Business Analytics for Management Decision

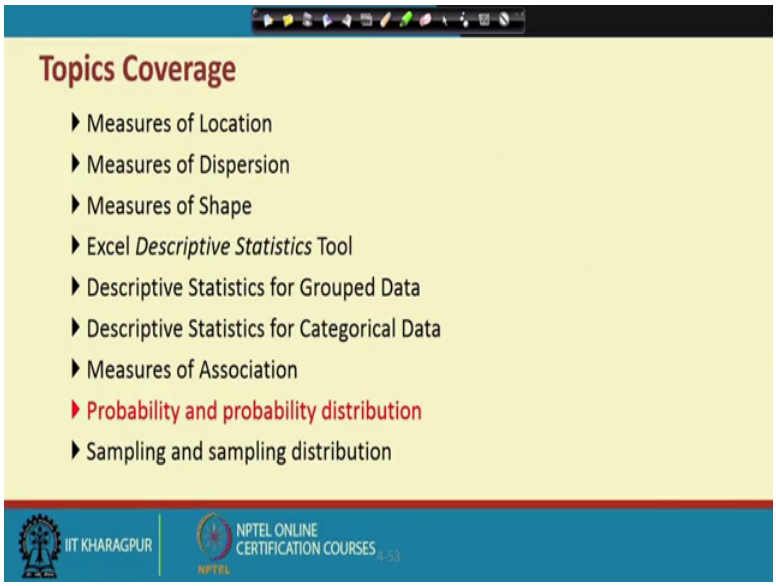
Prof. Rudra P Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

Lecture - 13
Descriptive Analytics (Contd.)

Hello everybody, this is Rudra Pradhan here and welcome you all to BMDA course. We are here to continue the component called as a descriptive analytics and in fact, this is the third lecture of this series on descriptive analytics. In the last two lectures we have already discussed about the descriptive test statistics and the association test statistics. Through descriptive statistics we can actually do or we can take some kind of management decision and again by using some kind of association techniques you can also solve some the business problem and can in a position to take some kind of management decisions.

But whether it is a kind of descriptive test statistics or some kind of association test statistics or something like predictive analytics or prescriptive analytics, we need actually something basics through which we can actually analyze all these techniques in a more attractive way. That means, what I like to say that some of the basic things we are supposed to discuss again in order to make these analytics you know in a kind of more attractive way. So, here two such requirement is called as a probability and probability distribution and sampling and sampling distributions.

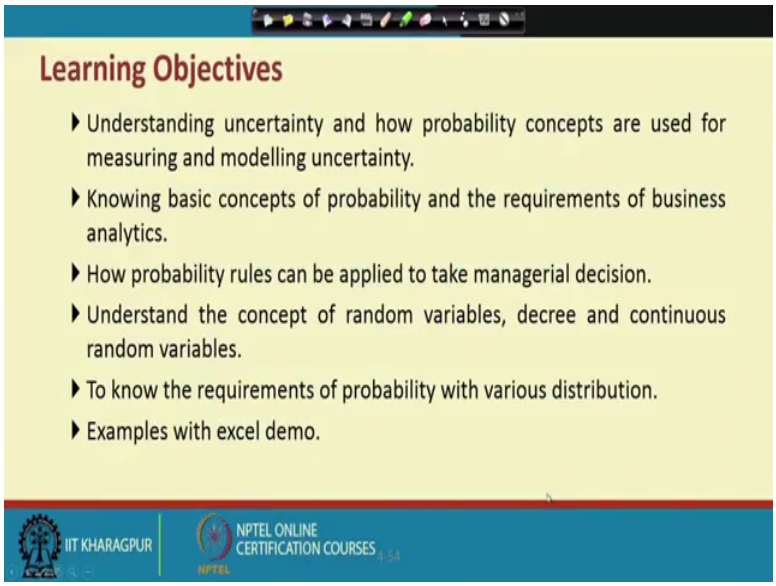
(Refer Slide Time: 01:43)



Because we are living in the kind of complex environment, all things are you know very dynamics and every corners you will find there is a kind of reaches there is lots of on uncertainty. So, dealing with your (Refer Time: 02:04) uncertainty. So, you have to actually predict the kind of environment and then solve the problems and you must be in a position to take some kind of management decision. So, since you know reeks and uncertainty are you know is most of you know, related to most of the problems. So, you are supposed to actually connect with the probability and probability distribution and sampling and sampling distribution.

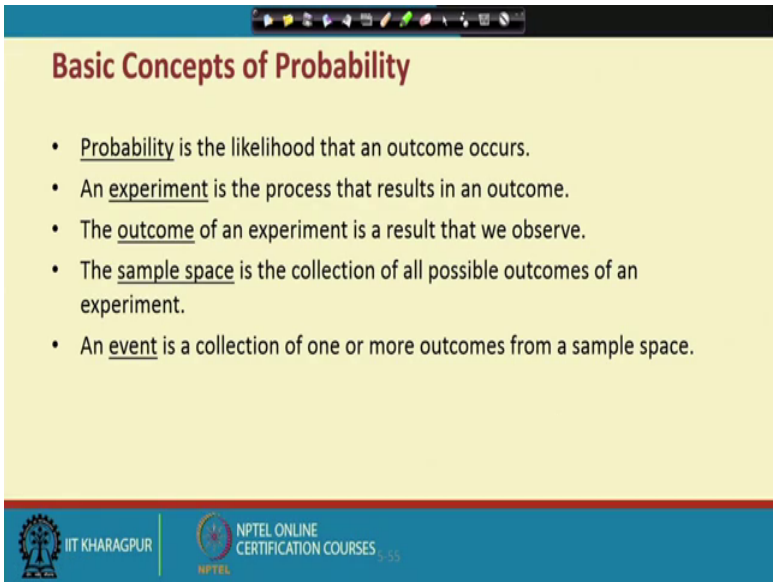
Knowing all these things are you know very handy or very useful for some kind of you can say better a better investigation or something kind of better inference. So, that is why we start with the first the concept called as a probability and probability distributions. In fact, it is not something new, but still we are in a position to know something related to probability and how it is useful for some kind of business analytics structure.

