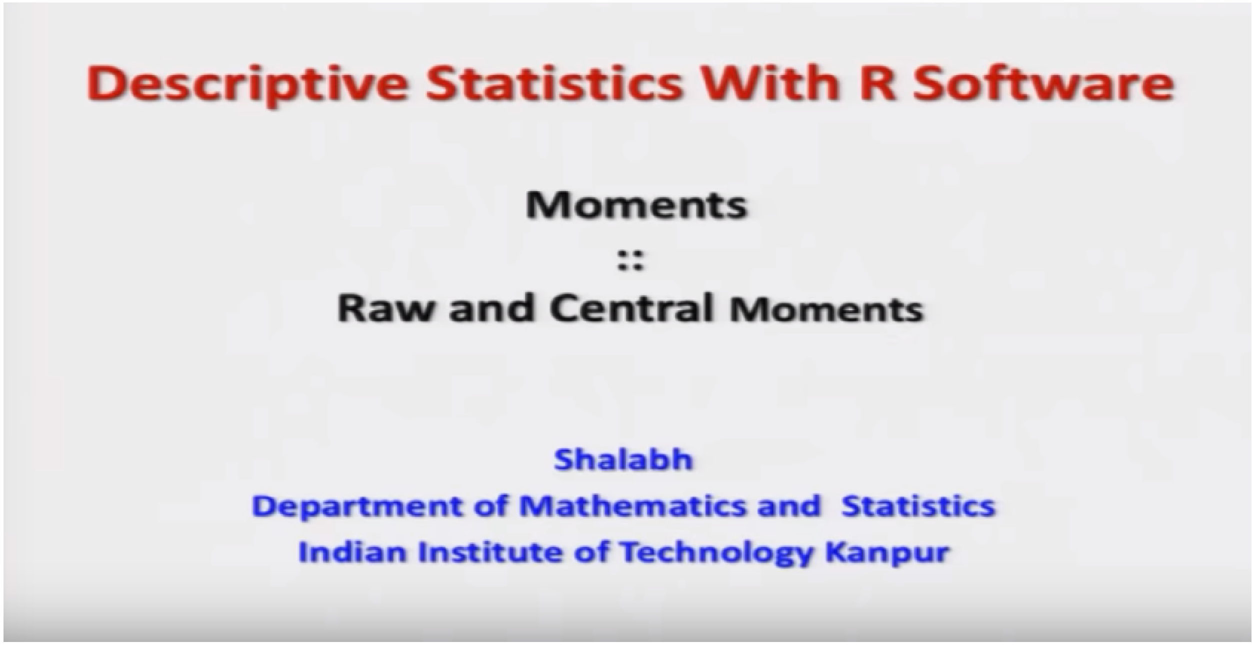
Lecture - 22
Moments - Raw and Central Moments

