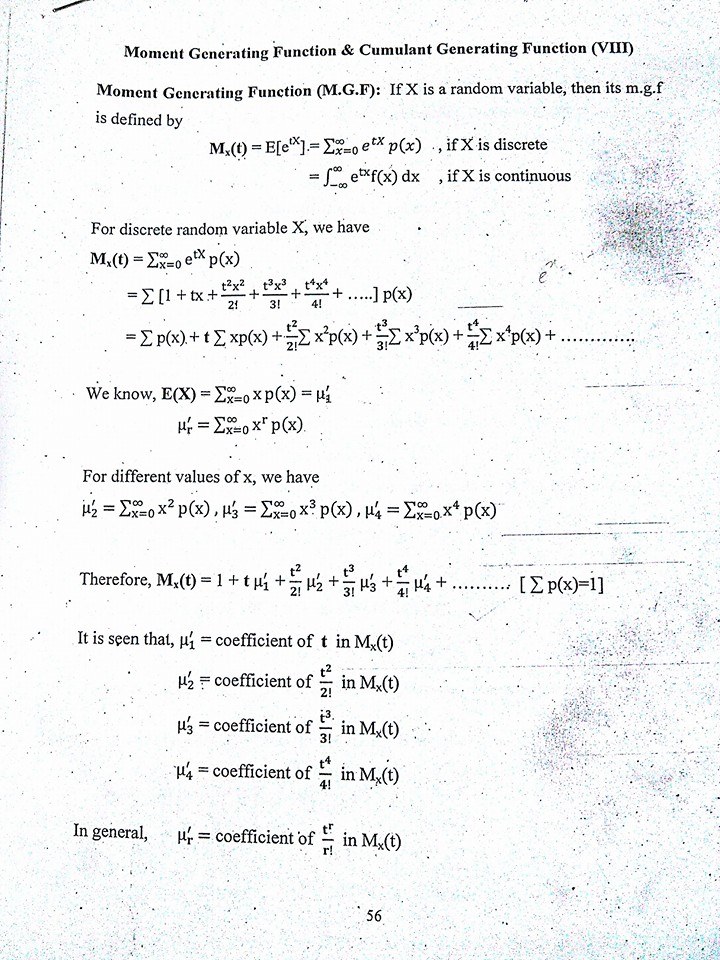
P(x)= =, x= 0,1,2…..