(Refer Slide Time: 02:48)



So, the learning objectives are here understanding uncertainty and how probability concepts are used for measuring and modeling uncertainty. Then second one knowing basic concepts of probability and the requirements of business analytics. And then again you like to know how probability rules can be applied to take some kind of managerial decision. Then understanding the concepts of random variables and discrete and continuous random variables, and to know the requirements of probability with the various you know distributions that is nothing, but called as you know probability distributions. And then we will discuss with some examples with you know excel spreadsheet demo right. So, now, with this we will like to move to the concept called as you know probability.

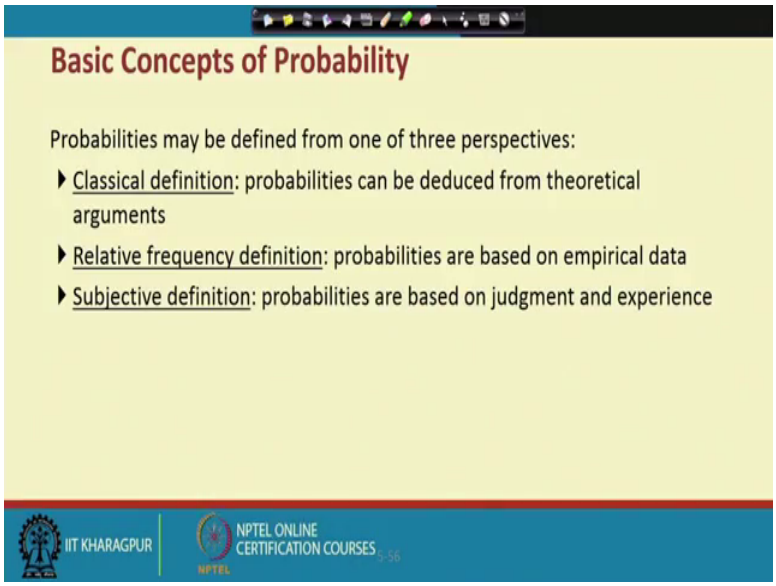
(Refer Slide Time: 03:48)



So, probability you know many different ways you can define it is the likelihood that an outcome can occur right. So, it is just chance sometimes it is called as chance of occurrence or chance or believe all these things will be you know represent the concept probability. And it is an experiment and that too it is the process that results in an outcome. The outcome of an experiment is a result that we observed in usual structure. Then there will be a sample space that is the collection of all possible outcomes of an experiment. For instance if you toss a coin or roll a die, so you will get some kind of possible outcomes.

Then an event is a collection of one more outcomes for a sample space so that means, it is the sample space where you know all the events can be taken into considerations then as per the particular requirement you have to specify a particular case. Like we have discussed the game between population and samples probability concept is also similar kind of nothing. So, here we have to see the what is the possible kind of cases then you have to choose your sample specific case, and the sample specific case out of the total possible cases you will give you some kind of clarity or some kind of concept through which you can come to a conclusion and then accordingly you have to take some kind of management decisions.

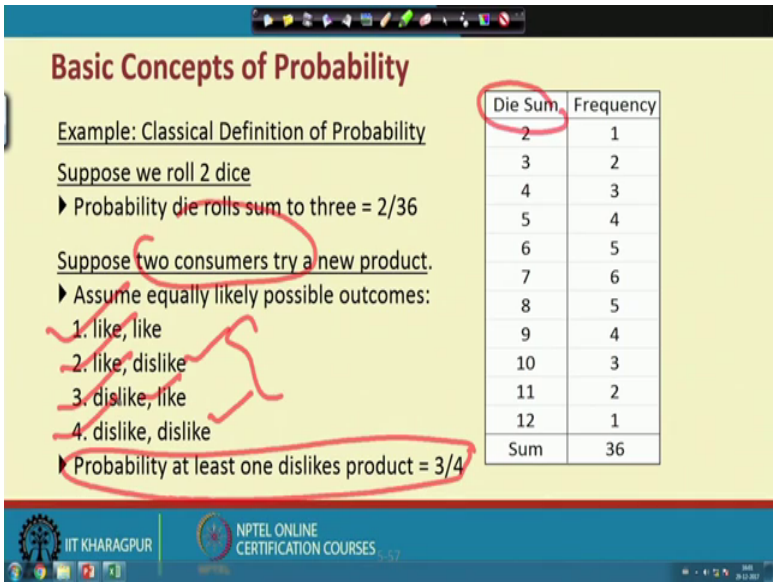
(Refer Slide Time: 05:31)



So, let me give more details about this particular you know concept. So, there are historically the probability is not something new it is a very old concept, but it is very interesting, very useful and it is the requirement for any kind of business analytics. Any kind of means its most of the analytics tools a probability kind of well connected is a kind of linkage through which you can have a better kind of judgment. So, there are you know various definition here starting with the classical definitions. So, here in the classical definitions probability can be obtained from theoretical arguments and then we have relative frequency definitions and probabilities are based on empirical data then we have a subjective definition probabilities are based on judgment and experience.

So, these are all various ways you can actually calculate the probability, to understand the probability and then you have to use the kind of know probability concept as per your particular requirement. So, we know whatever we as we have been notified. So, I will connect with some of the examples that you can actually you know easily understand.

(Refer Slide Time: 06:48)



For instance, here actually the classical is corresponding to classical definition of probability. So, suppose we roll 2 dice right. So, the probability die all sum to 3 right. So, when you roll to die simultaneously, then you will be find couple of you know occurrence, the occurrence will be 1 1, 1 2, 1 3, 1 4, 15, point 6 then again 21, 2 2, 2 3, 2 4, 2 5; that means, you will defined plenty of such you know combinations.

So, now, so the typical structure is here is that you know we need is, how to understand the kind of probability here is we, in fact, we have discussed something called as you know frequency then you know cumulative frequency in the previous units, but now this will be very useful in this particular case that to understand the concept of you know probability. Now here actually die sums., 2 die simultaneously you are throwing then you know what is the sum we can finally, obtain. So, these are the possible sums when you when will be parallely toss the 2 die then you will have 2 3 4 5 6 7 8 9 10 11 12. So, now, these are the sums we have actually then finally, how many frequency you can obtain from this sum, frequency 1 2 3 4 5 6 5 4 3 2 1. So, this is actually some kind of structure which you have to follow here. So, this by default will give you some kind of exposure about the particular you know case or to understand a probability concept.