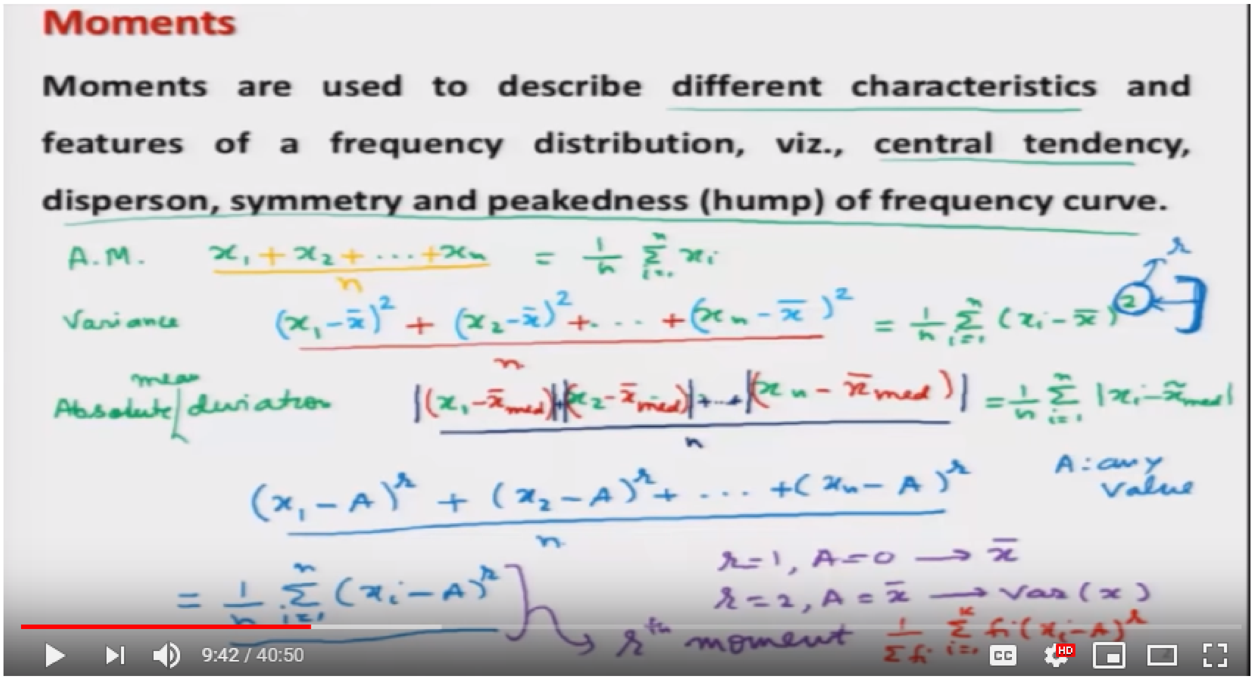
Welcome to the lecture on the course descriptive statistics with R software.

Refer Slide Time: (00:18)



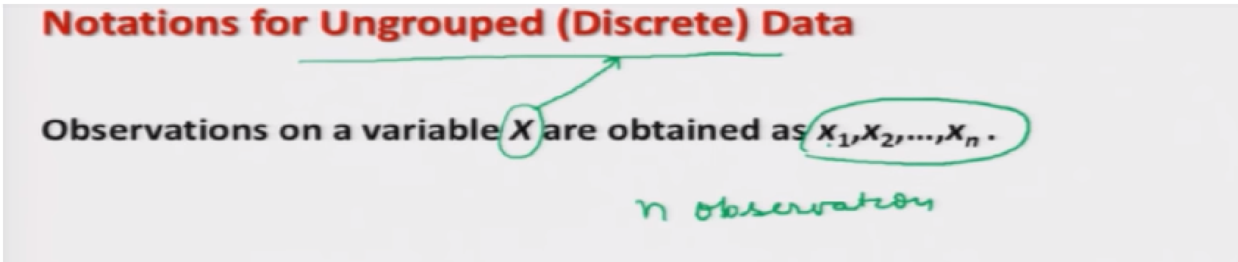
From this lecture, we are going to start a new topic, and this is about moments. What are this moments, and why do we need it, you may recall that, up to now in the earlier lectures, we have considered two aspects of the data information, first is central tendency of the data, and second is the variation of the data, and we have developed different types of tools like as arithmetic mean, standard deviation, variance and so so on, to quantify that information. Similar to this, there are some other aspects like as symmetry of the frequency curve, or how the values are concentrated in the entire frequency distribution, similarly, there is another aspect, what is the hump of the frequency curve? So, in-order to study all these things, and some other important properties of a statistical tool, we need a concept, which is called concept of moments. So, essentially you will see in this lecture that, moments are some special type of mathematical functions, and using these moments, we can quantify different types of information, which is contained inside the data or the frequency table. So, let us first try to understand what are these moments?

Refer Slide Time: (2:17)



