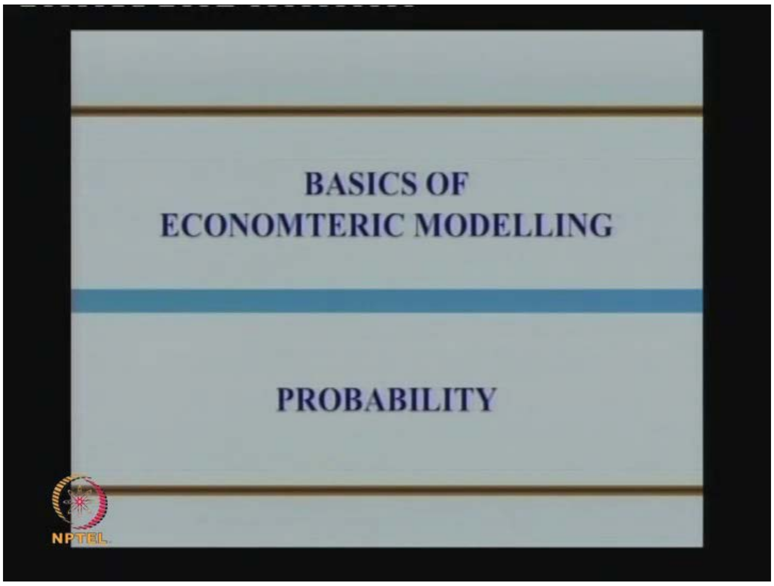
Econometric Modelling
Prof. Rudra P. Pradhan
Department of Management
Indian Institute of Technology, Kharagpur

Module No. # 01
Lecture No. # 06
Probability

Good afternoon. Welcome to NPTEL project on econometric modeling. This is Rudra Pradhan here. Today, we discuss the concept probability.

(Refer Slide Time: 00:30)



In the last lectures, we have discussed various aspects of econometric modelling. That is, the structure of univariate modelling and bivariate modelling. In the univariate modelling we have discussed central tendency, particularly mean, median, mode, dispersion; that is range, total deviation, mean deviation, standard deviation, coefficient variation and skewness; that is the shape of the distribution.

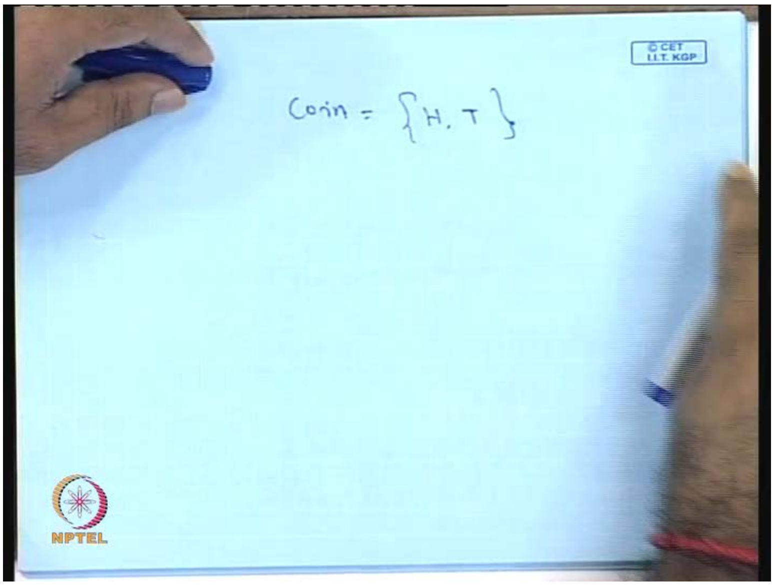
On the other sides, we have discussed correlation and regression; that is, to know the association between the two variables and the causality between the two variable. Now, the issue is whether you go for univariate modelling or bivariate modelling, we must like to know whether the univariate modelling or bivariate modelling has its uniformity or has its feasibility. So for, as a feasibility is concerned or strengthness of this modelling is concerned or validity of the model is concerned or significance of the model is concerned, we have to know or we have to add certain things here. The issue is that, we like to justify whether the particular setup, that is univariate setup or bivariate setup are relevant or significant one. To justify the same, we like to know the concept of probability, estimation and hypothesis testings.

So probability estimation hypothesis testing will has a fantastic role. So for, has a significance of a particular variable in a modelling setup. So whether you apply univariate statistics or bivariate statistics or multivariate statistics, you must have complete information and or complete structure of the models and that model has to be statistically significant. So now, in these particular lectures, we like to know how probability can be applied or can be used to know the significance of a particular variable or significance of the model; that means the overall fitness of the models.

So now, what is all about this probability? Probability is basically, you can say chance of occurrence. So it is quantitative measurement of uncertainty. In the real world situations, you will find some of the things are very certain and some of the things are very uncertain. So that is how, in the beginning, we have mentioned the difference between the social science structure and you can say mathematical science structures.

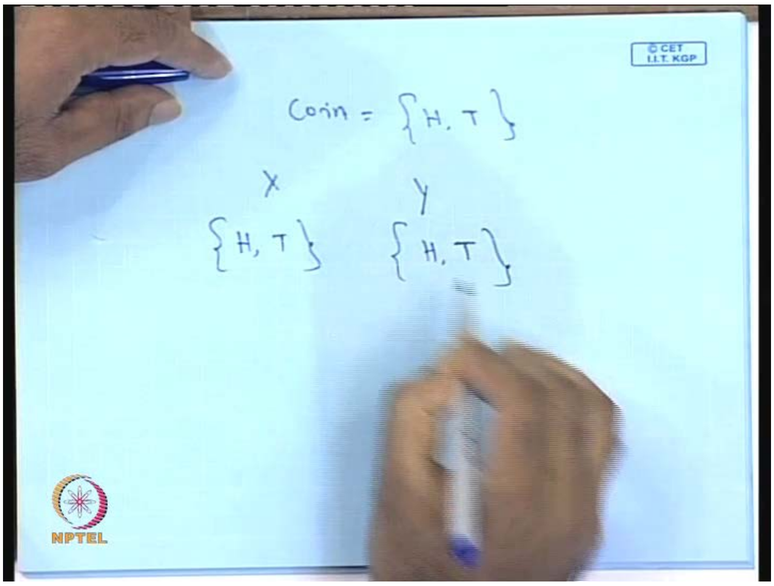
In one sides, the relationships are exact in nature. In the other side, the relationships are inexact in nature. The moment you will say relationships are exact, then obliviously this is one way, we can say that the situations are very certain, very clear. So now, point to go for verification or you can say any statistical to estimate or to check reliability, etcetera. But, in reality there are so many things are there. We are now sure about the fact, so that means there is always question of uncertainty. So now when there is question of uncertainty, we try to justify how it can be possible to measure or you can say how you can justify the situation. So, when it is uncertain, so what is the percentage? So, that means how much different from the certainty. So now, in that case probability plays a fantastic role. So, the basic meaning of probability is the quantity measurement of uncertainty. So, it measures the degree or chance of occurrence of an, you can say uncertainty event.