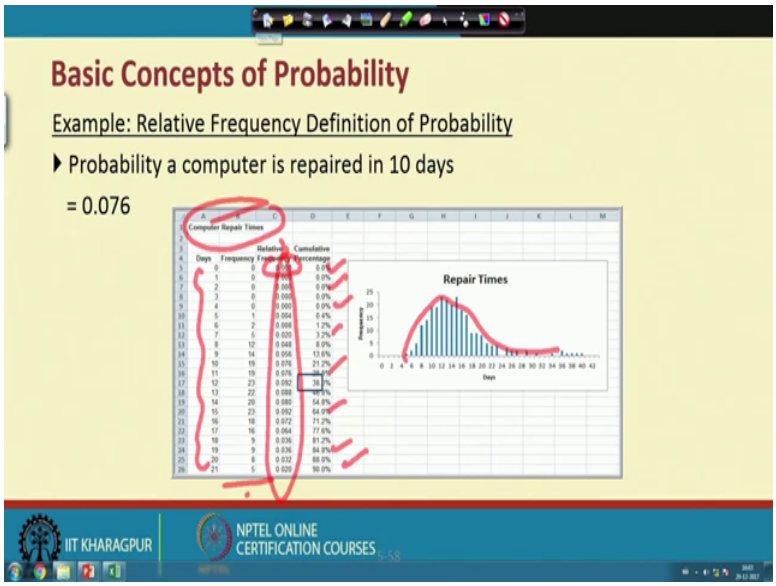
So, now, the another kind of an example see here is let us assume that there are two consumers and they consume two you know products and the observation which you have here. So, far as a probability is the concern is you know they may means two

products simultaneously there is a high chance that particular guy like both or like one dislike one again dislike the first one and liking the second one and there is a high chance that the guy may dislike you know both. So that means, all to get that there are 4 different possible cases.

So, probability of at least one dislike probability of, so here the question you know this is a specific you know questions which you like to address; that is what actually we need actually kind of management decision. So, what is the probability of you know at least one dislike. So, that you know you can apply this strategy accordingly and you can design the problem as per your requirement. So, at least one dislike means you see here, at least one dislike. So, here one and here one and here also at least one so that means, there are you know 3. But what is the total possible cases? Total possible cases are you know 4. So, this is 1, this is 2, this is 3 and this is 4.

So obviously, probability of you know at least one dislike product will be by default 3 by 4. So, this is how the literary understanding of you know concept called as you know probability. So, in order to know more about the probability then we will move to you know next structures see here.

(Refer Slide Time: 10:16)



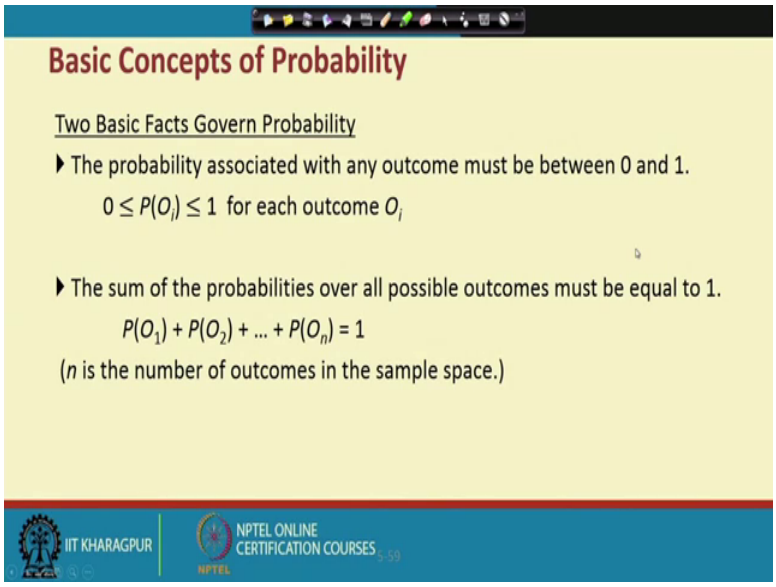
So, this is another kind of structure which is connected with you know frequency. So, now, let us say there is a kind of computers and computer you have just have a kind of computer systems and you are editing something else and how many times you need

actually the kind of maintenance. So, that is the kind of repair times. So, starting with you know 0 days to 21 days. So, this is how I have you know this is what the game and the day wise information is available here. So, first day, so frequency of repair is 0 and second first day, second day, third day and continue up to 21.

So, now you will find the frequency, frequency of repair times. So, it will be 0 0 0 0 then 1 2 5 12 14 like this. So, now, you can prepare a relative frequency and that relative frequency by default will give you snap sort of actually the concept called as probability. So, what we supposed to do here. So, this is actually you know individual frequency you take actually total frequency here, then every individual frequency by total frequency. So, then the probability can be obtained. So, for each individual case you just divide the total frequency then you will be find. So, this series will be obtained for each individual case. Then you can actually just add you know one by one then you will get you know cumulative frequency. These are all cumulative frequency then ultimately if you plot all these you know with respect to day and the frequency the frequency curve will be coming like this.

So, this is what we have already discussed in the last, in last lecture and that to you know you need unit 2. So, this is again you know same concept, but it is connecting to the concept called as you know probability. So, in order to know a little bit more, then I will give you here the kind of basic things about the probability.

