Total Quality Management-II
Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

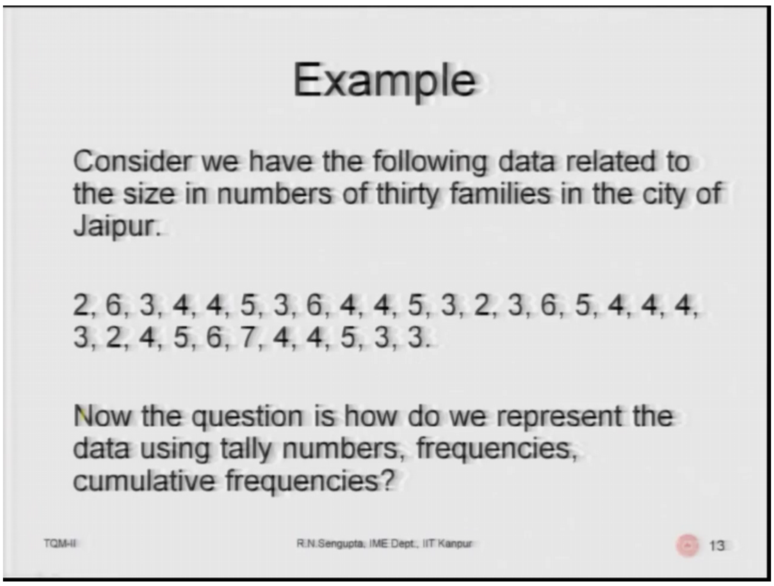
Lecture – 02
Introduction to Probability Theory

A very good morning, good afternoon, good evening my dear friends; welcome to this TQM 2 class or lecture number 2, and as you remember we had just started about giving a very brief introduction about general statistics and we said that we will continue for 2 or 3 classes and then basically go into the depth of TQM and design of experiments.

So, continuing the discussion further we were discussing at length at the last fag end of the first class about property mass distribution probability density functions and then the concept of say for example, cumulative CDF values cumulative distribution function of the less than type greater than type. And how we could basically understand this is CDF for the regress that greater than type less than type considering that you are rolling a die and your main focus is to find out the CDF values for all the faces which are coming up less than equal to 4 or greater than 4.

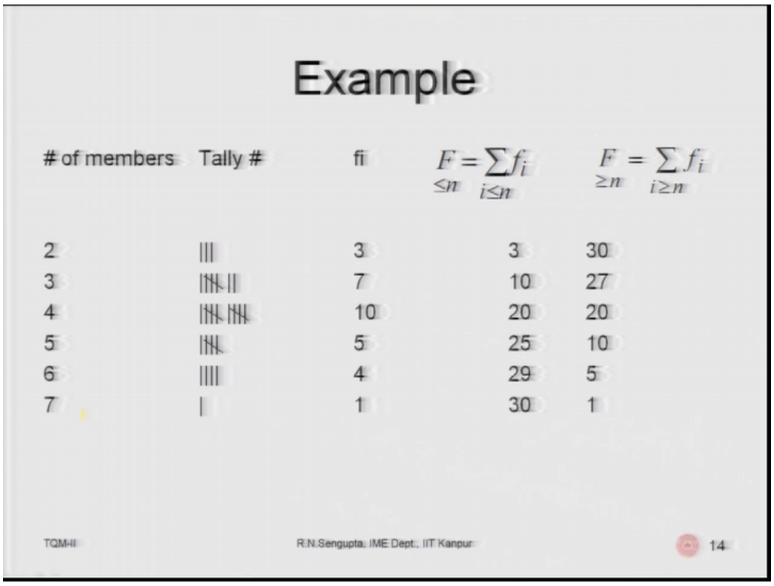
So, you will be basically try to find out the concept and further and concept extend that for future discussion.

(Refer Slide Time: 01:39)



Now, consider this we have the following data related the size in the number of thirty families in the city of Jaipur. So, they are basically 2, 6, 3 and goes to the last 2 numbers and so and henceforth till 3 and 3 now the question is how do we represent the data using the tally numbers frequencies and cumulative frequency. So, the main focus is frequencies.

(Refer Slide Time: 01:56)



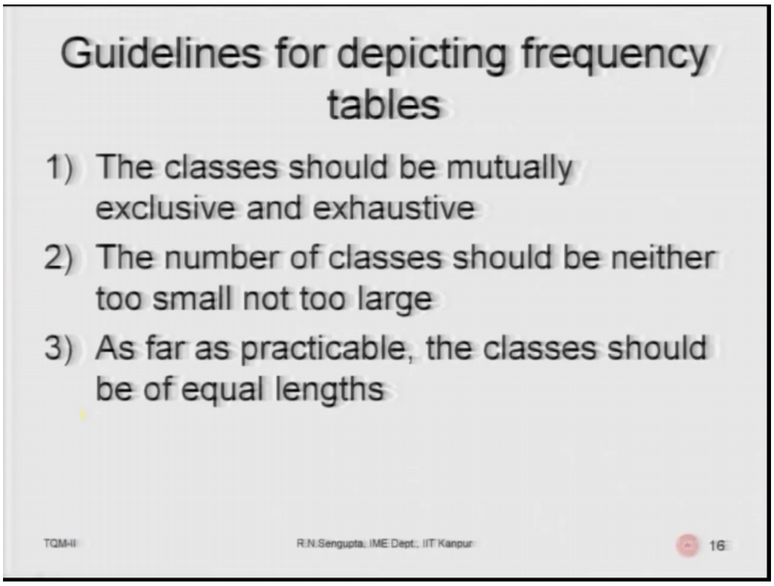
# of members	Tally #	fi	$F_{\le n} = \sum_{i \le n} f_i$	$F_{\ge n} = \sum_{i \ge n} f_i$
2	|||	3	3	30
3	~~||||~~ ||	7	10	27
4	~~||||~~ ~~||||~~	10	20	20
5	~~||||~~	5	25	10
6	||||	4	29	5
7	|	1	30	1

So, these; obviously, we know, but I am not going to go into detail. So, you find the number, and write the number of family members write the tally numbers which is the second column, then f i is basically the frequency which I have been mentioning about time and again frequency relative frequency chance and the probability in the long run.

So, the frequencies are given in the third column starting for number of family members with 2. The number is 3 then family number 3 7 family number of size 4 is10 and the last one being family number with size 7 being 1. And if you find out the less than type of greater than type, then the corresponding addition on the, of the frequencies are given. So, it basically starts which is the second last column starts from 3 to 30, and for the last column basically goes in the reverse direction, because you are trying to basically add up all the probabilities from the for the greater than type and they basically come down from 30 to 1.

So, if you draw the curves you have the ogive curves if you remember we did in class 11, 12 and in statistics. So, the midpoint of the ogive curves or the less than type greater than type gives you the median you can find out the median this just for a information and from there we can find out on the other problem information.

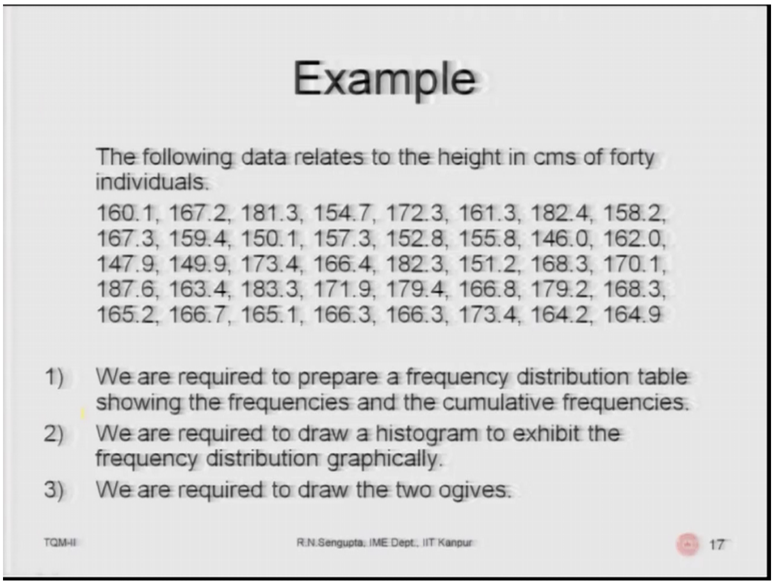
(Refer Slide Time: 03:15)



Now, guidelines for depicting frequency tables would be the classes should be mutually exclusive and exhaustive the number of classes should be neither to too small, too large as far as practicable the classes should be of equal lengths.