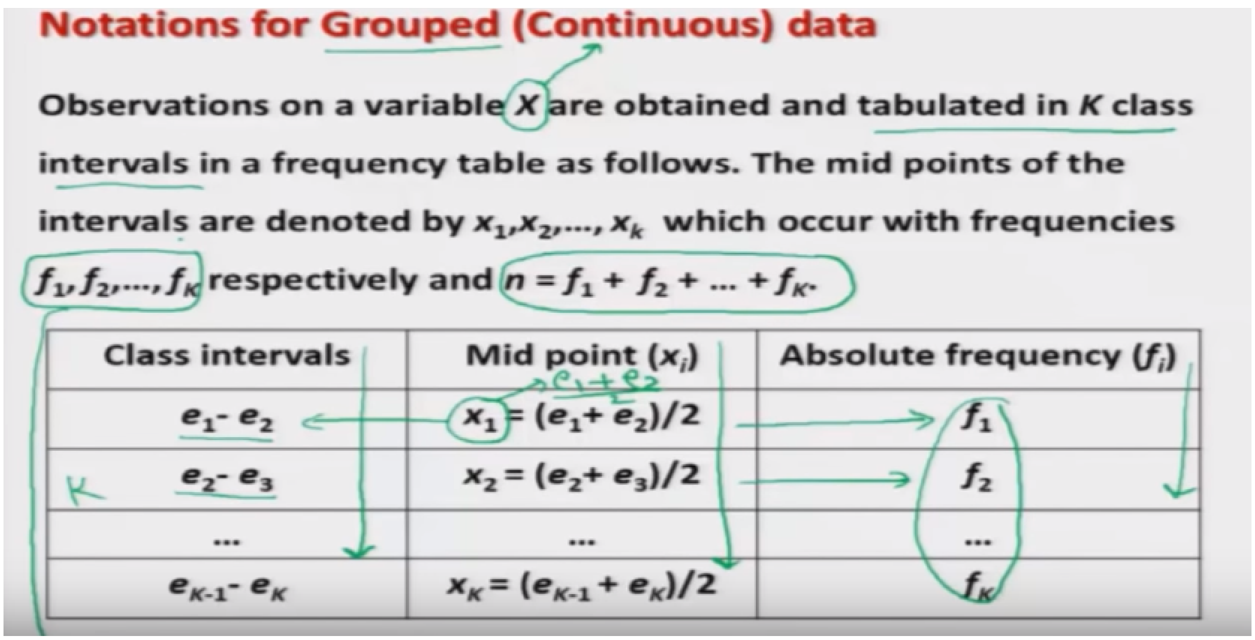
So, moments are essentially used to describe different types of characteristics of a frequency distribution, different features of a frequency distribution like as central tendency, dispersion, symmetry peaked-ness etc., Now, if you try to understand that, how these things have been developed? So, if I try to see here, when I wanted to study the central tendency, then we defined arithmetic mean, and what we did, we had observations x_1, x_2 up to here x_n, then what we did, we just added them, and we divided it by the number of observations, say here n. So, this gave us the information about the central tendency of the data that, we're the mean or the average value of the data set lies in the frequency distribution or a frequency curve, Right! Similarly, when we studied the variation in the data, we define a quantity like variance or absolute deviation, I will try to show you both, what we did, in case of variance, we had the data x_1, x_2 up to here x_n, and then what we did, we took the difference of these observation from the arithmetic mean, which we call as deviations. So, we took the deviations of each observation around the arithmetic mean. So, we obtained x_1 minus x bar, x_2 minus x bar, up to here x_n minus x bar, and after this, we squared them, all the deviations were is squared, and after this what we did, we simply took the arithmetic mean of these squared deviations, and similarly when we computed the absolute deviations, then we had considered the observations, x_1, see here x_2, up to here x_n, and then when we wanted to study the absolute, say mean deviation, then what we did, we simply found the median of these observations, and we took the deviation from their median like this, x_2 minus x bar median, and x_n minus X bar median, and after this what we did, we simply took the absolute value of these things, and after this we just took the arithmetic mean of these values, now you can see here, what are we trying to do, in the case of arithmetic mean, we are trying to find out here, 1 upon n, summation i goes from 1 to n, x_i, in the case of variance we are trying to find out, 1 upon n, summation i goes from 1 to n, x_i minus x bar whole square, and in the case of absolute deviation, we are trying to find out the arithmetic mean of the absolute deviations of x_i, from the median. So, now looking at these three examples, you can see here, what are we trying to do, we are essentially trying to consider the deviations, say x_1 minus, some arbitrary value here A, the deviation of second observation from some arbitrary value here A, and deviation of the another observation from some arbitrary value A, and then I am trying to use a suitable power function on this deviation. For example, if you try to look over here, please try to observe my pen, here you are trying to use the, the square. So, I can make it more general, instead of using here the square, I can replace it by here some general quantity, see here r. So, I try to consider, now here the deviation of observation x_1 minus A, x_2 minus A, x_n minus A, that is the deviations of observation around any arbitrary value A, A is say any value, any arbitrary value which we have to choose. So, now I try to take the r power of these expressions, and then I try to find out their average, this I can express as 1 upon n, summation i goes from 1 to n, x_i minus A raised to the power r, and now you can see that this function has several advantages, for example, if you try to take here r equal to 1, and A is equal to 0, you get here nothing, but mathematic mean x bar, if you take here r equal to 2, and A equal to say here, x bar, then you get here ,variance of x. Similarly, if you try to choose other values, we can define some other functions which are going to represent the different characteristics of the frequency distribution. So, in general I can call this function as, say r^{th} moment, and now you know that this is what I have defined for the data, which is observed as, x_1, x_2, x_n, and that is ungrouped data, and if you want to define the same quantity for, say here, some grouped data,, then this can be simply expressed into summation of, of f_i, i goes from 1 to K class intervals, fi, x_i minus A, with power of a r, and so on. So, now this is the basic idea behind defining a function which is called as moment. Now, what is the advantage of this thing, you will see here that when I try to take specific values of r and A, we get different types of characteristics, for example, when I choose r equal to 1, and A equal to 0, I get arithmetic mean, which is giving other information on the central tendency of the data. Similarly, if I take r equal to 2, and A is equal to sample mean, then this quantity is giving us the information about the variation in the data. Similarly, we can think about that there are some other properties of the frequency curve like a symmetry or it's hump, hump of the curve, how to quantify those information? So, on a similar line as these two values are going to give us the information on central tendency and variation in the data. So, we can use the concept of moment to quantify the information, about the symmetry of the frequency curve or the hump of the frequency curve, but in order to do so, first we need to understand the concept of moment, and what are the basic definitions in grouped case, in ungrouped data case, and beside the idea which I have given you here, there are some other aspects. So, in this lecture, I am going to consider about the raw moment and central moments, and in the next lecture I will try to show you, how to compute them on the R software. So, in this lecture please try to understand the basic concepts, and they are going to help you in the, for that lectures and in future,