(Refer Slide Time: 12:21)

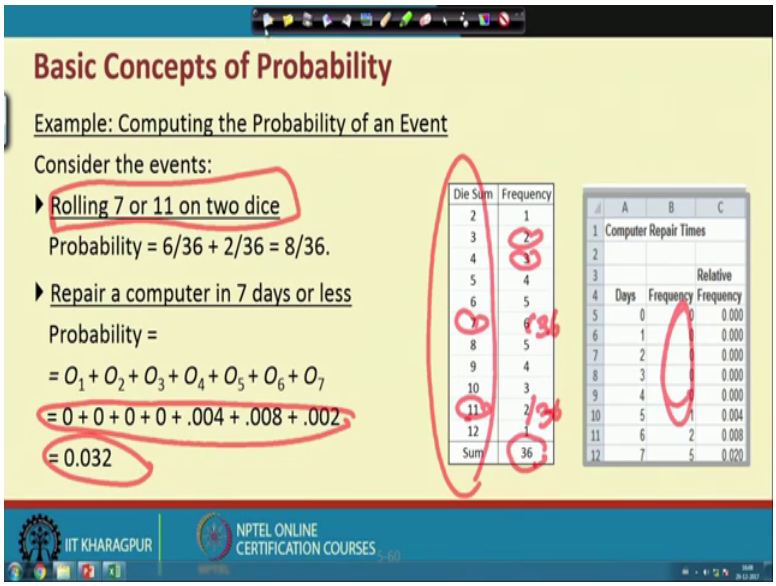


So, like you know earlier discussed the concept called correlations which is always between minus to minus 1 to plus 1. So, here the probability is also having basic features without any kind of understanding or any kind of issue whatever may be the problem structure or whatever may be the problem background. So, when you are dealing with the probability the value of the a probability will be always positive and it will be always lying between 0 to 1 and depending upon the possible you know outcomes. So, the sum of the probability will be exactly equal to 1.

So, these are the 3 important features you always keep in mind when you solve any kind of problems connecting to probability or you are using probability for any kind of experimental or empirical kind of investigation process. So, here some of the things are you know, so what we have already mentioned here. So, this is what actually is the kind of concept of probability. So, this is what actually the concept of probability. So, here, this is the probability of occurrence and this will be in between stand between 0 to 1. So, 0 means no occurrence, but 1 means you know it is fully predictable, then sum of the probability will be actually exactly equal to 1.

So, this is individual probability case 1, case 2, case 3. So, number of possible cases you fully add the probability of all possible cases then this sum will be always equal to 1. So, this is again something called as you know standardized structures like you know correlation matrix. So, now, this will be, this would be very useful for any kind of investigations and then accordingly you can actually take some kind of management decisions.

(Refer Slide Time: 14:24)



So, now, move to further understanding. So, whether it actually means how this value of probability will be positive and whether this sum will be exactly equal to 1. So, we will connect with a particular example here. This same example which you have discussed earlier then 2 die we are actually rolling and then what is happening. So, this particular you know case which we have already which you have already highlighted. So, this die sum. So, now, there is a high chance that you know I would say you know sum is it 2 means then there is a high chance you know 1 1.

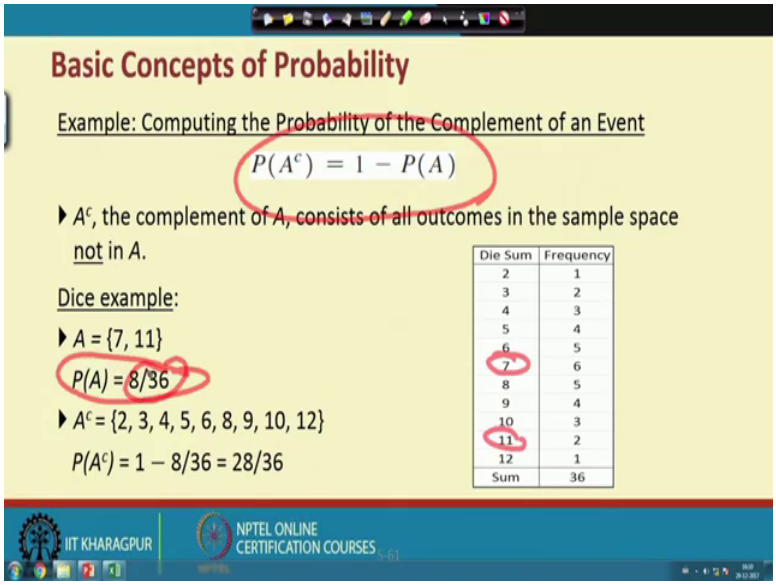
So, if it is sum is it 3, a 3 means. So, there is a high chance that the possibility is 1 2 or 2 1 so that means, there is a 2 number of frequency. So, now, if the sum is 4 then there is a high chance that you know, so the occurrence will be 1 3 or 3 1 or 2 2 so; that means, the possible cases will be you know 3. So, then these are all individual frequency corresponding to the die 2 die sum and then the total frequency will be 36. So, now, corresponding to individual frequency if you divide the total frequency then it will be transferred into the concept called as probability right.

So, now here you see here. So, the question is actually rolling 7 or 11 on 2 die right, 7 or 11. So, 7 is here and then 11 is here right, so that means, it says 6 divided by the particular item 36 and then this is 2 by 36. So, now, if we will add up these 2 then the total occurrence will be 8 by 36. Repair a computer in 7 days or less, 7 days less means up you know, so first second third 4 5 6 and 7. So, now, accordingly, you have the

following you know frequency corresponding to the particular frequency you have to actually calculate. That means, with respect to these are all you know relative frequency and up to 7 days right, so 1 2 3 4 5 6 7 then corresponding the probabilities are here. So, you have to just you know add up all these you know possible cases then you will be find the total probability.

So, what is the total probability when you have the kind of, when you need actually you know some kind of managerial output to repair a computer in 7 days or less. Then this is how the structure you have to apply then accordingly you have to interpret that is the here 0.03 means it is actually 3 percent plus to 3 percent. So, now, again moving towards the concept of probability we can cite with you know more examples.