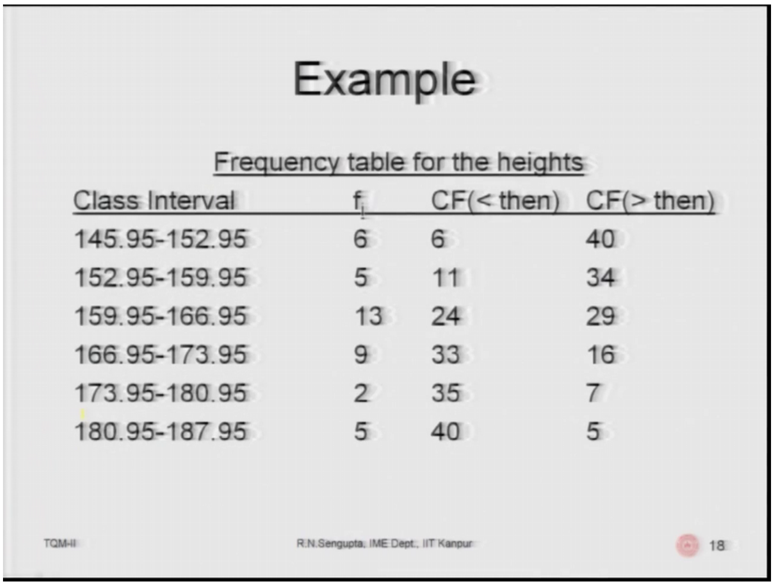
So; that means, when you are taking classes and other than anyone number, consider that you want to find out the number of family members who have between 2 to 4 both inclusive then; obviously, it can be the 2 members in the family 3 members in the family 4 members in the family then the next group would be 5, 6, 7 and so on, and so, so the class intervals of 3 and you do the problems accordingly.

(Refer Slide Time: 03:51)



The following data represents the relates the height in centimetres forty individuals do starting from 160.1, till 160.49, we are required to prepare a frequency distribution table we are required to draw a histogram we are required to draw the two ogive calls; I am just reading the bullet points.

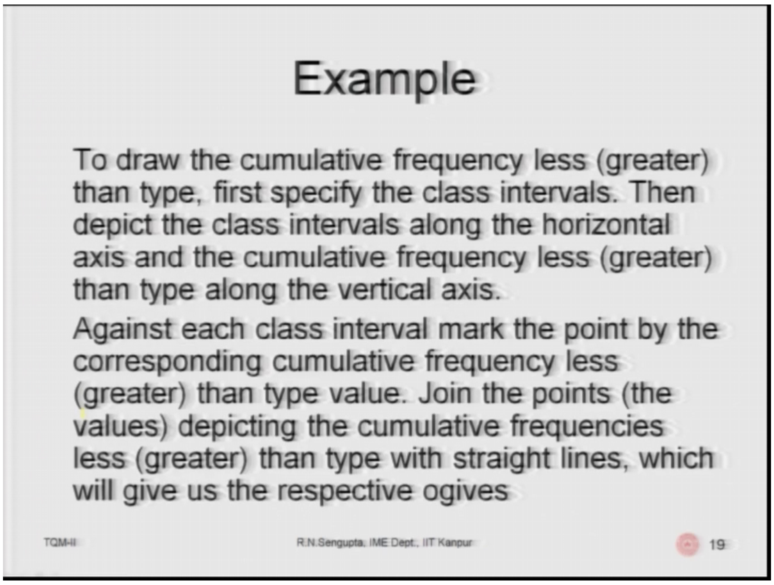
(Refer Slide Time: 04:06)



So, you draw the frequency tables for the heights the leftmost column of basically classes which have been made 1 for the first, interval is 145.95 to 152.95 so; obviously, the class intervals would be made in such a way that none of the values any particular value rises at the end at the boundary. So, they have to be either in the interval 1 or interval 2; that means, I am considering them side by side. So, they can they the values cannot lie between interval and interval 3 at the boundary cannot lie between interval 3 and interval 4 and so on and so forth.

So, the class intervals have to meet accordingly based on that we have made these 6 intervals starting from 145.95, I am again repeating till 152.95 and the last value being 180.95 to 187.95, the frequencies are given from 6 to 5 this addition of the numbers in the in that class then the cumulative frequency less than type starts from 6 to 40, and the cumulative frequency from greater than type basically starts from the top to 40 to 5.

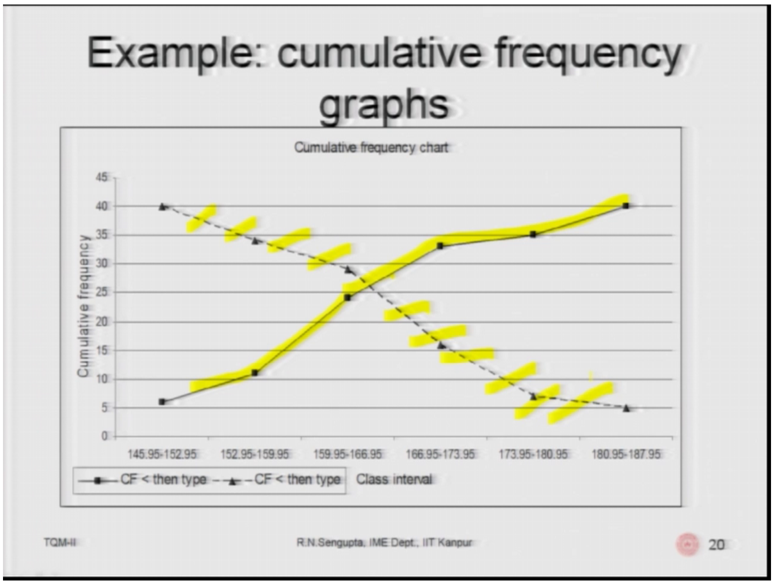
(Refer Slide Time: 05:13)



So, if you basically draw those I of course, you will get the median and do the problems accordingly, to draw the cumulative frequency less than or greater than type first specify the class intervals then depict the class intervals along the horizontal axis and cumulative frequencies less than or greater than whatever in the y-axis.

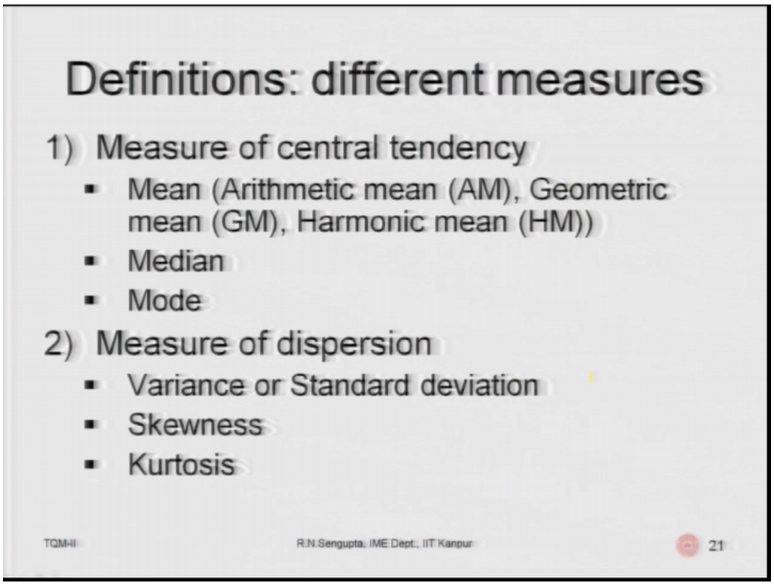
Against each class interval mark the point by the corresponding cumulative frequency or the less than type of greater than time join the points, depicting the cumulative frequencies that gives you the ogive curves find the intersection that gives you the median. So, this is the problem, problem which you have solved.