(Refer Slide time: 05:19)



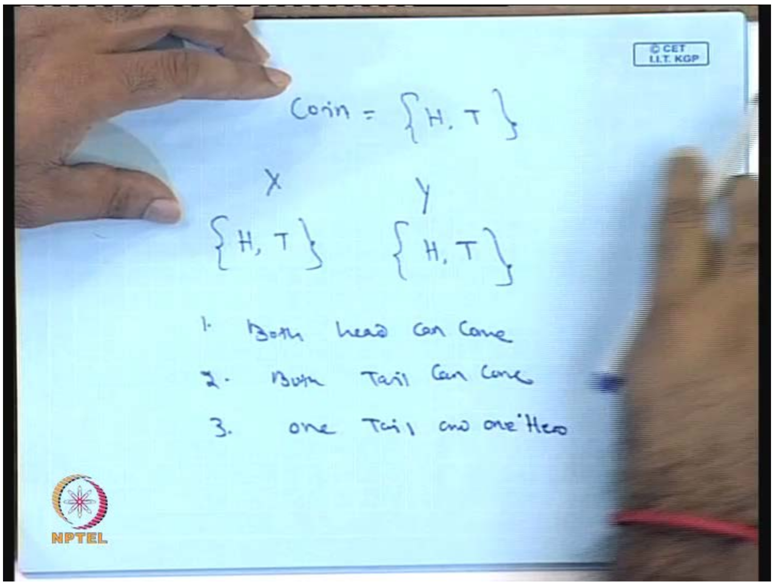
So I will take a case here. Let us take a case of, I will toss a coin, let say coin is a examples. Now, if I will toss a coin, then obviously its outcome is either head or tail. Now, how probability can plays a role here? So now, the moment you will toss a coin and obviously we have two possible outcomes. Either you will get head or you will get tail. Now, the question is whether you will get head or whether you will get tail. So now, this is not at all certain. Now, uncertainty has to be applied. Now, take another case.

(Refer Slide Time: 06:02)



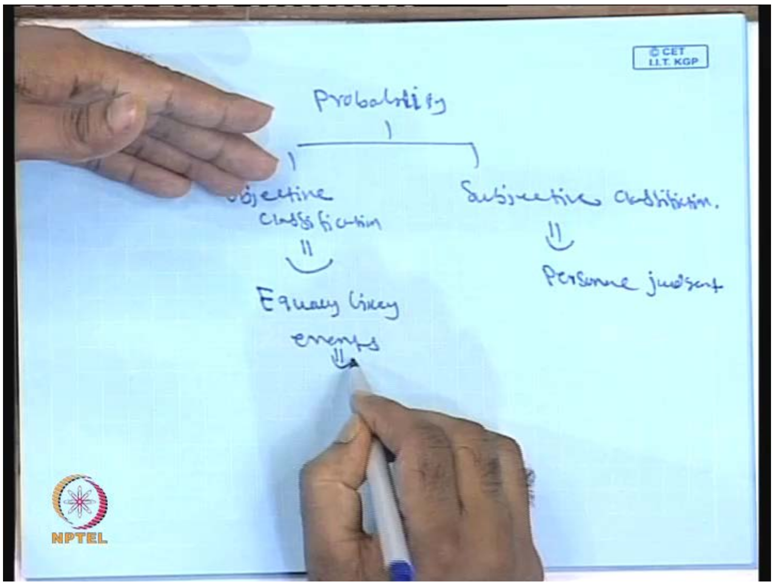
Now, there are two, there are two persons, say X and y. So, they like to toss a coin. Now, in the X case, the possible outcomes are head and tail, and in the Y case, the possible outcomes are also head and tail. So now, if both the person at the same time toss the coin, then obviously what is the chance of occurrence or what is the chance of belief or chance of occurrence? Now, the moment both the person, you can say try to toss the coins, then obliviously there is several possibilities.

(Refer Slide Time: 06:40)



First both both heads can occur, both the cases, both head can come, then second case both tail can come and third case, one tail, one tail and one head. So, there are three possible situations are there. So, we have to make an experiment, what is the possible outcomes and at what extent, we have to bear it? So, in that context, you know probability plays a fantastic roles. Let me explain, what is all about these probability structures.

(Refer Slide Time: 07:33)



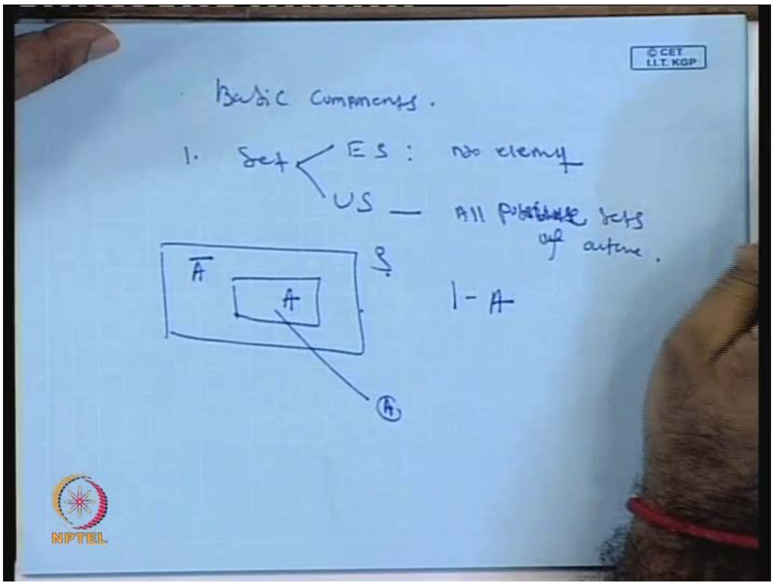
Probability, basically divided into two parts. Probability, we can measure into different angles. One is called as a objective classification, then another is called as a subjective classification. The case of objective classification, it basically depends upon equal likely events, equally likely events. In the second case, subjective classification, we usually look for personal judgment. For instance, take a case of election, snow falls, etcetera. So, in the case of objective classification, equally likely events. The same thing, toss a coin, P k card or toss a dice, etcetera. So, in that context, so subjective classification of probability plays a fantastic role. So, probability basically can be measured in objective angle and can be measured in subjective angles. So, in one case, equally likely events must occur. In another case, it is must on the personal judgments. Equally likely events means the cost of a particular outcome may not depend on the cost of other item. So, this is how the objective classification and subjective classification.

Now, the issue is here, so what is the exact structure of probability and how it is, you can say applied in the econometric modelling? Basically, probability can be applied purely on mathematics and purely on statistics. But, in this econometric modelling, particularly when we will go for multivariate framework? When the problem involves so many variables at a time, then we like to know which particular variable is most impact and if it is or if it has impact, then how much and whether the impact is statistical significant.