(Refer Slide Time: 17:09)



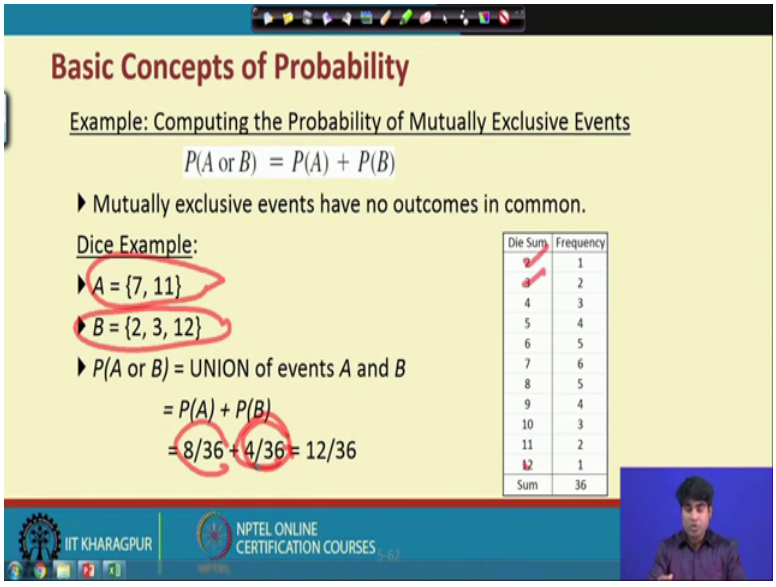
So, now, here is there is the concept there is event say A and since total probability is equal to 1 then you know event occurrence and not occurring occurrence. So, the sum will be exactly equal to 1. So, as a result, there is a concept called as you know A to the power c that is called as A, they complement of you know A this just the opposite side. That means, if the event is the occurring in your particular situation then the chance of non occurrence will be 1 minus probability of you know occurrence. So, that is how the structure is the structure is actually written like this. So, this is what actually a complementary of A. So, now, I here, sum you know examples, dice examples. So, here A is a kind of event and that the outcome is 7 and 11 which we have already highlighted

in the previous examples. So, now, the item will be 6 by 36 and 2 by 36 by default it will be having 8 by 36 that is the probability of A event.

So that means, if you are not a you know including a then 1 minus remaining you know 1 minus you know this much 1 minus this much means 8 by 36 will be the a complement of A. So, as a result, either you calculate this once then you know 1 minus this much will give you probability of you know A complement or else you go through individual kind of just you remove this ones then individually calculate for rest of the things. Then automatically you will get the same value.

Since it is a kind of standardized concept since probability exactly equal to 1 total probability exactly equal to 1; obviously, if you calculate the first one the second one will be by default obtained by deducting in the first one from this sum equal to 1. So, this is how you have to actually calculate and then connect as for the particular you know need.

(Refer Slide Time: 19:36)

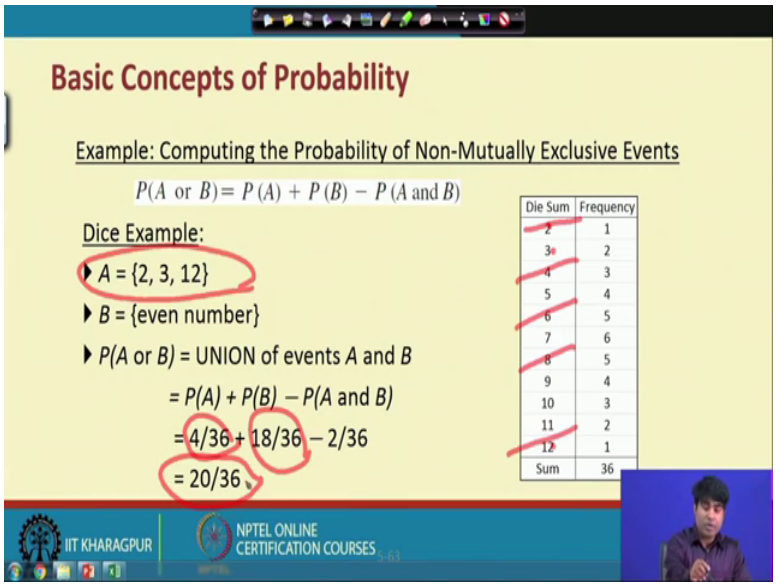


So, then moving towards you know further discussion about the probability, so I will connect with you know some of the rules behind the probability. So, there is a concept called as a mutual exclusive. So, P A or B so that means, event A so; that means, 2 events going simultaneously. So, probability of you know A or B, A or B. So, is if they are mutually exclusive then the probability of A or B equal to probability of A plus probability of B.

So that means, same things you know here is A is the event of you know occurrence, events of occurrence you know 7 and 11 and B is the occurrence of 2 3 and 12 right. So, then 7 11 which we have already discussed earlier so that is nothing but 8 by 36 and the occurrence of 2 3 12 so; that means, this is 2 3 and 12, so 1 by 36 2 by 36 and 1 by 36. So, this will be coming again 4 by 36. So, probability of A or B equal to probability of A plus probability of B so that means, technically probability of A is coming 8 by 36 and probability for B is coming 4 by 36. So, then probability of A or B will becoming 8 by 36 plus 4 by 36 that will be exactly equal to 12 by 36.

So, this is another way of calculating the probability and this may be useful for a situation and then you can actually take management decision accordingly. So, now, this is what. So, this is another kind of examples and the same examples.

(Refer Slide Time: 21:20)



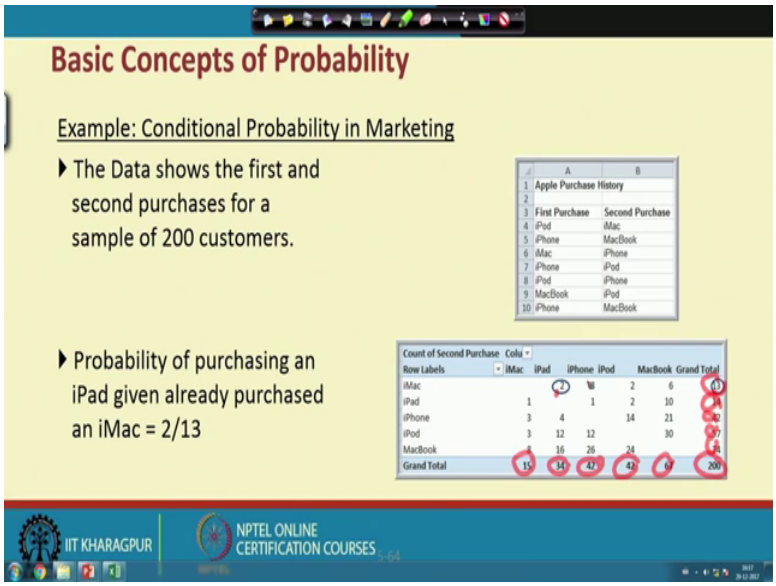
So, here we are again reporting it means we want to report probability of A or B when the events are not actually mutually exclusive, non-mutually exclusive events. So, two events A and B when they are not you know mutually exclusive then the probability of A or B will be probability of A plus probability of B minus probability of A and B, that means, here is for examples. So, this you know A is a event which is the sum of which is the sum of 2 3 and 12, so 2 3 and 12 then A B is the event which is the occurrence of you know even number on c, even number only, then that is that is nothing, but actually 2 4 6 8 then 12. So, these are you know possible case.