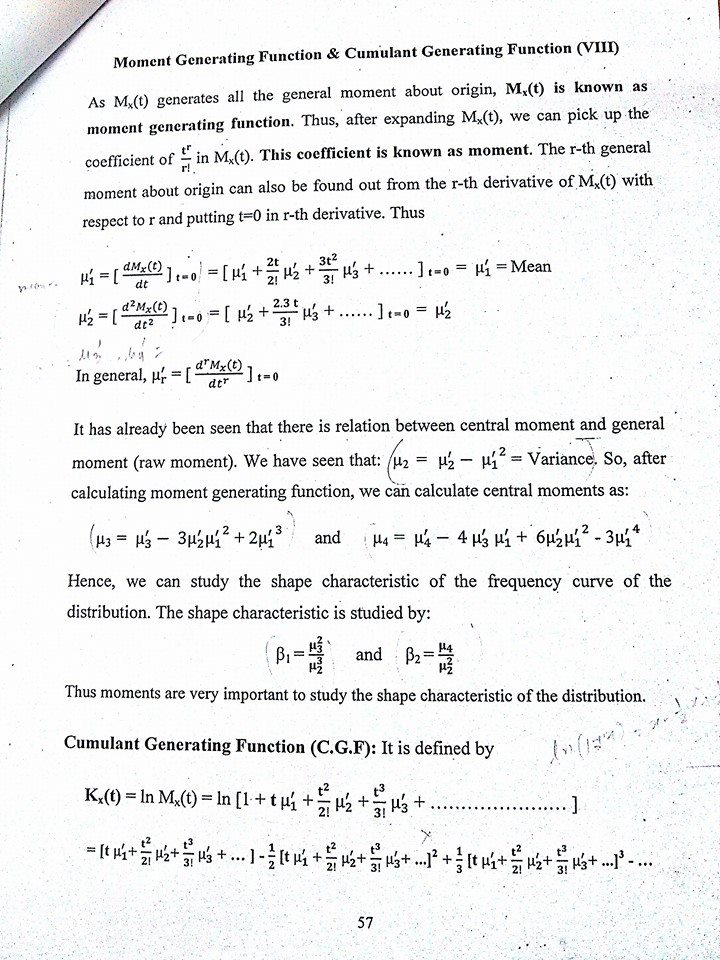
=,0 otherwise,

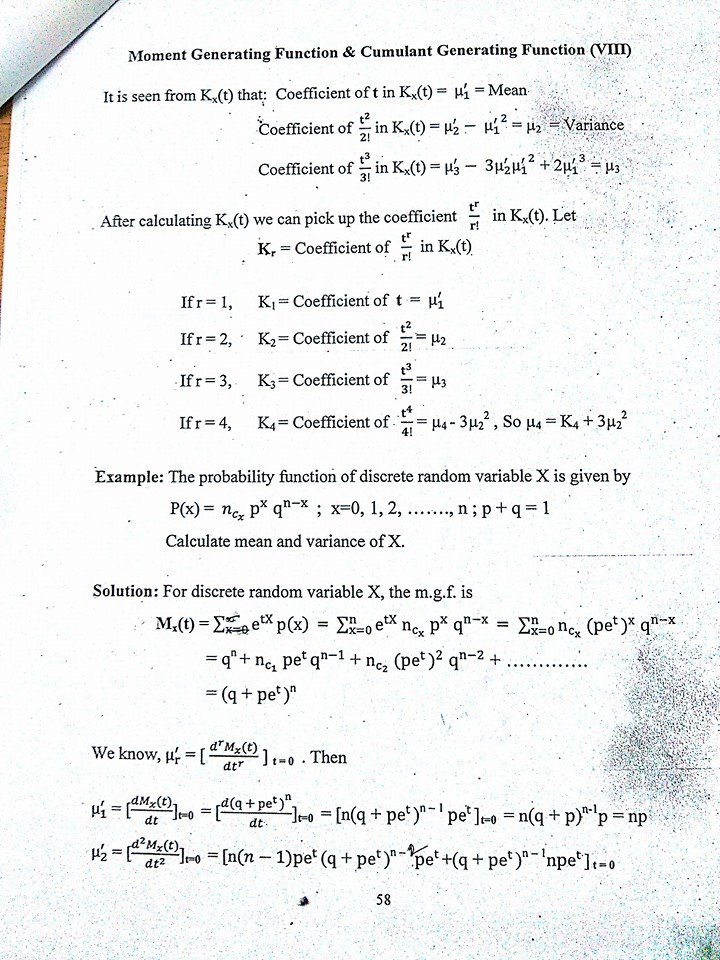
Find the probability function of y=x+4