So, in order to know all these answers, then probability has to play the game. Now, probability will give an indication, whether it is it is significant or if it is significance, what is the level of significance? So, that is how we have to justify. In that context, probability plays a fantastic role. So, we like to know the detail about the structure of probability, before you apply it, in econometric modelling.

Now, so the concept of probability is very, you know very big. It is a, it is a broader angles we have to describe, but in the mean time, we have to just discuss the concept very briefly. So, our point is here to know, what is exactly probability, how is it setup and how it can be used in the econometric angles. Now, so know the detail concept of probability, so we must know something about the concepts of probability. So, to interpret the probability, you must have a thorough knowledge or its you know components.

(Refer Slide Time: 11:09)

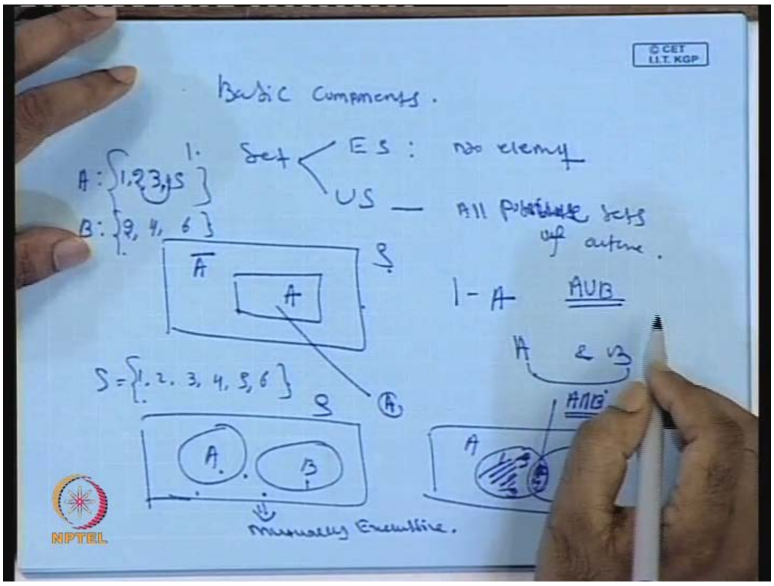


So one of the basic components, that is what we we we will say basic basic components of probability. First is the set. Now, in mathematics there is concept called as a set. Set means well defined objects. Object may be in different way, may be different set, different structure, etcetera. Here, the starting point of this probability is you can say, this set theory. So, this set theory plays a fantastic role in probability. So, we have various concept of set in the probability theory. Now, set we we are describing in two ways: one is called as a empty set and another is called as a universal set. Empty set means, if the set has no elements, no element and universal set means all possible all possible set so all possible all possible sets of outcomes possible sets of outcome.

So now, let us take a case here. This is the set, we will call it universal set S. Now S, inside S, we have a set called A. Now, this is a particular set A and our you can say event A, then S is the complete set. So, this is otherwise called as a universal set. Universal set means it is the integration of all sets at a times. So within the universal set, we have a subset. For instance, just like we have discussed the multivariate, bivariate framework. Within the bivariate framework, framework we have bivariate analysis and under bivariate analysis or again multivariate analysis, we have again univariate analysis.

So now, this structure is almost all same here. So, we are discussing about the universal set. So, within the universal set, we can describe or we can pick, in point a particular set also. Now, so what is all about set, what is empty set, what is the universal set? So, empty set means there is no element in the particular set. Universal set means all possible outcomes in the set. And in addition to that, there is concept called as a A complement. So, it is nothing but the set 1, it is usually depend as a one minus A. so, what about elements are in set A which is not consider, then total set minus, you can say total items involve in A, represent the complement A.

(Refer Slide time: 14:07)



So, now there is another questions. Suppose, we have a two sets A and B. So, how they are related to each other. Just like we have discussed, there are two variables, how they are integrated each other. In the probability case also same, so it can be analyzed under one particular variable or one particular event or it can be analyzed under two particular variable or two particular events. So now if it is, if the question is only one particular event or one variable and how is the setup? If it is, with respect to two variables or more than two variables or more than two events, so then how is the setups?

So, now, in the case of two variables say A and B, so then I can represent this universal set like this. Set A, this is universal set S, so within the set I can give a picture here like this. This is set A and this is set B. I will put another another examples here. Set A and set B. So, this is A, B and this is S set. S represents the total number of or total set of, set of, you can sum up all the individual sets.

So now, in the first case, set A and set B are totally independent. This is what you usually call it, mutually exclusive events, mutually mutually exclusive exclusive events or it is otherwise called as a disjoint set; that means there is no common element in between. So, A is a set under universal set and B is a set under universal set, but there is no integration between A and B. Just like we have discussed in the multivariate framework, there are several variables in the in the setup, initial setup. So, we like to know, what is the dependent cluster and what is the independent cluster?

So, now, the specialty of that modelling is that within the independent clusters, we must, we are very much interested whether there is again any integration among the independent variables. If there is so, then the problem may be inconsistent or it may be very complex in nature. So now, there may be situation within the particular setup all these variables may be independent. If it is independent, then obliviously as far the econometric modelling structure, then this is the right track. But, in in in a certain situation or most of the situations, you will find all these independent variables are not independent. There may be some association or relationship between the two or then in that case, we have to find out the way, how to tackle all these, you can say relationship or if you cannot avoid at least, you can minimize this particular problem. So, this is how we have to represent the set of mutually exclusive events.

So, now, in the other case, there are two sets. But, there is some integration between A and B. Now, if there is two sets and we like to know the integration, so there are three different structures here. now, this particular component is one clusters and A outside of the environment, it is called as a complement and this is B set, outside of the B, it is called as B complement. So, now, within A and B, so we like to know how they are related to each other; that means whether they have some common element or they have not some common element. Now, in that case, there is no common element, that means two sets are completely independent to each other. For instance, so I will take examples here.

For instance, I will take A equal to 1, you can say 3, 5. This is the set of you can say odd numbers. Then, I will take another case B equal to 2, 4 and 6. This is the set of even numbers. Now, this is event A and this is event B. But, both are, if I will say S equal to 1, 2, 3, 4, 5 and 6, this is the example of tossing a or you can say die. So now, there are six possible, you can say outcomes. So, within the six possible outcomes, I have categorically divided into two parts: one is the event which have odd numbers and the event which have even numbers.