So, now, so far as you know first event is constant 2 3 12. So, that is nothing, but one by 36 2 by 36 then 1 by 36. So, this is coming actually 4 by 36 which you have already discussed right. Again so far as you know B even number is concerned then 1 by 36 then 3 by 36, 5 by 36, 5 by 36 then 1 by 36. So, this will be coming again 18 by 36. So, now, P A and B so that will be coming actually, 2 by 36, so 2 by 36, as a result, P A plus P B minus P A intersect P A and B will be giving you 20 by 36.

So, now compared to the previous situation compared to previous situation. So, when they are mutually exclusive, when they are mutually exclusive, then the probability then the probability is actually coming probability is coming 12 by 36 when they are mutually exclusive where the third component will not be there. But when there is actually you know they are not mutually exclusive then the common elements you know this P and B, need to be actually subtracted, as a result it will be giving you know 20 by 36. So, this is actually 12 by 36 and here the result is coming 20 by 36. So, depending upon the kind of problem, you have to actually calculate the, if the 2 events are mutually exclusive then the kind of prediction will be different when they are not mutually exclusive the kind of prediction will be different.

So that means, probability will give you some kind of exposures to predict the particular you know business problem depending upon the problem requirement and problem structure. So, accordingly, we can actually use probability and then predict the kind of environment.

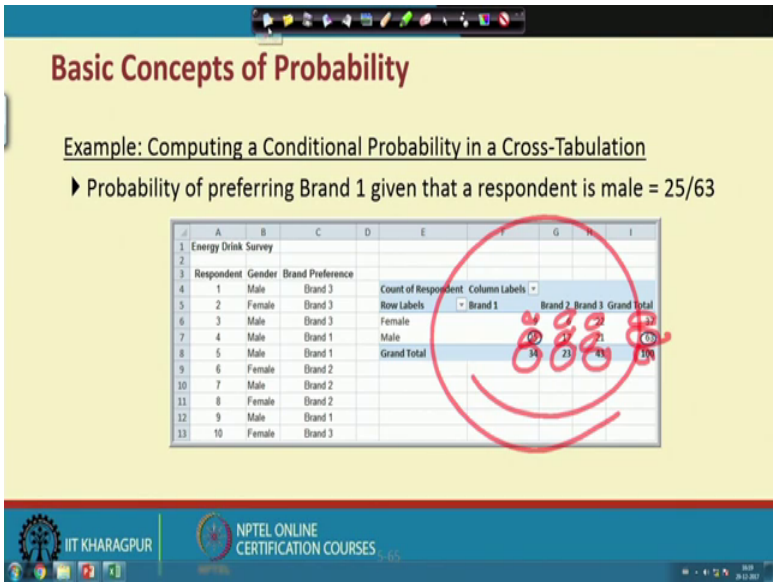
(Refer Slide Time: 24:18)



So, this is another kind of examples. So, this is a conditional probability there is a concept called as conditional probability and most of the fields you know we frequently use like operation finance marketing. So, I am giving here is a kind of example called as you know marketing example. So, the example is actually first purchase and second purchase and that is with respect to iPad and iPhone, iMac all these things will be there. So, we like to actually, we like to check you know what is the probability of purchasing and iPad giving already purchased an iMac. So, it is actually with these particular you know structures you may be in a position to calculate right.

So, this is what the first hand informations and this is what actually called as a cross tabulations which we have already discussed in the previous unit and then this concept actually, so these are all called as column totals, these are all column totals with 2 different for purchase and then this is actually row total and these are all row total and by default then sum of the column total and sum of the row total will be giving you the total sample. Then obviously, the probability of particular items will be like this. So, here probability of purchasing and iPad given already purchased and iMac. So, it is nothing, but actually core you know 2 and corresponding to this probability is a 13. So that means, it is a 2 by 200 and 13 by 200 then finally, it will be coming actually 2 by 13. So, means again this is little bit complicated kind of examples, but it is highly required or it is very useful some for some kind of business related problems.

(Refer Slide Time: 26:12)



So, now, corresponding to the corresponding to this particular concept of you know conditional probability. So, I will have I can cite here actually another examples. So, here, we have respondents and corresponding to respondents we are actually citing the gender and then branch reference.

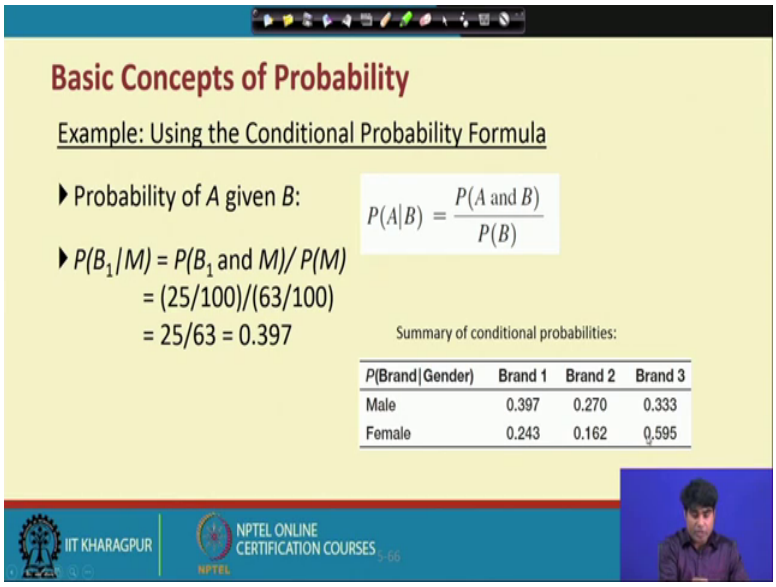
So, then you know in the last lectures in unit 2 we have already discussed the concept called as cross tabulations, now with the help of this ones and you go to the excel spreadsheet then you prepare a cross tabulations. Then cross tabulation finally, this particular structure in the transport in to this particular structure and by default, this is nothing, but actually called as you know probability structures.