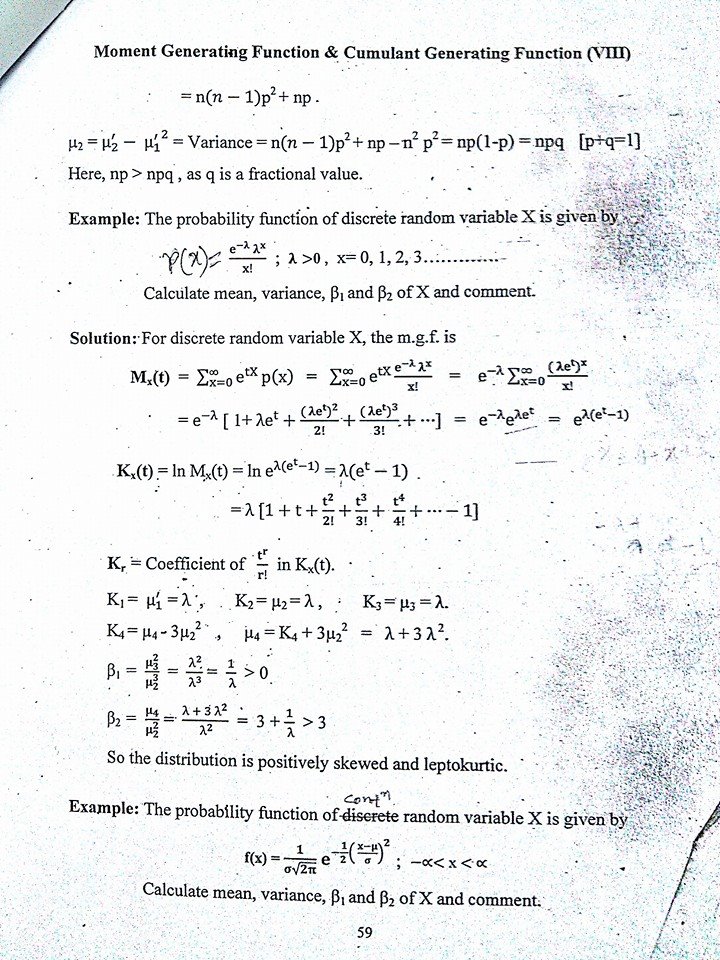
Solution: P(Y=y)= P(x+4=y)=P(x=y-4)

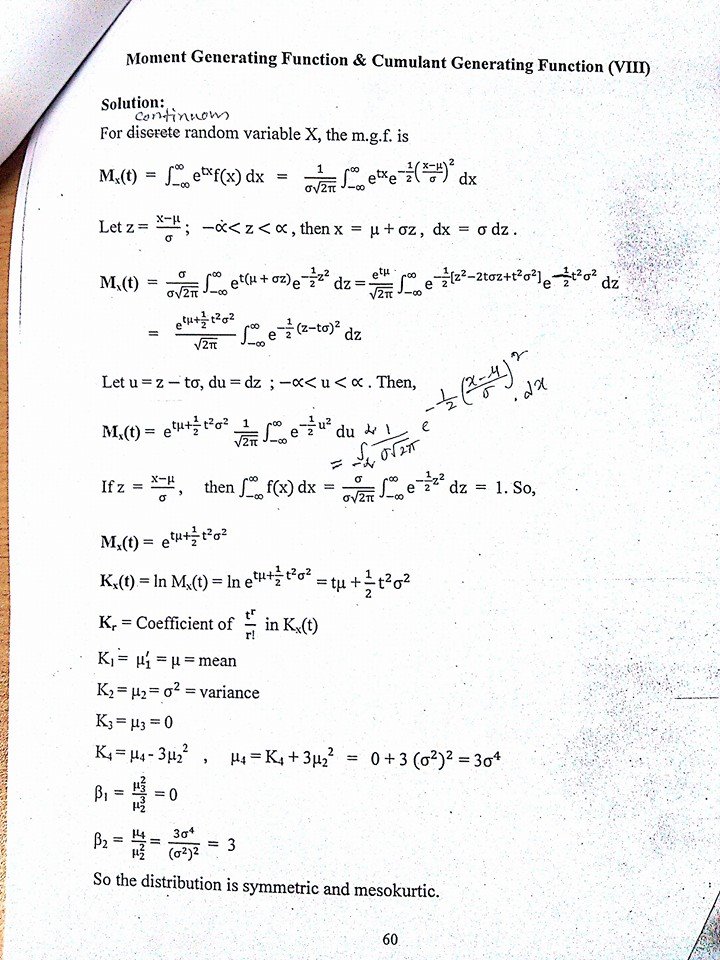
= , y=4,5,6…….

****

****

****

****

****