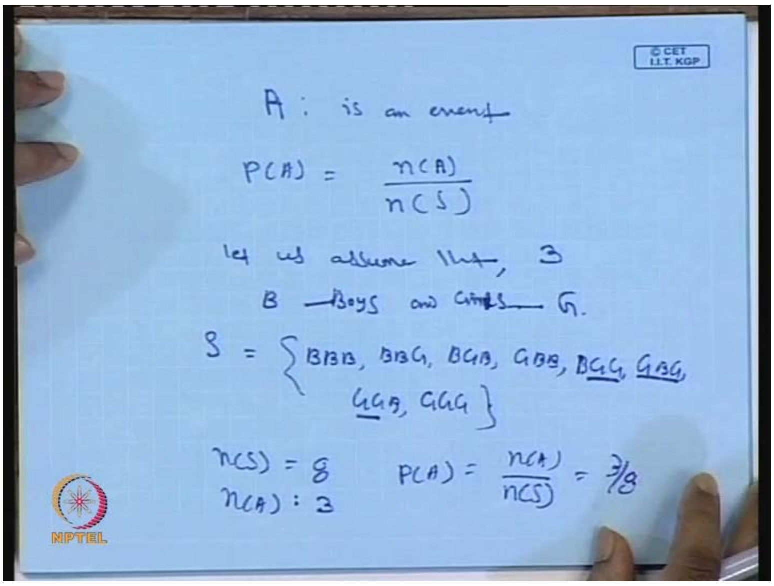
So, now, when I represent the odd structure, we have three elements that is 1, 3, 5 and when we will go for even elements, then we have again three elements, that is 2, 4, 6. But, set A and set B are completely independent. So, that is what, it is called as a disjoint set. So, this is one typical problem in probability theory. Another typical problem is that both the sets are in the universal set, but there is some common element. For instance, I will take here, the set 1, 2, 3, 4, 5. So, then another case, I will take 2, 4, 6. And obviously 2 and 4 has a common element. Now, I will take it here. This is 1, this is 3, this is 2, this is 4, this is 5. Now, if B has a even number elements, then 2, 4, 6; that means 2, 4 is a common element. So, these 2, 4, 2 and 4 is represented as a A intersection B. So, this, the mathematical notation is called as A intersection B.

So, now, if there is the intersection or this common elements between A and B, so that means, if we will clog together and it must be the universal set, provided the two sets are, A and B are only only two sets in the systems. So, then obliviously, there is problem called as a A union B. So, a union represents the set of all elements between set A and set B. so, this is how, you like to know, what is a particular set, component set A, what is the particular set, component set B, then we like to know what is the association between A and B. So, there are two way, they integrate. Either, they are completely independent or they are dependent. If they are dependent, how they are dependent to each other. So, this is how the complete structure of, you can say probability theory.

So, now, there is the, there is certain other items. For instance, one one typical word is called as experiments. So, since the term probability is the quantity measurement, then we have to, we have to make so many experiments to get you know possible outcome because the situation is totally uncertain. So, we have to make so many experiment to get a particular objective or to fulfill the particular objective.

So, now, experiment plays a fantastic role in the probability concept. So, similarly, there is concept called as a sample space. Sample space consist of, you can say all observation at a time. So, this is otherwise called as also universal set. Now, there is concept called as a event. Event means in a particular case, just like we have discussed set A and set B. Set A may be a particular event, set B may be a particular event. So, now, within a particular system, event A, event B may be may be integrated, may not be integrated. If they are integrated and how they are integrated to each other? Now, so, how do we measure the concept of probability?

(Refer Slide Time: 21:41)



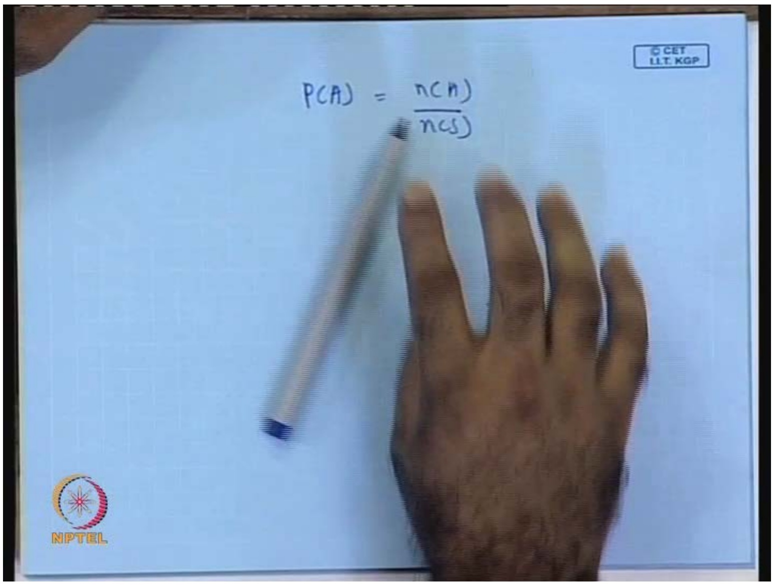
Generally, probability concept for a particular event say, let us assume that A is an event, A is an event. So, then probability of A is measured as a number of favorable outcomes divide by number of possible outcomes; that means total number of possible outcomes in a system and number of particular event. So, the ratio is called as a probability. So take a example here. Let us assume that, let us assume that there is a family consists of three, three members. So, there is a family family consists of three members and the family

family cluster is with respect to boys and girls, boys and girls, girls. So, we like to know, what is the probability of having exactly two girls in the system and not more than two girls in the systems?

So, now, if there is question of choosing three members in a particular team, then how they are, you can say integrated? And how the problems are formulated here? Now, the total space will be like this way. Now, I will call it instead of boys, I will call it B words and instead of girls I will call it G words. So, three particular situations can be occured simultaneously. So, there may be different situation B B B; so, that means there is no girls settles, then B B G, then B G B, then G B B, so then then B G G, then G B G, then G G B, then G G G.

So, now, this is the complete setup. Now, the question is what is the probability of exactly two girls at a time? So, exactly two girls means, we have to see where only two girls are there. So, this is one case. This is another case. This is another case. So, there is no more. So, there are three outcomes, three numbers. So, total samples space on n S equal to 1, 2, 3, 4, 5, 6, 7, 8 and individual sample set n A which is exactly exactly two girls, which is nothing but, you can say 3. So, probability of A is nothing but, n A by n S; which is nothing but, 3 by 8. So, this is how probability can be calculated. So, now, we like to know various issues of the particular term probability. So, let me explain here what is that issues.