Refer Slide Time: (11:58)



Before going into the further details, let me specify my notation for grouped and ungrouped data. Now, I'm going to consider two cases, case one where the data is ungrouped, and the variable X is discrete, and we have obtained n observations on X, which are denoted as small x_1, small x_2, small x_n.

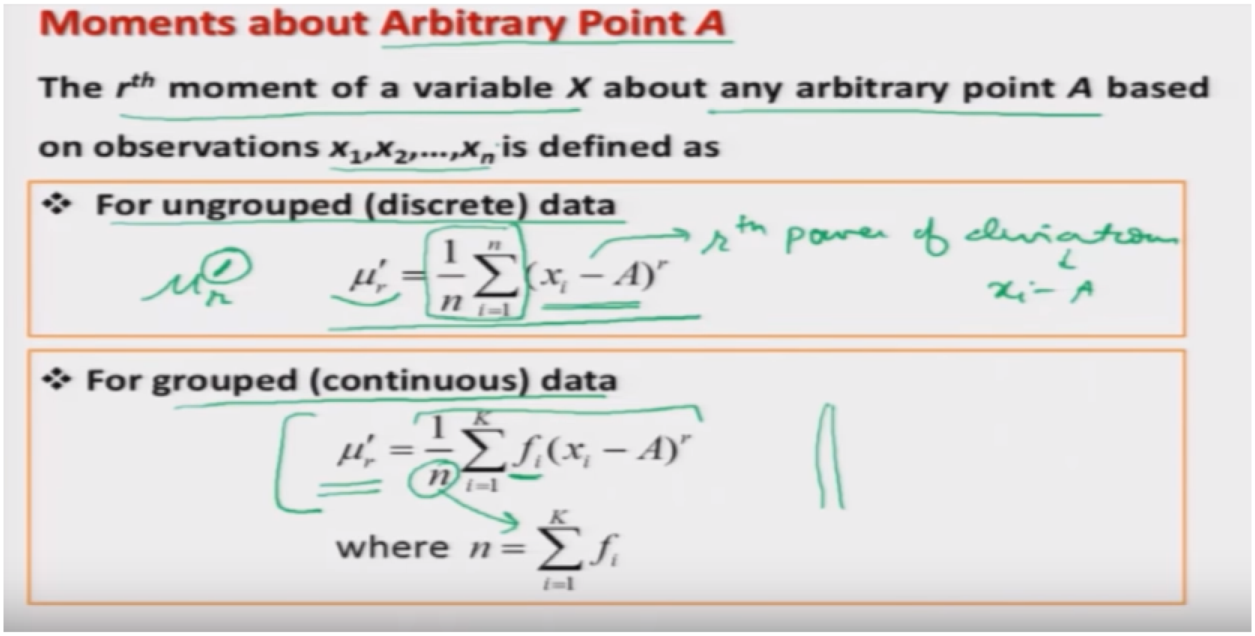
Refer Slide Time: (12:28)