(Refer Slide Time: 05:46)



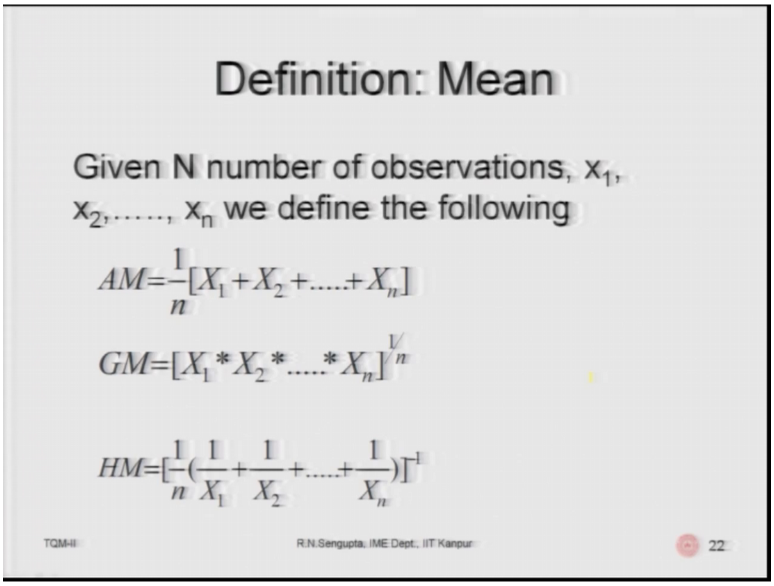
So, the black one where I am is now hovering. So, and this is the yellow coloured highlighter which I am going this basically gives the less than time and where I am just hashing it marked, so that is basically the less than type.

(Refer Slide Time: 06:02)



So, now, we will consider definitions of different measures which are measures of central tendency mean then which is the arithmetic mean, geometric mean, harmonic mean, and then is the median there is the mode and measure the dispersions would be variance standard deviation skewness kurtosis and on all this values so we will come to them as we proceed.

(Refer Slide Time: 06:29)

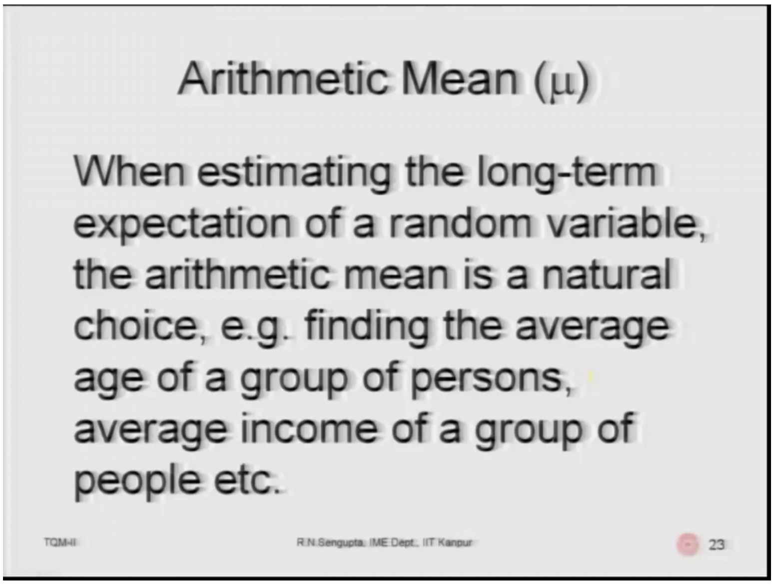


So, for the mean you can have arithmetic geometric and harmonic in arithmetic mean for considering the probability of each occurring is 1 by n considering the n of number of observations and each are distinct, and the distinct values are given by X 1 to X n. So, the corresponding probabilities being multiplied by x i, i being 1 2 n would be 1 by n.

So, you basically technically add up all the values divided by n and get the average for the geometry mean; obviously, you know you will multiply them find out the nth power root and basically solve the problem. So, arithmetic mean may be related to the high geometric mean may be related to interest rate harmonic mean may be related to the concept of speeds when you are travelling from city a to city B and coming back from city B to city a you want to find out the average speed.

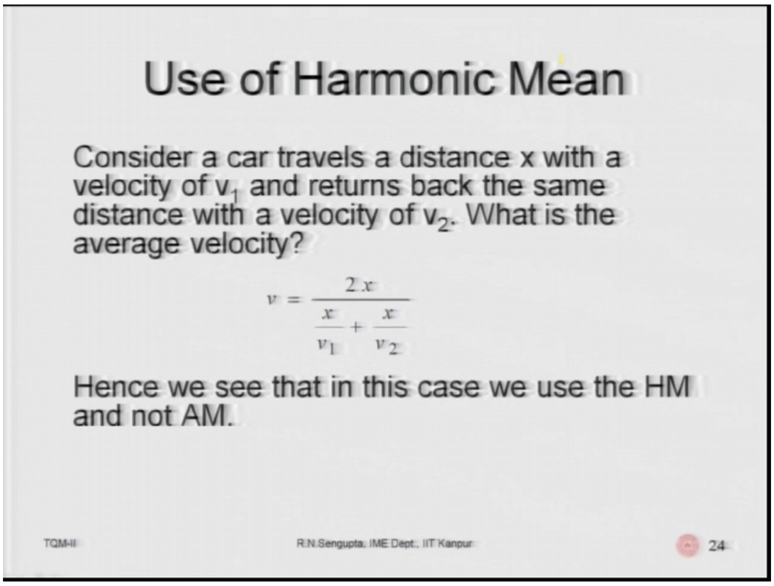
So, harmonic mean is given by this formula. So, these we have solved many of these problems when we were small as I said in class 11 or 12 or maybe in some of the schools did a the boards they teach in10 also.

(Refer Slide Time: 07:25)



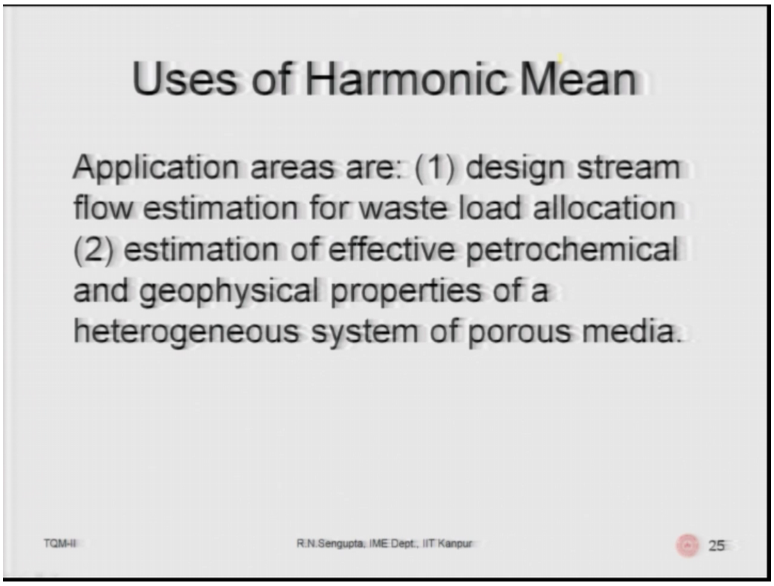
So, arithmetic mean so this is the symbol which we use this will become apparent time and again why we are using mu, mu would be a general common variable used to denote the mean; whatever the distribution is when estimating the long term expectation or random variable the arithmetic mean is a natural choice finding the averages age of the group of person average income of a group or person we use the arithmetic mean.

(Refer Slide Time: 07:48)



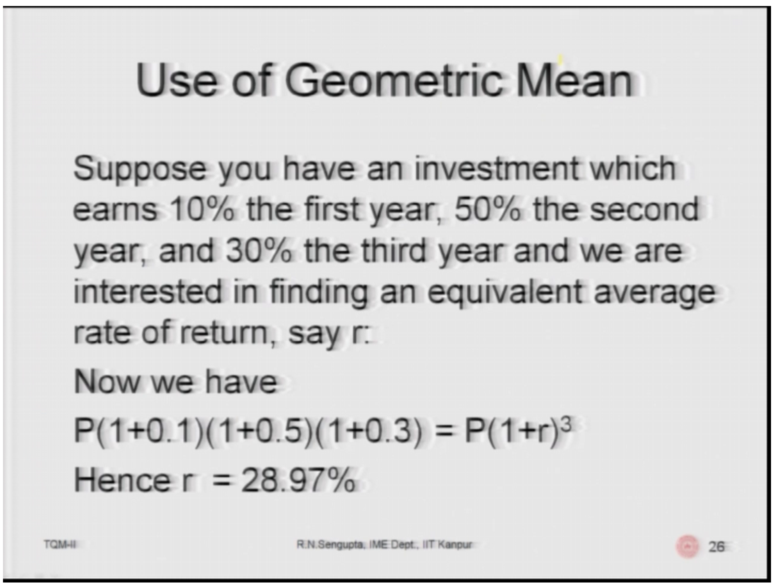
So, for the harmonic mean as I told you consider car travels between city a distance x between, city n to be with first in the forward direction from a to be with velocity v 1 and returns back with this in the same distance x from B b to a and with a velocity v 2. So, if you want to find out the average velocity you can find it out as given in the formula, so you see that you use the harmonic mean.