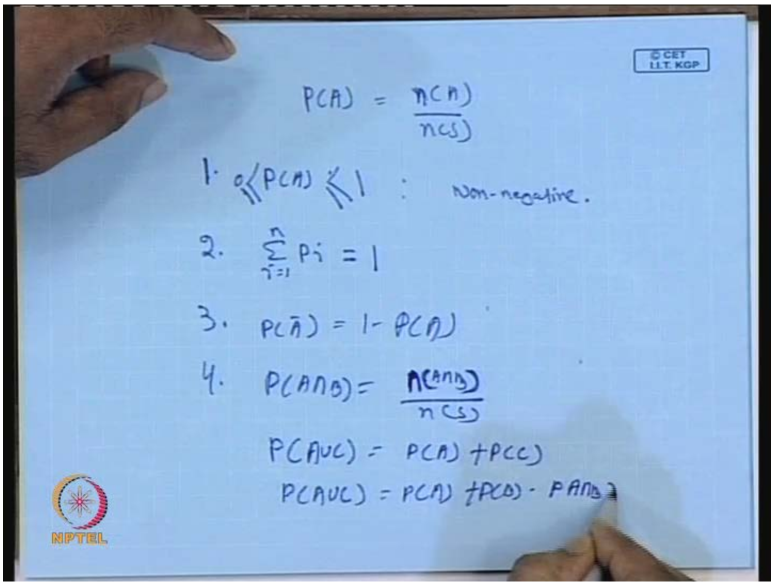
(Refer Slide Time: 24:40)



So, the basic starting point of probability is like this, for a particular event A; P A equal to n A by, you can say n S, but you know like last class discussion every statistic has its you know advantage and had its disadvantage. You know disadvantage means it has lots of limitation and shortcomings. So, it has also lots of advantage, so that we can apply or we can solve a particular problem. For instance, take a case of covariance. The specialty of covariance is that it tresses association between two variables. However, the limitation part of this particular statistic is that, if two variables or if there is any comparative analysis, then the technique covariance cannot be used properly because it is not at all unit less measurement.

So, now, in order to solve that particular problem, we have to apply correlation. So that, with the help of correlation, we can get to know the answers. But again correlation has an advantage over covariance but certain situation, correlation itself has a lots of limitations. For instance, correlation is unit less measurement and it is advancement covariance, it is better technique to know the association between the two variables but in the same time, it cannot measure the cause and effective relationship between the two variables. So, for that again, we have to go look for something else. So, this is how the problem, very complex in nature. In every in every statistic, has its advantage and has its disadvantage. Now, in the probability aspects, we like to know what is specialty of particular probability.

(Refer Slide Time: 26:34)

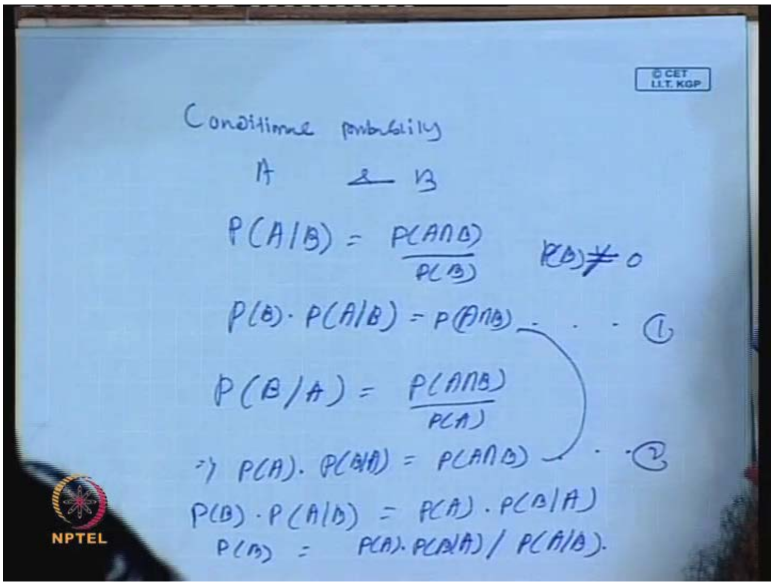


So, like you know, we have discussed the correlation correlation issue. One of the fantastic feature of correlation is that, the value correlation coefficient lie between minus one to plus one. So, similarly, in the case of probability, we have also certain advantage or you can say interesting features. So, one of the interesting feature is that probability of A is always greater than to 0 and less than to 1. So, that means the value of probability is always non-negative. So, it is always non-negative; that means it is always positive. This is the most important properties of probability.

Second, sum of P i, i equal to 1 to n is exactly equal to 1; that means sum of all probability in a particular setup must be exactly equal to 1. For instance, in a particular setup there are two events A and B. So, obliviously the probability of A occurrence and probability of B occurrence should be exactly equal to 1. First of all, the probability for first event A should be positive and probability for second event B should be positive and in a final case, total probability must be exactly equal to 1. So, these two property has to be fulfilled, otherwise the concept of probability is inconsistent.

So, there are other other tricks also, like you know P A, probability of A complement which is nothing but, 1 minus P A. Then, we have also discussed A integration, such you can say P A intersection B is equal to, you can say probability of A intersection B divide by, you can say n total sum; in fact, its number of possible outcomes divide by total number of outcomes. And for mutual exclusive case, for mutual exclusive case P under A union C is equal to P A plus P C, where P A union C usually P A plus P B minus P A intersection B; Sorry, A intersection C. So, this is when there is mutually there is mutual exclusive or disjoint, then obliviously the common element or intersection term will equal to 0. So, simply P union C equal to P A plus P C.

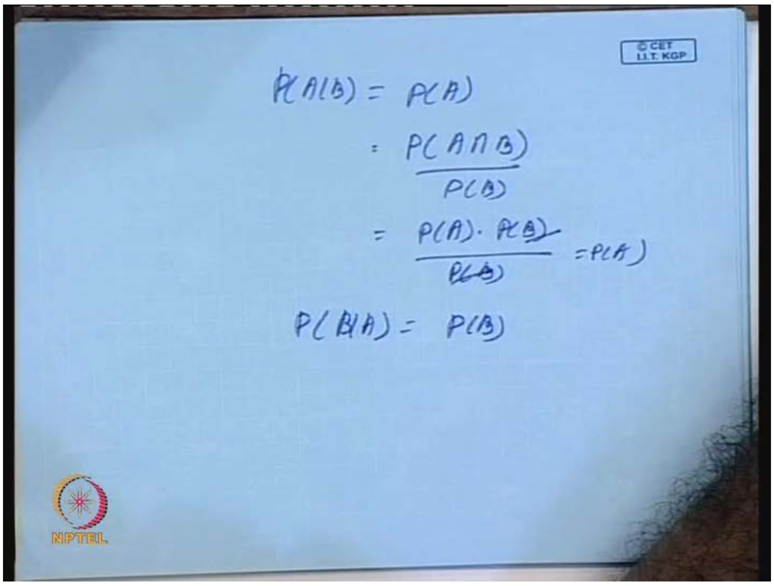
(Refer Slide Time: 29:13)



So, this is how the structure of statistics of probability or we can say, we can discuss the feature of probability. Then, there is concept called as a conditional probability. There is concept called as a conditional probability. So, we take a event. There are two events A and B, then A and B. Then, there are two way, we have to define the conditional probability P A given B which is nothing but P A intersection B divide by P B provided P B must be positive. Positive means that that is not be equal to 0. Then, P obliviously P B into P A given B is equal to P A intersection B. So, this is case one. And another situation, probability of P given B given A is equal to probability of A intersection B divide by probability of A. That implies probability of A into probability of B given A is equal to probability of A intersection B. Now, this is equation number two.