So, now, each individual item has a kind of meaningfully interpretations right. So, you know brand ones so for as a female is concerned so what is the probability. So, now, this is the total samples then out of which female, female sample is 37, male sample is a 63, then brand 1 is a 34, then brand 2 23, brand 3 43. So, all details are there and same information, this is the first hand input and through cross tabulation will transfer into this particular structure and this will give you again you know better a kind of structure through which you can actually project the particular problem and then analyze the problem as per the particular need.

So, corresponding to this shift you know structures. So, again moving towards the probability concept so that means, here actually in this case we are interested to know

probability of preparing brand 1 given that a respondent is male. So that means, actually this is the humans game between actually brand 1. So, this is the brand 1 and then the male. So, this is what the brand 1 and male. So, by default the target is here. So obviously, the probability will be 25 by 63 and obviously, this is nothing but you know 25 by 100 and divided by 63 by 100 so obviously, the total probability will be 25 by 63 here.

(Refer Slide Time: 28:25)



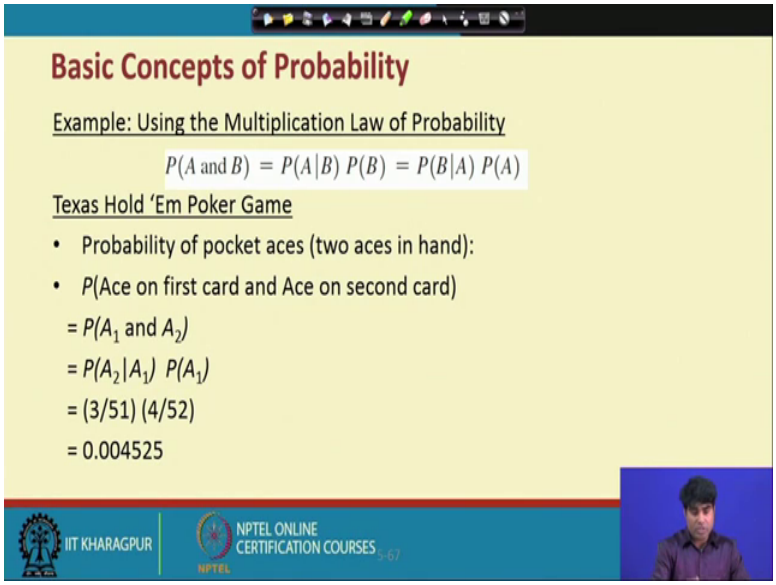
So, now, this is how the structure about the conditional probability that is the mathematical kind of structure obviously, so P A given B, it is nothing but you know P A and B divided by P B. So, that is what actually happening here.

So, what we have already discussed 25 by 100 and 63 by 100. So, you just go through here 25 by 100 and then 63 by 100. So, then you will get you know finally, 25 by 6, 25 by 63 and that is what the probability is all about when you are calculating probability of event A given B. So, that is nothing, but actually 0.397.

So, this is the actually summary sheet corresponding because the original or original information is here original information is here and then this is what actually. So, this is the original information and then this is the cross tabulation, this cross tabulation will finally, help you to predict the particular situations. So, this is the cross tabulations results and on the basis of cross tabulation results you are in a position to predict this

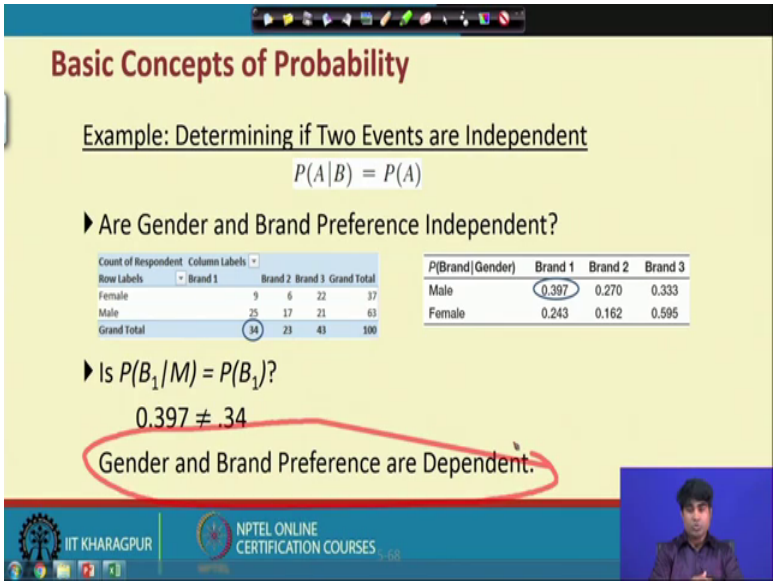
particular situations with the given conditions. This is what is the beauty of you know conditional probability.

(Refer Slide Time: 29:46)



And same again you know it can be huge for complex kind of environment and these are all complex problems for example, you know probability of pocket s 2 ace in hand and, probability of event A 1 and event A 2. So, with the help of probability theorem you can in a position to predict the particular environment with multiple kind of scenario or multiple kind of informations.