(Refer Slide Time: 08:16)



Further for the case of harmonic mean other examples are designed for the stream of flow estimation for waste loaded location or estimation of effective petroleum and geophysical properties on the on of a heterogeneous porous material you will basically use the concepts of harmonic mean.

(Refer Slide Time: 08:36)

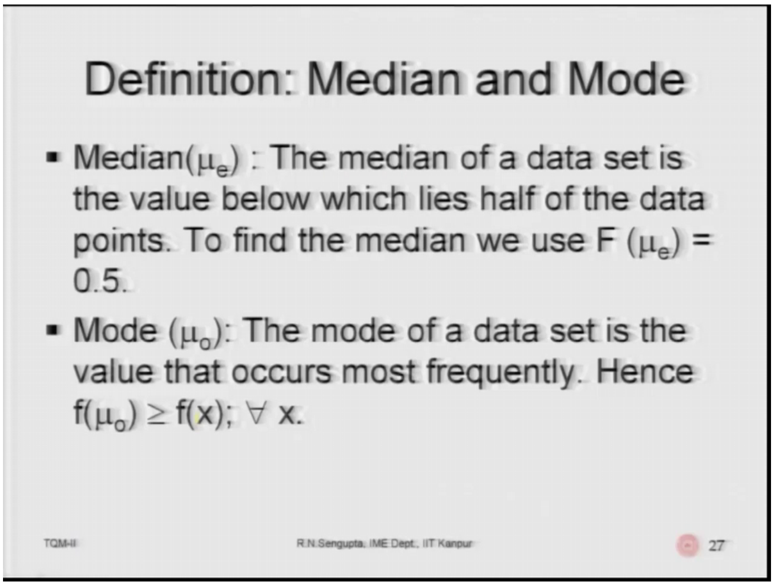


For the geometric means suppose you have an interest investment which earns10 percent in the first year 15 percent, in the second year 30 percent in third year and we are interested in finding an equivalent average return r. So, you basically find out if given P is the principal amount. So, the total amount of money which you get back after the first year considering the interest rate is point 1 would be P into 1 plus 0.1.

So, that is 1 plus 0.1, in the bracket as shown here where I am just hovering the this pointer, then it will be if we if I invest take out that money again invest in the second year it will be multiplied by, by the by the interest rate corresponding interest rate of the second year which is now 50 percent. So, it will be 1 plus 0.5 that will be multiplied by the first term from the money which you get back and basically go correspondingly. So, if I want to find out the average interest rate it will be P into 1 plus r to the power n, so here it is 3 years so it n will be 3 and you can find out r.

So, it comes out to be 28.9, so the median and the more definitions they would not be used. So, much in design of experiments, but will still, define them for our own convenience.

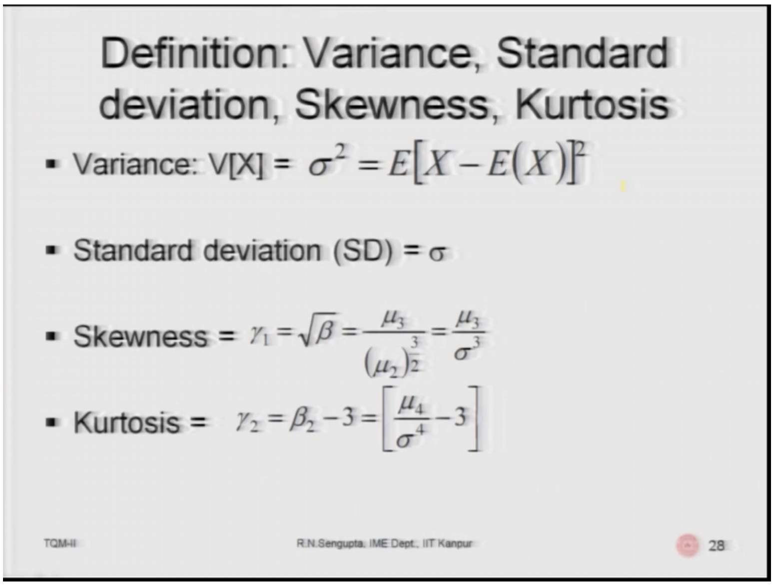
(Refer Slide Time: 09:57)



So, the medium of our data set is the value below which lies half of the data, which means the probability is being divided at that point to 0.5 0.5, to find out the medium we use the this capital F as we know is basically the CDF value. So, the CDF value till that point the sum of the probabilities till that point will be 0.5 or 50 percent and mode will be given the mode of a data set is the value that occurs most frequently.

So, hence we will use the value as given that the maximum num number for which value the frequency of the relative frequency of the probability the highest and we will find out that value as the mode, mode is may not be may unique it may be they can be more than 1 modes also.

(Refer Slide Time: 10:43)

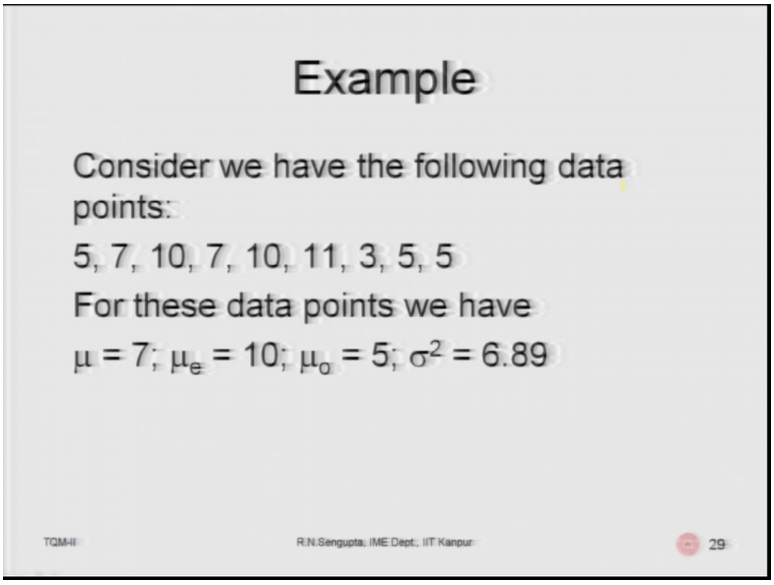


So, definitions of the variance standard deviation and skewness are for dispersion. So, we will use this 3 term, but mainly we will be using standard deviation or variance. So, variance is given by the expected value of the average of the square of the difference.

So, in each term we take the difference between the realest value and the and the average value square it up and add them or you find out and multiply by the corresponding probabilities, or add them and divide by n whatever it is or in the case of continuous case it will be an integration.