So, now, if we will compare these and these, so, then P probability of B into probability of A given B equal to probability of A into probability of B given A; so that means probability of B is equal to probability of A into probability of B given A all divide by probability of A given B. So, this is what the theorem of conditional probability provided in all the cases the value of probability should be positive and it should be in between 0 to 1. So, it means no situation, it should be negative and it should be more than 1. It should be, the limit should be in between 0 to 1, like the term called as a coefficient determinant which we discussed in the last class, the square of the correlation coefficient.

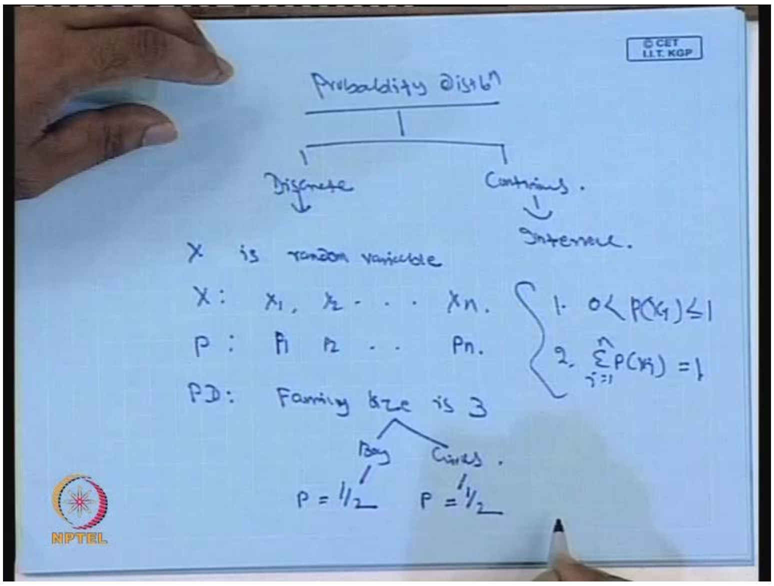
(Refer Slide time: 31:34)



So now now, for a for independent events, for independent events, if the two events are independent or disjoint, then probability of A given B is simply equal to probability of A because it is nothing but probability of A intersection B divided by probability of B, but probability of A intersection B is nothing but, probability of A into probability of B because they are independent. So, probability of B, so it is simply equal to probability of A. Similarly, probability of B given A is equal to probability of simply B. Similarly, P A into P B divide by P A. So, obliviously the simple answer is P B. Now, we have, we like to know what is what is the issue of probability, various aspects of probability, various theorems under probability and various conditions of probability. So and within the condition, we will like to know the conditional problem. This is very interesting component which we discuss in details later stage.

So, in fact, there are so many other issues also in probability, is what we called as a mathematical probability. So, we are not going to discuss the details about this issue because we are very much restricted on econometric modelling. So, our our idea or our agenda is to know little bit about the probabilities. So, that it is a a means, it will it will help you lot for for econometric modelling, particularly to test the reliability issue and to check the significance of a particular variables.

(Refer Slide Time: 33:23)



So, now, so I will discuss here the concept called as a probability distributions. So, we we have the issue here, probability probability distributions. So, before before we, before we move to probability distribution, so we must have a thorough knowledge on probability, then we have to discuss about its distributions like we have discussed the univariate modelling. So, that means we like to know what is the central structure, variability structure and then, we have to come for size of the distribution or step of the distribution. So, in the case of probability also same things. So, we must have the theoretical background of the probability and its structure conditions theorem, then we have to see how is it exactly distributions.

So, now, suppose as a probability distribution is concerned, there are two ways it can be discussed. One under discrete series and another under continuous series. So, let us what is let us we discuss first, what is probability distribution? Continuous means the variables in intervals, it will be obtained in a intervals. So, here it may be finite or infinite, but it is not in interval structures. Now, so, to to analyze the probability distribution, we assume that X is a random variables, X is a random variables. So, that means its outcome completely depend completely depends on chance. So, it it is a it is obtained through experiment only. So, this is how it is called as a a random variable. X is a random variable, X consists of X 1, X 2 up to X n and corresponding probability is equal to P 1 P 2 up to P n. For instance, what is all about the distributions?

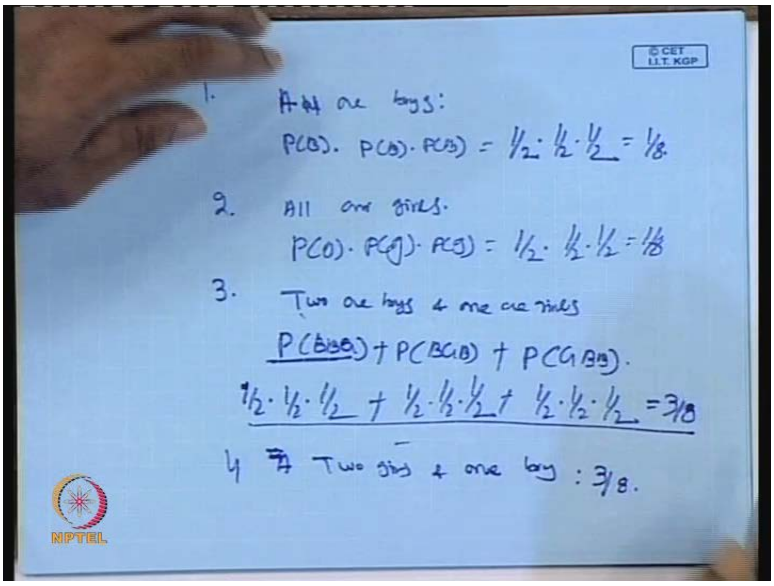
Let us take a case; you like to toss a coin. So, obviously, there are two possible outcomes, either head or tail. So, how do we analyze the situation? So, of course, the total outcomes will be two; head and tail. Now, when we will make an experiment, then the situation is totally uncertain. So, now, we like to know what are the possible outcomes in that particular structure or systems. Now, first structure is means, there may be 0 head, then there may be 0 tails. Now, it is possible when there are two outcomes at a times, but when there is question of only one, then the game is very simple one. So, in that case, so either there is possibility of head or there is possibility of tail. But, now we have to see, how much tails are there or how much heads are there.

So, now, similarly, X is a random variables which represents the number of occurrence. P represents the corresponding probability. So, obviously the condition is that here, so in the first case, 0 less than P i P X i less than equal to 1 and second case summation P X i i equal to 1 to n exactly equal to 1.

So, let us take a case here. We like to formulate a probability distributions, we like to formulate a probability distribution. The condition is that so, the same problem. Let us take a case of a family size, family size is three and it consist of boys and girls.