(Refer Slide Time: 30:19)

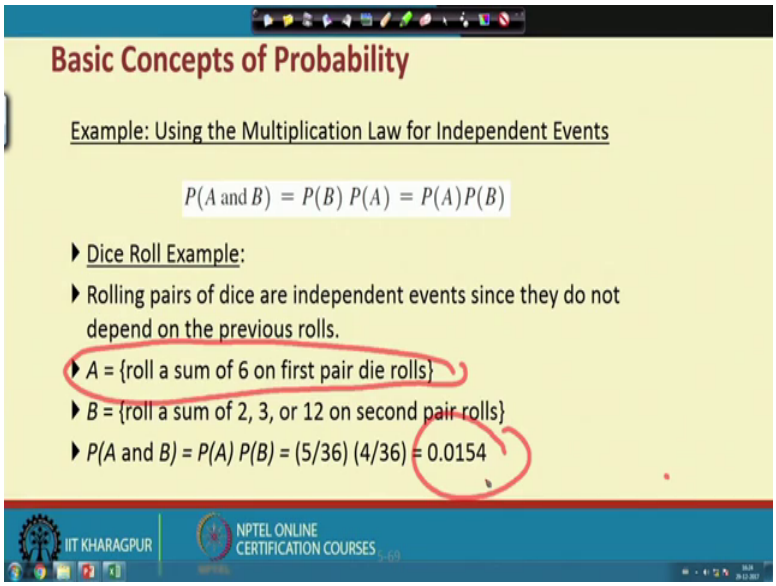


So, this is another kind of example. So, I am putting here this examples to analyze the particular task so obviously, here same problems. So, if 2 events are independent then we have already discussed in the case of you know a mutual exclusive and non mutual exclusive. So, here also same things, 2 events are you know independent then you will be find you know P A intersection B will not be there. So, as a result, probability of A given B will be actually P A. So, that is nothing but here the case.

Now, the second example are gender and brand preferences are independent. So obviously, the gender and brand preference you know are dependent, it cannot be actually independent because it has a well connection with you know branding with the genders. So obviously, means the idea is that you know this is what actually we actually we need actually or not I mean, when we talk about the use of analytics for management decision. So, now, this is a excellent kind of inference which you can actually find through the use of analytics and then you can predict the particular situation. That means, in this case the indication is that gender and brand preference should not be actually, should not actually break the link. So, there is a linkage and the kind of prediction and the kind of marketing strategy should be used for any kind of branding. So, gender has a linkage.

So, you have to take care of the kind of issue when we are you know doing something related to brand preference. So, this is how the conditional probability will help you to calculate the particular situation and predict the environment as per your requirement.

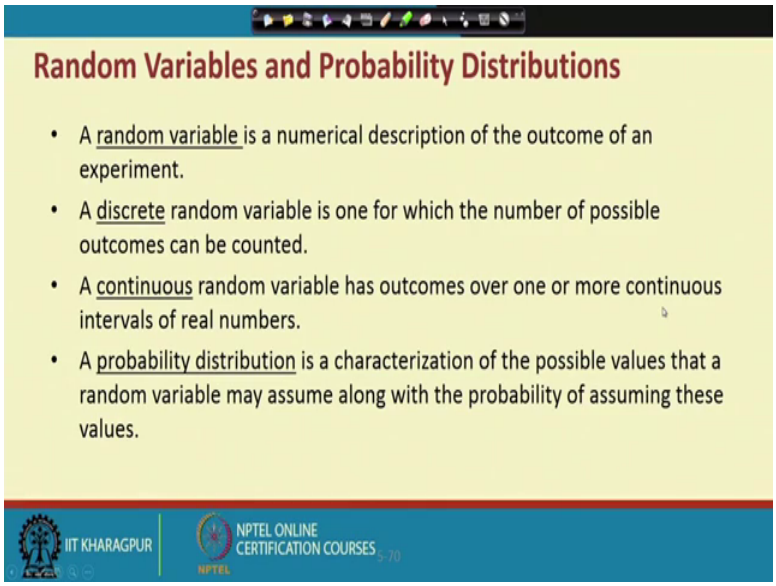
(Refer Slide Time: 32:10)



Again a similar kind of you know probability concepts. So, here some of the rules are there. So, when again, probability 2 event means probability case of probability case when there are 2 independent events and they are completely independent to each other then the probability of A and B can be simply probability of A into probability of B.

So, the cited example here, now, the examples of here you know if you say same example which you have already a highlighted in the previous example dice roll, then roll a sum of 6 on first pair die rolls and roll of 2 3 or 12 in a second pair rolls; obviously, 2 concepts are you know or 2 particular structure are completely independent. So, as a result the final outcome will be P A into P B only. So, as a result 5 by 36 and multiplied by 4 by 36 and this will give you the probability of 0.0154. So, it is nothing, but actually 2 percent only. So, likewise you know you have a multiple kind of situations through which actually you can actually absorb or you can predict the kind of a requirement and then with the help of this probability concept you can solve some of the management problems and then you can take your decisions.

(Refer Slide Time: 33:28)



So that means, actually this is very very useful component all together and it is not actually a for (Refer Time: 33:47), using some kind of game like tossing coin and rolling in die etcetera, but the thing is that you know it is very beautiful component through which actually you have to predict the business environment. And the concept of probability starting with you know gambling activities, but actually in the real world or in the modern time of you know modern kind of environment, it is you know why it is frequently used in most of the activities, so far as you know prediction is concerned and judgment is concerned. So, all together my summation is that you know, this particular concept is a high importance and it is a highly required for any kind of business environments so for as a business prediction is concerned and that to their with the use of you know any hard core analytics tools or without any hard core analytics tools.

I mean sometimes what I like to say that sometimes without using any complex kind of business analytic tools probability can probability itself can give you some kind of inference or some kind of output through which you can take management decision. So, in the probability case, some of the things we which you can discuss in more in details or you can connect with you know other kind of concepts which is called as probability distribution. So, when you are doing something or doing something kind of management activities then it is a kind of continuous process and which we have already discussed in the case of unit 2 and unit 1, slightly because you know every time we wanted to know when you have a samples and that to data for a particular variable or a particular event

the first hand check is actually the kind of visualizations and then we try to extract or try to visualize is how is this particular distribution.

So, whether you know it is equally spreads or it is a skewed distributions or it is a kind of symmetrical distributions. So, the probability structures will give you better kind of indication. So, for as you know understanding of data is concerned and the use of the data is a concern and then you can actually analyze this problem as per the requirement and then you can take some kind of management decision. So, now, how this probability concept can be connected with the various distributions so that you will discuss you know after this break.

Thank you very much. Have a nice time.