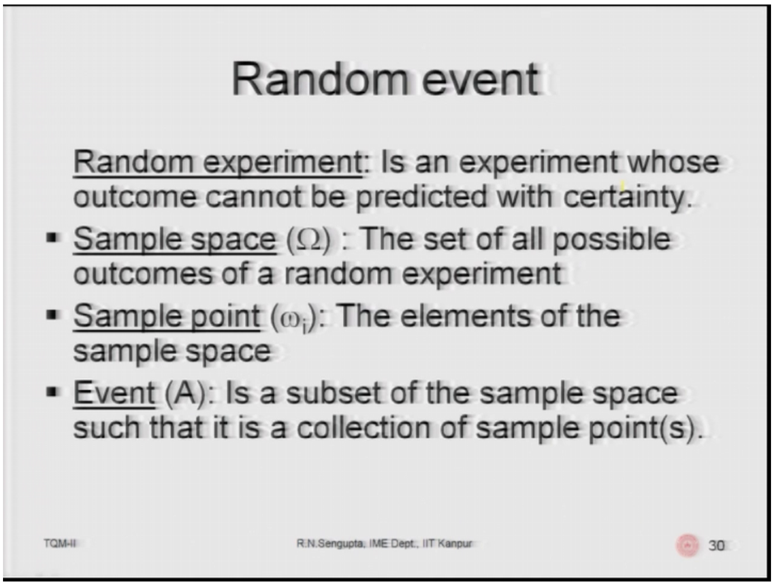
So, square standard, standard deviation will be the square root of that and the skewness and kurtosis are given as the formula they are just I am giving you just for the interest rate we will utilize that if at all required in the in the later cases.

(Refer Slide Time: 11:30)



So, consider we have the following data points 5, 7, 10, 7, 10, 11, 3, 5, 5 from these data, data points we have mu value as 7 and then mu e as tenth mu not as 5 and so on and so forth. So, based on that we find out sigma square this comes out to be 6.89, now what is a random experiment.

(Refer Slide Time: 11:53)

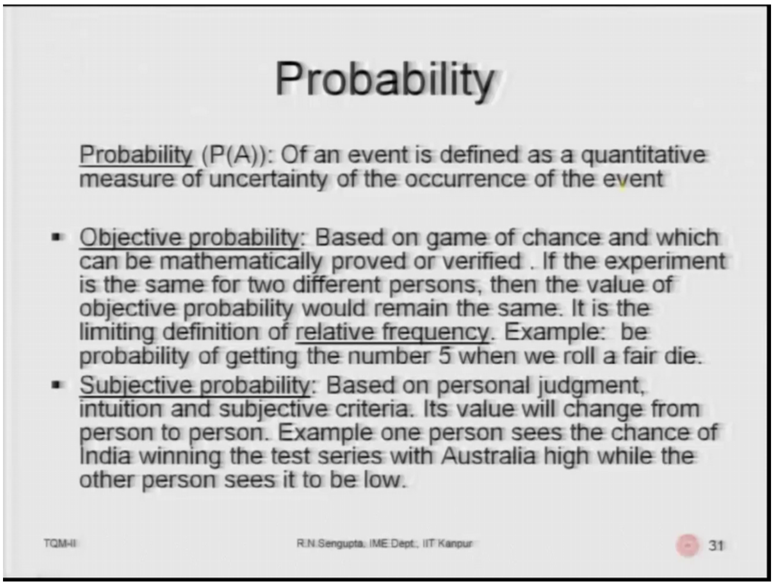


So, now, random experiment is an experiment whose outcome cannot be predicted with certainty, so consider you're rolling a die or you are tossing a coin or you picking a chits from a box. So, in that case it is a random event because we do not know which what will be the value of the outcome.

So, so the random event or the random value or the random variable would be the case based on which you are going to study and the realest value would be the actual value of the random variable or the event. Sample space is the set of all possible outcomes of a random experiment and sample points would the elements of the sample space based on which you will try to find out how many sample points are or they can be infinite also in the actual sense.

Well event is a subset of the sample space such that is a collection of all the all the sample sets or points, corresponding to the experiment which you are divide trying to basically find out say for example, I am rolling a die. So, the sample points would be 1, 2, 3, 4, 5, 6, so now, if my event or an experiment is to find out the even number only. So, the numbers coming out favourable would be 2 4 6. So, the collection of those sample points would be 2 4 6 based on which will lead to our experiment, and will define, define the event accordingly, so probabilities.

(Refer Slide Time: 13:15)

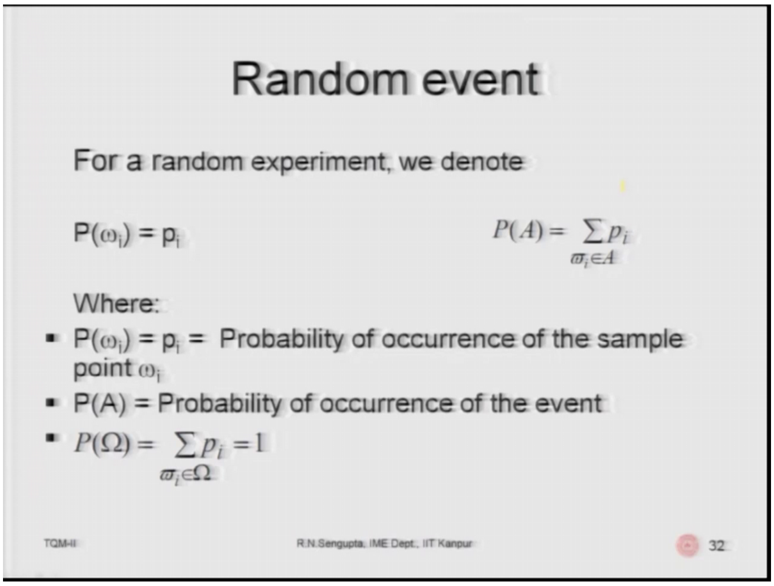


So, probability of an event is defined as a quantitative measure of uncertainty of the occurrence of the event, objective probability basically would be defined on the game of chance and which can be mathematically proved or verified and the experiment is the same for 2 different persons then the value of the objective probability would remain the same, it would not change it is the limiting definition of the relative frequency because if you remember we are taking the chances of the relative frequency in the long run as I increase the number of observations the sample points goes to infinity, so that limiting ratio basically is the probability.

Example can be the probability of getting the number 5 when we roll the fair die for your day means unbiased die, so if you roll it. So, the probability we know is 1 by 6, but actually it would happen that is you keep rolling them in the numerator will have the how many number of times the number 5, came and in the denominator it would be the number of times we did the rolling or the did the experiments.

So, subjective probability is based on the personal judgement intuition and subjective criteria it is value will change from person to person, example 1 person sees the chance of India winning a test series with Australia high well the other person sees it low so; obviously, subjective. So, it cannot basically be based on the fact that there is some underlying very sound theoretical concept or material concept based on which we can pass the judgement.

(Refer Slide Time: 14:37)

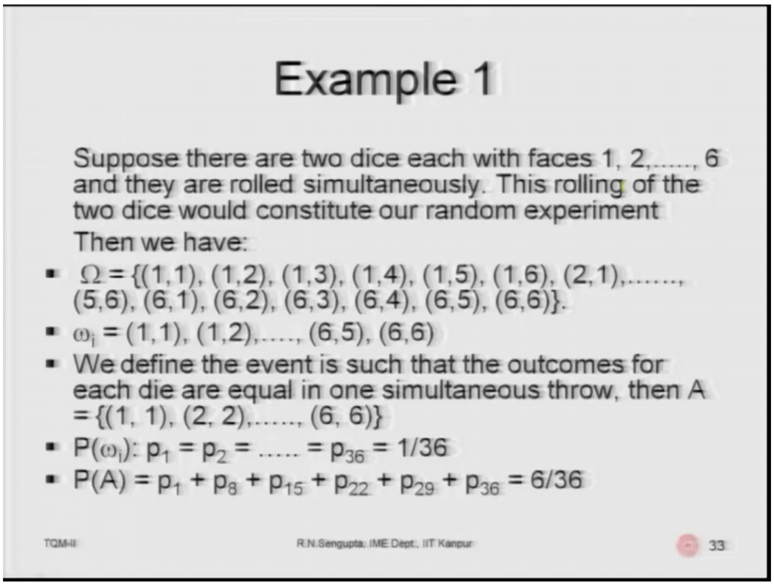


For a random experiment we denote the probability corresponding to value when, when the random value is basically takes 1 1 sample point. And that sample point would have basically have a probability corresponding to that which will be denoted by small P suffix I so; obviously, if I collect all the probabilities for all the sample points. So, it should be 1 that is point 1 which is basically given in the last bullet point where I am where I am pointing my finger.

And if I want to find out say for example, any if set a corresponding to the event e or a so; obviously, it would mean that I add up all the probabilities for all the sample points which are under the set a and basically add it up and find out the probability. So, now, here is what, what I mentioned are there are the important points I am just repeating it.