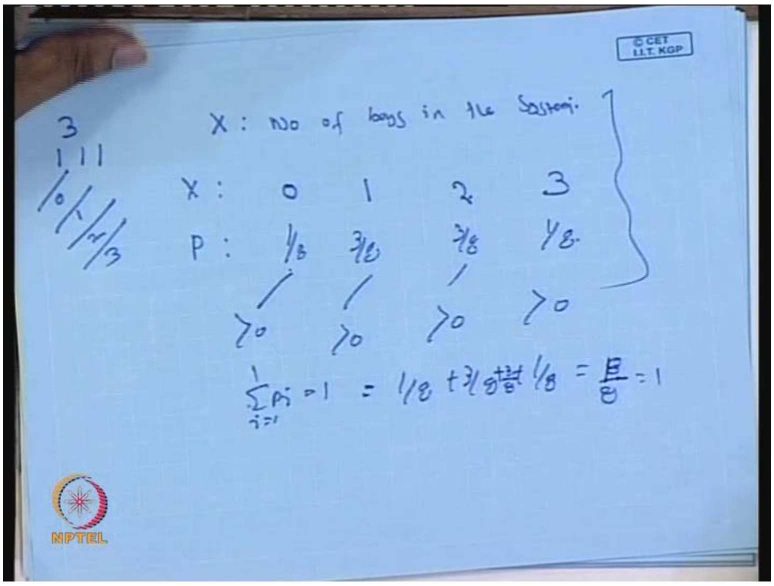
So, so how is the step of the probability distribution boy chance of occurrence is 1 by 2 and girls chance of occurrence is 1 by 2. So, let us assume that the probability of success P equal to 1 by 2 for boys and probability success for girls is P by 1 by 2, just like head and tail. So, the possible chance is 1 by 2 and 1 by 2. So, if it is head, of course, then obviously, it is 1 by 2 and if it is tail, of course, it is also 1 by 2. So, now, in this particular setup, also see the situation is almost all same. Here, there is two alternative only, either boy or girl. Now, how we have to formulate the situation.

(Refer Slide Time: 38:23)



Now, there are certain case. First case, if all are boys; if all in all the cases all are boys. So, then there are three possible outcome. So, I mean family size is three. So, P B into P B into P B. So, obviously the total outcomes will 1 by 2 into 1 by 2 into 1 by 2. So, this is 1 by 8. Now, let us another case, all are girls, all are girls. Then, probability probability of G into probability of G into probability of G, so which is equal to again 1 by 2, 1 by 2 1 by 2 equal to 1 by 8. So, now, similarly, so we like to know, we like to know what is the case of two boys and three girls case. In that case, so let us say two are boys and one are girls. So, then P into B B G B B G plus P B G B plus P G B B, in all the cases here this is P B B G means 1 by 2 into 1 by 2 into 1 by 2 plus 1 by 2 into 1 by 2 into 1 by 2 plus 1 by 2 into 1 by 2 into 1 by 2 because in the first case B, 1 by 2. Then, again B 1 by 2, G also 1 by 2. So, in all total, if we solve the particular problem, then you will get 3 by 8. So, now, fourth case, if two are girls and one boy, it is two boys, one girl. Then, next case, two girls and one boy, so result is also similar. So, that means it will come also 3 by 8. Now, we will formulate a probability distributions.

(Refer Slide Time: 40:41)

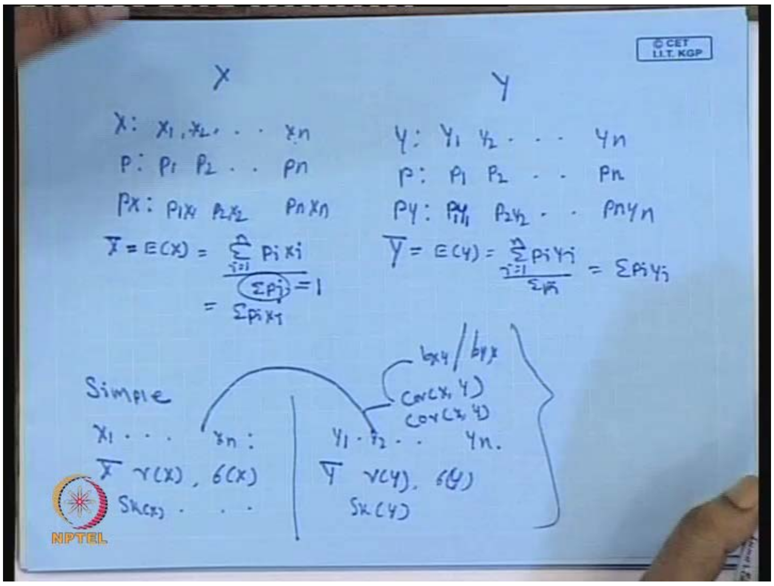


So, now, probability distributions X is a random variable which consists of, you can say let say number of boys in the systems, number of boys in the systems. Now, so, X contains, it can be it can be 0, it can be 2, it can be 2. Sorry, it can be 1, it can be 2, it can be 3. So, that means the family size consists of three. So, if there are three, then there may be possibility 0 number of boys, one number of boys, two number of boys and three number of boys. So, there are 4 possible situations. So, now, if there are three possible situations what is the chance of probability? Now, corresponding probability is for 0, it is already 1 by 8. This is 3 by 8, this is 3 by 8, this is 1 by 8. Now, the complete setup is called as a probability distribution provided it is, it must be satisfy the condition. So, what is that condition? Now, in every case, it is greater than 0, greater than 0, greater than 0, greater than 0 and sum of P i should be exactly equal to 1. So, that means 1 by 8 plus 3 by 8 plus 1 by 8 should be exactly equal to 1. So, it can it is a 3 by 8 plus 3 by 8 so it is nothing but, 8 by 8 it is exactly equal to 1.

So, now, so there are many ways you can explain or many examples you can sight to explain this probability issue and also the probability distributions. Now, the interesting aspect of probability distribution is that, it can be integrated with the bivariate framework. For instance, we have already discussed, we already know what is the concept of univariate and what is the concept of bivariate. Now, if we will make a look in probability how probability can be applied or can be integrated to univariate structure and can be integrated to bivariate structure. The way we have discussed till now, it is more or less univariate structure of probability distributions. So, that means every time we are discussing a single variable or single event, then we are finding out the various possible outcomes or various possible scenarios.

So, now, if there are two different variables, for instance take two random variable say X and Y. And they have corresponding probability and how X and Y are integrated to each others, just like you know the correlation structure and covariance structure. We can also a justify here, how it is the problem setup.

(Refer Slide Time: 43:53)



So, now, take a case of two variables. X is one variables and Y is another variables. X contains X 1, X 2 up to X n corresponding probability is P 1, P 2 up to P n. So, in the case of Y, the corresponding situation is Y 1, Y 2 up to Y n corresponding probability is P 1, P 2 and P n. So, now, you make a look here. When we look for simple structure, so we have series X 1 up to X n, then another series Y 1, Y 2 up to Y n. So, what is our agenda here? We like to know, what is X bar? What is variance of X? What is standard deviations of X? What is this skewness of X and so on.