So, probability of the sample point is pi is the probability of occurrence of the sample point, whatever the i-th 1 is and P a is the probability of the occurrence of that event corresponding to those case, where you add up the probabilities of all the sample points which are a constituent assembly of the event and; obviously, the probability of the universal set there is addition of all the sample points probabilities would definitely be 1 and a for the null set; obviously, it will be 0.

(Refer Slide Time: 15:58)



Suppose there are two dice each with faces 1, 2, 3, 4, till 6 and they are rolled simultaneously. So, they are two dice this rolling and the 2 dice will constitute our a random experiment. So, these and the values of the universal set can be in the first roll you will get a 1 in the second roll the second now you get a 1, other case can be in the first roll you get a 1 second roll you can get a 2. So, if you find out all the combinations it will be 1 1, 1 2, 1 3, 1 4, 1 5, 1 6.

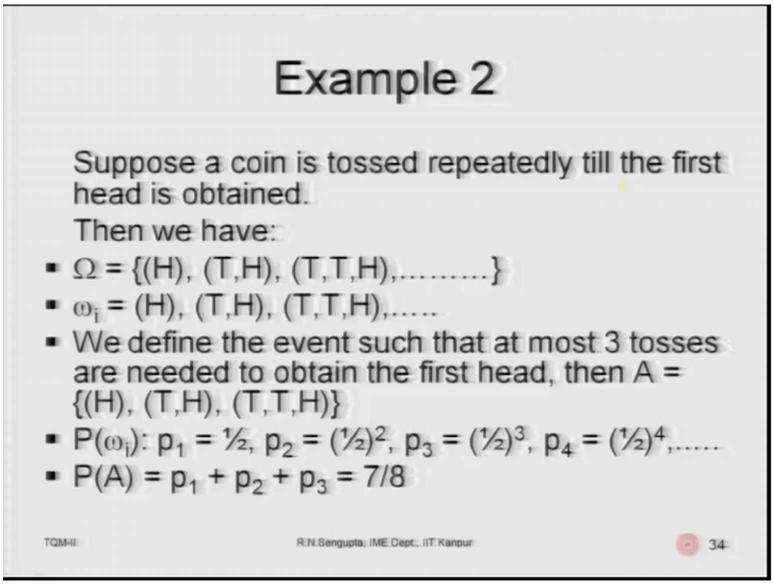
Now, if you basically go into the next stage it will be I roll that first dice comes a 2. So, corresponding to 2 you can have again from 1 to 6, then corresponding to 3 again you have basically 1 to 6, corresponding to 4, 4 means 2 3 4 I am basically mean the outcome for the first die to corresponding to 4, it will be again 1 to 6 this 1 to 6 are basically for the second die, then for 5 it is 1 to 6 6 it is 1 to 6.

So, if I basically find out all the combinations it will basically start if you see the first bullet point from 1 1 to 6 6. So, 6 6 is here, so if you want to denote the case that many roll the die the 2 faces are same that the total combinations corresponding to that experiment or event would be the set 1 1 that is 1 1 1 sample point second 1 is 2 2, third is 3 3, fourth is 4 4, fifth is 5 5, sixth is 6 6.

So, collection of that even basically satisfy the experiment which we have done. So, you define the events in such a way that the outcomes for each die equal in 1 simultaneous row then a would basically as I said 1 1 to 6 6. So, if I want to find out the probabilities, so each of them would basically be 1 by 36, because 1 by 6 is the probability of getting either 1 or 2 or 3 or 4 or 5 or 6, considering is unbiased die similarly for the second 1 it is also 1 for all the various from 1 to 6 1 by 6.

So, if I find out the probabilities it will be 1 by 6 into 1 by 6 which is 1 by 36. So, I want to find out the probability of that getting for this experiment which I just defined of getting the numbers coming out in the face 1 of die 1 and a face of die 1 and face of die 2 are equal so; obviously, they are 6 events. So, it will be 6 by 36 and the whole probability would be 1 by 6.

(Refer Slide Time: 18:14)



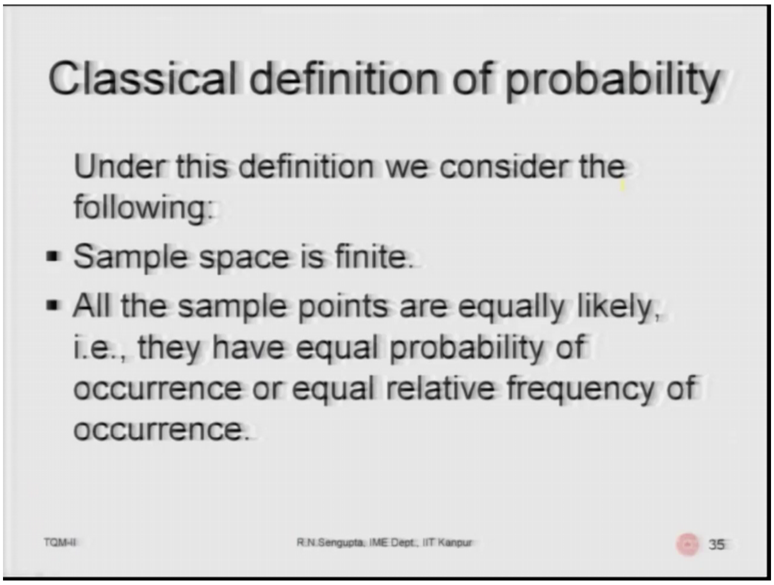
Second example is suppose a coin is tossed repeatedly till the first head is obtained, and we have the mu as there I am the first head obtained when the this the value of the universal set is basically if I get the first head. So, or if I get the first head in the second row if I get the first head in the third row if I just get the first head in the fourth row and so on and so forth so; that means, you can continue till infinity.

So, the overall universal set would be the sum the sample points would be H or T H or mean sorry. So, it should be basically a combination of all the points, so say the third point would be T T H, fourth point would be point means the sample points will be T T T H and continues till infinity.

Now, if I want to basically find out the corresponding probabilities. So, I want to find out that we need to define an every element that at least at most 3 tosses are needed to obtain the first head. So, 3 tosses at most needed being I get head in the first, I get head in the second, I get head in the third.

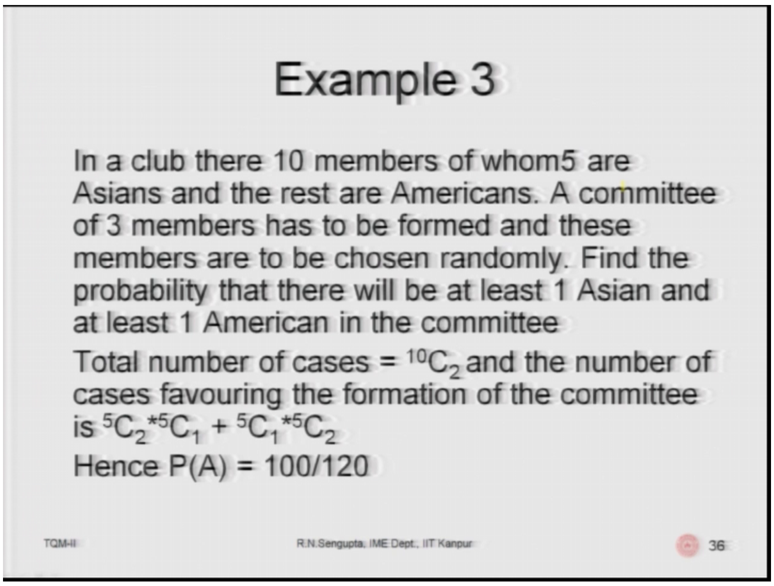
So; obviously, the probabilities would be in the first case for the sample point would be half for the second sample point half into half for the third sample point would be half into half into half. So, if you find out all the probabilities it comes out to be 7 by 8.

(Refer Slide Time: 19:42)



So, the classical definitions of probability defines that under this definition we consider the following, sample space is finite all sample points are equally likely that is they have equal probability occurrence or equal relative frequency of occurrence and based on that; we basically define the problems.