Similarly, this is Y bar, variance of Y, this is sigma upon Y and skewness of Y. This is the entirely or complete structure of univariate setup. Now, when we will move to bivariate structure, then we will like to know how they are integrated to each other. And for that, we either apply covariance upon X Y or we can apply correlation upon X Y or you can also have it through B X Y or B Y X. So, this particular structure is not at all attached with, you can say the concept of probability. Now, the same setup can be integrated or can be interpreted through probability issue because lots of instances or you can say many occasions, your variables are random in nature that means the the outcome of a particular variable entirely depend upon chance, in that case probability has to be applied.

So, now, we like to know, if it is applied in that particular case, then how is the setup? So, we like to know how the probability a structure can be applied to this univariate setup to bivariate setup. Now, when the outcomes are like this, only X equal to X 1, X 2, X n, no probability. Then, the situation is completely certain. So, that means there is no question of uncertainty. So, that means the chance of occurrence is not at all issue. So, now, whatever your expecting this same results with you. So, this is what the certain issue. But, when there is question of uncertainty, then probability has to be applied. So, that means, now the issue is we have to multiply probability probability with the original variables. So, corresponding term is P 1 X 1 P 2 X 2, then P n X n.

So, like here, samples sample means, so we will also call X bar which is nothing but, we will call expected value of X because since it is quantitative measurement of uncertainty, then obviously, we are expecting something. Now, expectation has to be applied. Now, so, your expecting something, so what is that expectation? Is it hundred percent or is it less than that? If it is hundred percent, then obviously the probability chance of probability is exactly equal 1. So, if it is you can say no, then its probability almost all impossible, its 0.

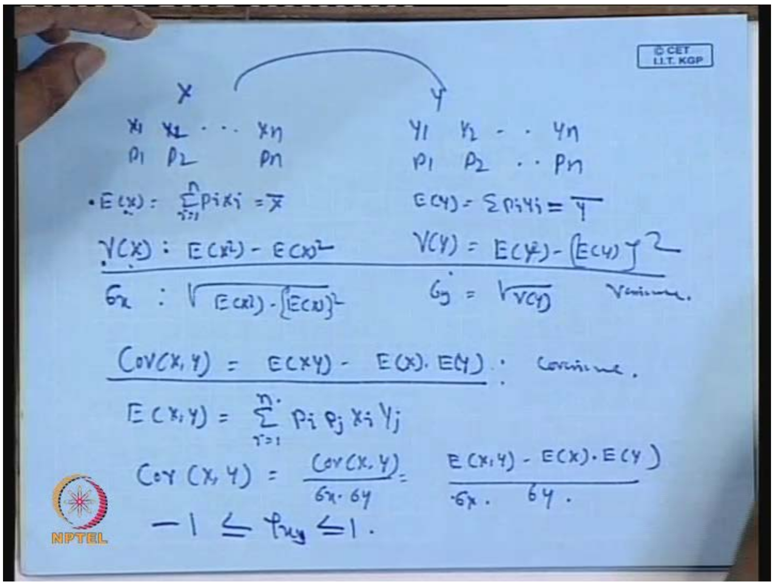
So, now, X bar is nothing but X, so which is nothing but summation P i X i i equal to 1 to n divide by summation P i. You remember one thing, in the in the first lecture under univariate modelling, we have discussed the univariate statistic particularly with respect to mean, median, modes. And under mean, we have different setup like arithmetic mean, harmonic mean and weighted average, weighted average mean. So, now that weighted average is nothing but the probability issue. This probability is just like a weight issues, so it is assigned weight way. But, every items now with this with this setup, if you understand the concept of weighted average mean, then obviously there is no confusion about this expectation in the probability theory.

So, now, in the case of weighted arithmetic means, so we applied summation W i. Now, we have summation P i. Now, this is a X, our expected value of X or mean of X under the probability distribution. Similarly, so here we will we will get P Y. So, this is this is actually, Y structure and this is also P structure. So, we will get P Y. So, P Y is nothing but P 1 Y 1 P 2 Y 2 up to P n Y n.

So, corresponding corresponding to Y and its probability, so we can get expected value of Y. Expected value of Y is nothing but summation P i Y i, i equal to 1 to n. And remember, this is nothing but simply summation P i X i because sum of P i is exactly equal to 1. So, obviously it is also summation P i, so which is nothing but summation P i Y i. Now, this complete setup is called as, we can say mean of A and mean of B.

So, now, we like to know we like to know the moment from univariate structure to bivariate structures, with simple setup, where the situations are very certain in nature. In another case, the situation are completely uncertain in nature, where the structure is not at all simple. The structure is assigned or designed with weight vectors or probability vector, chance of occurrence of that particular items.

(Refer Slide Time: 50:28)



So, now, moving to that particular issue. Now, for X and Y, so X 1 Y 1. Sorry, X 1, X 2 up to X n. P 1, P 2, P n. So, then Y 1, Y 2 up to Y n. P 1, P 2 and P n. So, E X is equal to summation P i X i, i equal to 1 to n. So, this is nothing but, X bar. Similarly, E upon Y is equal to summation P i Y i which is nothing but, you can say Y bar. So, this is E X.

Now, we can get it also V X. V X represents variance of X. So, now, variance of X is nothing but E X squares minus E X whole squares. Similarly, variance of Y is equal to E Y squares minus E upon Y whole squares. So, you have E X and V X. So, that means it is the univariate structure with respect to central issue and this is with respect to dispersion issue. Now, if we will a we will know standardization, then it is nothing but square root of E X square minus E X whole squares. So, similarly, sigma Y is nothing but, variance of Y factors.

So, now, so you like to know E X V X E Y V Y. Now, we like to correlate X upon Y. So, for that we need to have covariance covariance of X Y. So, covariance X Y is nothing but E upon X Y minus E X into E Y. So, now, what is the E X Y? E X Y is simply represented as summation P i P j X i Y j, i equal to 1 to n. So, with this issue, so this is the covariance issue, this is variance issue and we need to have a final outcomes; that is correlation between X and Y. So, correlation of X Y is nothing but covariance of X Y divide by sigma X into sigma Y. So, that means, it is nothing but E X Y minus E X into E Y divide by, you can say sigma X into sigma Y. So, like correlation, so this also satisfied. This is usually denoted as 0 X Y is equal to 1.

Now, so we we get to know what is the issue of probability, its various structure various setup and how it is useful for a real world business problem and how it is helpful when the situations are not certain at all; that means, it is the question of uncertainty. So, that means probability has to be applied when the situations are very much uncertain in nature. So, that is why, it plays a very fantastic role in econometric modelling. So, like you know, original samples, we discuss about univariate statistic and bivariate statistic; that is, with respect to covariance and correlations, so we have to apply probability with the particular samples, then we can also get to know the, what is the univariate setup and bivariate setup.

So, this is you can say, very much helpful or you can say, it is very useful for further econometric modelling. Now, in additional to probability the estimation and hypothesis testing plays a fantastic role which we will discuss in the next class in details. So, with this, we can conclude this class, today. Thank you very much. Have a nice day.