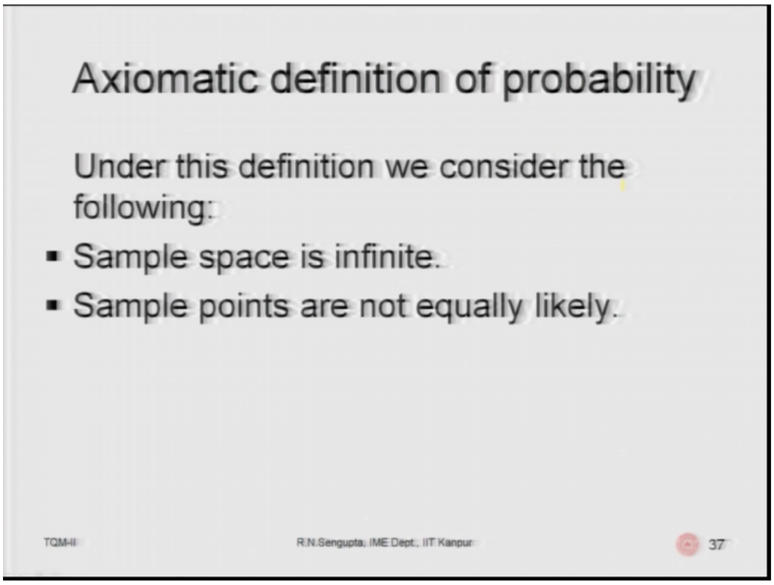
(Refer Slide Time: 19:58)



In a club consider there are10 members of whom 5 are Asians and the rest are Americans a committee of 3 members have to be formed. And these members are to be chosen randomly, find the probability that there will be at least 1 Asian and at least 1 American in the community the total number of cases would be10 C 2 and the number of cases favouring the formation the committees would be given; like if there are the at least 1 Asian and at least 1 American the combination can be 2 Americans 1 Asians and or 2 Asians and 1 American.

So, the combination will be 5 C 2 into 5 C 1 that is the first case; that means, consider 2 American and 1 Asian and the second case would be 1 American and 2 Asian again it will be 5 C 1 into 5 C 2. So, the total probability comes out to be 100 by 120.

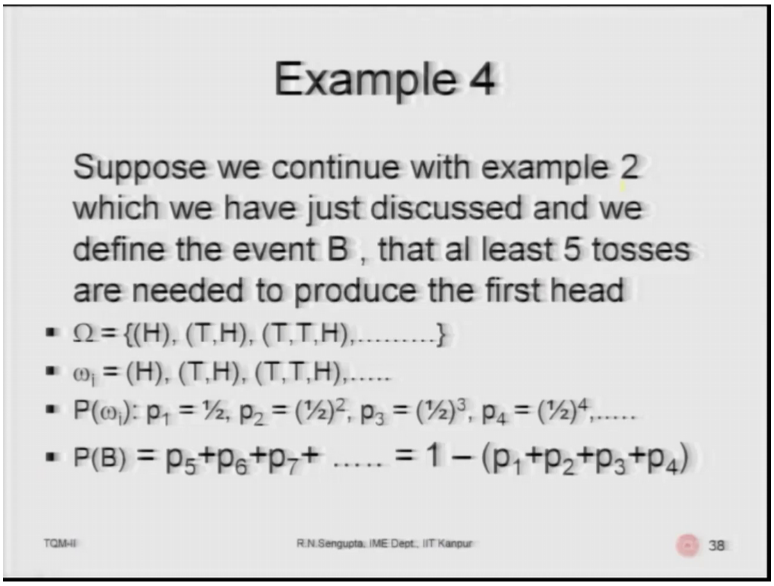
(Refer Slide Time: 20:44)



So, the axiomatic definition probability says that under this definition we consider the following sample space are infinite sample points are not equally likely. So, if you remember the problem which we solved like getting the first head. So, the sample points are w 1 w 2 small w 1, someone till infinity, but the corresponding probabilities are different because if I get the head the first time the probability is half. If I get the head the second time in the second throw it will be half into half.

If I get the head in the third throw it will be half into half into half which is 1 1 by 2, cube and I cut deduce it concentrate on that it would mean that the probabilities are unequal and we can solve the problems considering them the axiomatic definition is true, sample months are not equally likely because that is what I said that the probabilities are half one-fourth, one-eighth and correspondingly it goes on and they are infinite because the, when you will get the head we do not know so; obviously, the set of a sample point is infinite.

(Refer Slide Time: 21:52)

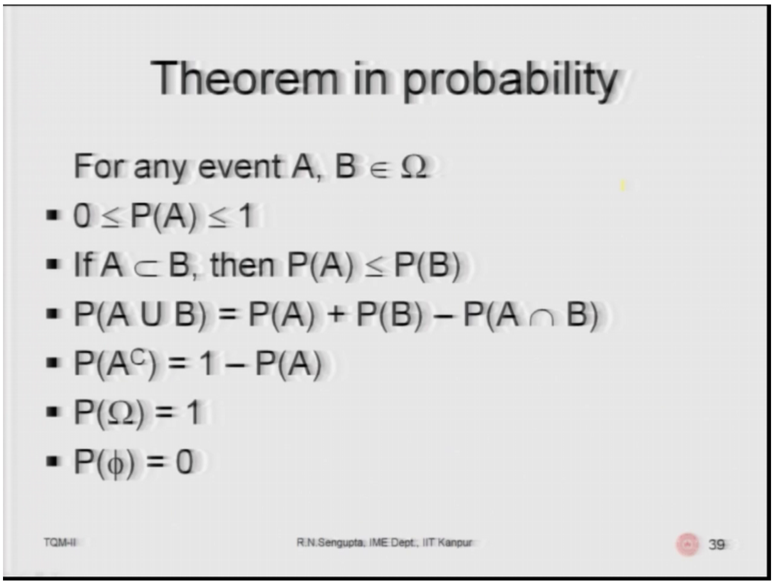


Suppose we continue with example two where we have just discussed and we define the event B, so that at least 5 tosses are needed to produce the first head. So, the sample space would be either a head, then its tail head then its tail, tail head and continues and the sample points are its the head which is standard like the first sample point then the tail head is the second sample point, then tail, tail head is the third sample point, and continues though the corresponding probabilities are half, one-fourth, one-eighth and then one-sixteenth , and if I want to find out at least 5 tosses I needed to produce the head so; obviously, the head can come in the in the fifth toss.

So, first four are tails, fifth is head or the second one would be first 5 excuse me, 5 are head sixth is it 5 or tail sixth is head.

Next would be for sixth a tails seventh is head and it continues till infinity. So, if I want to find this it would be better if I subtract the particular value what is that particular value I will tell you. So, I basically I get the head the first time the head the second time head the third time the first second third fourth what I mean is basically I am getting in the first trial second trial third trial fourth trial. So, I basically add these values and subtract this value from 1 I will basically get the answer. So, it as is given 1 minus in the bracket P 1 plus P 2 plus P 3 plus P 4 would may the answer.

(Refer Slide Time: 23:27)

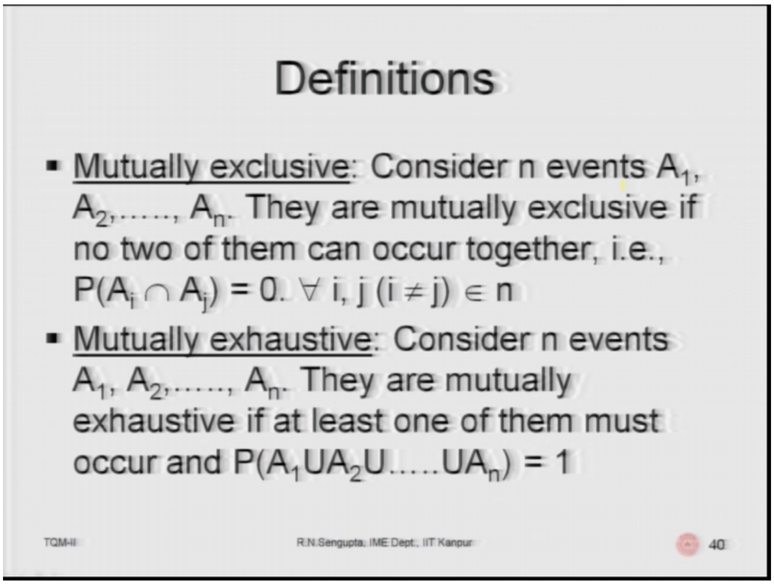


So, in theorems in probability is that for any event A and B being n member on the universal set, the probability of P on any event between is always between 0 and 1 if a is a subset or A proper subset of B then the probability of a would definitely be less than equal to B because probability cannot go negative.

So, it has to basically increase it can be 0 so; obviously, in that case P A is equal to P B, then the concept of the of the union sets or probability of a intersection B P a would be plus P B would be basically intersection B would be P A plus P B minus the overall area common between A and B and P A complement would be 1 minus P A.

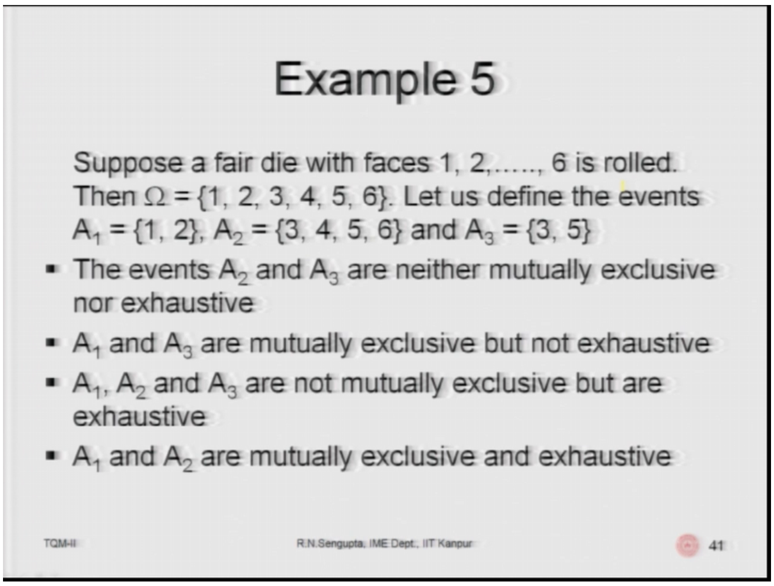
So, if it is a complimentary set and a A complement plus A is the universal set then the probability of a would basically be 1 minus the probability of a complement or vice versa, probability of a null set is 0 as given the last point and the probability of the universal set is 1.

(Refer Slide Time: 24:29)



So, we will basically discuss now mutually exclusive. So, consider n events A 1 to A n they are mutually exclusive if no 2 of them can occur together, which means P A i intersection A j is equal to 0, for any i and j being a member of n. And mutually exclusive exhaustive would be considered and n events A 1 to A n they are mutual exhaustive least 1 of them must occur which means the union of A 1, A 1 union A 2 union A 3 dot till a union A n would basically be 1.

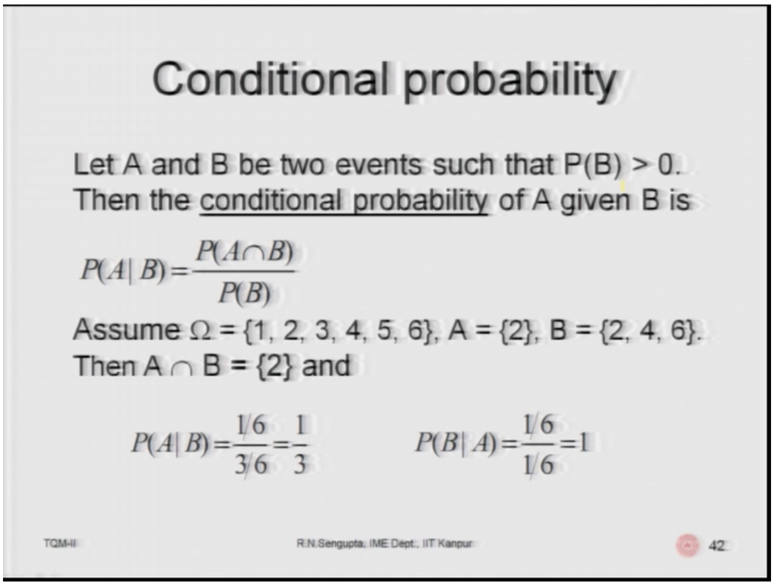
(Refer Slide Time: 25:05)



So, suppose a fair die with faces 1 to 6 are rolled the universal set; obviously, would be 1 to 6 let us define this events accordingly. So, A 1 is given by 1 2. So, the sample points are 1 and 2 A 2 is given by the sample minus 3 4 5 6 separately A 3 is given by the sample points 3 and 5.

So, if I basically look at A 1 and A 2 and A 3. So, I will basically can define the following the events A 2 and A 3 and either few mutually exclusive and exhaustive A 1 and A 3 are mutually exclusive, but not exhaustive A 1, A 2, A 3 are not mutually exclusive, but are exhaustive and finally, A 1 and A 2 are mutually exclusive and exhaustive.

(Refer Slide Time: 25:43)

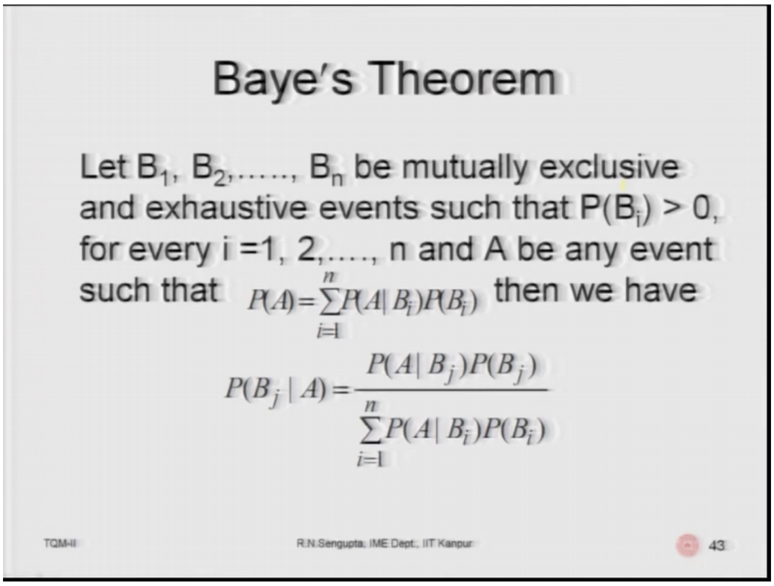


So, when you are coming to the concept of conditional probability, so let A and B be 2 events such that probability of B being greater than A, A is true. So, it is not as is null set then the conditional probability gave of A given B would be given by in the numerator you will have the probability of a intersection B, and in the denominator you will basically have P B, so let me define this.

So, assume the universal set is again the rolling I mean the rolling of the die dice unbiased dice and the numbers are as I mentioned is 1 to 6. So, and now I define, define events A as the even number 2 and B as the set of all the name event numbers between and between, between 1 and 6 which is 2 4 6. So, if I find out the intersection of A and B it is 2 and then if I find out that given B has occurred what is the probability as the A B to occur is given by if you note down, but I am just now pointing my fingers, it would be 1 by 6 which is the probability of a intersection B and the probability of B would be there are 3 occurrences of the total 6.

So, it will be 3 by 6 and the corresponding probability is 1 by 3, if I want to find out the probability of A of B given A has occurred; obviously, it will be 1 by 6 by 1 by 6 because the moment A occurs B has occurred, so the Baye's theorem.

(Refer Slide Time: 27:02)



Let B 1 to B n be the mutually exclusive and exhaustive event such that P B i. So, there is the P B i's being greater than all of them are not equal to 0 and for every 1 to n and define A to B an event. So, that is basically the conglomeration of the base the conditional property is coming into picture which will be if I consider the overall summation. So, that will make sense that probability of a would basically be the summation of these n elements what are those n elements probability of A intersection B i divided by P B i.

So, i tends to 1 may 1 to n, so, on what I have is basically I find out the intersection of A and B 1 that probability divided by the probability of B 1 would be giving the first element, next element which will add would be probability of A intersection B 2 divided by B 2 and I continue and find out the sum which will give me A, and if I want to find on the other way around probability of B j or B i, whatever it is given a is given by this formula which is basically the concept of conditional distribution on the Baye's theorem.

So, with this I will end the second lecture and continue the discussion till we start the design of experiment have a nice day.

Thank you.