Similarly, the second cases, when we have the data which is grouped, and we have a variable X, capital X which is continuous in nature, and the data is obtained on X, and the same data has been tabulated in K class intervals in a frequency table like as here, first column of the frequency table is giving us the class intervals here, like even e_1 to e_2, e_2 to e_3 and so on. So, these are essentially, you are here k class intervals, and the second column is giving us the information on the midpoints of this class interval. So, the idea noted here is x_1, x_2, x_k. So, x_1 is going to give us the information on e_1 plus e_2, divided by 2 and so on, which is obtained from here. So, the second column is giving us the value of x_1, x_2, x_k, but here x_1, x_2, x_k

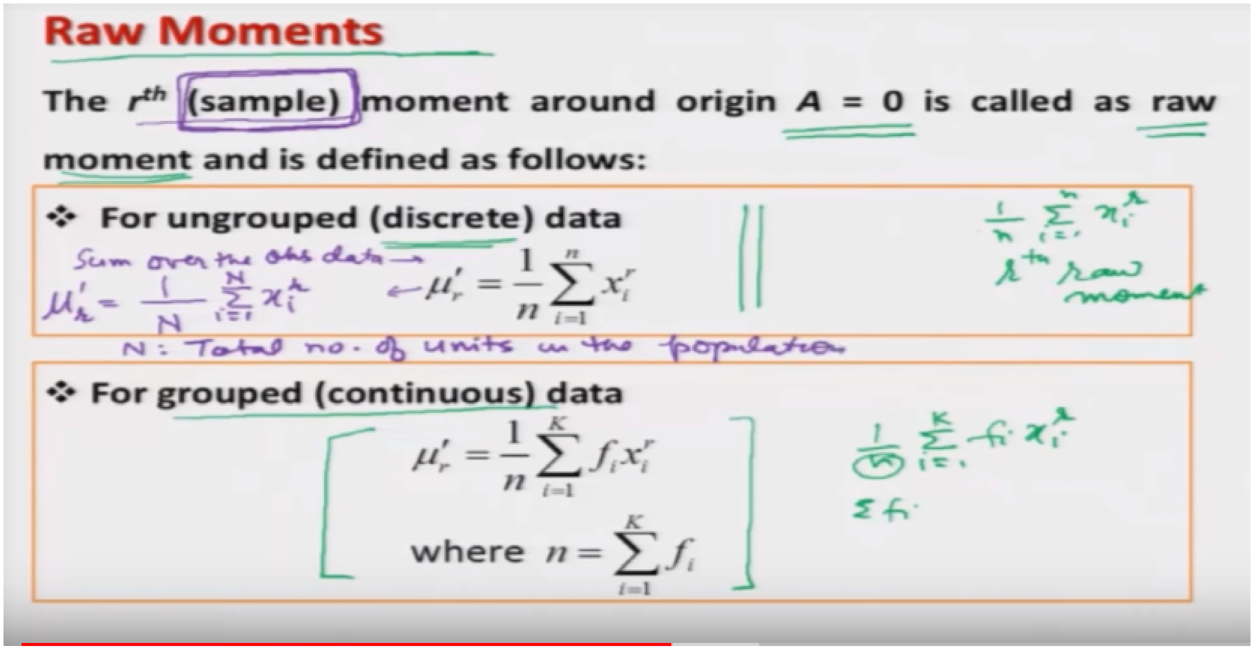
are the midpoints, and similarly in the third column we have the frequency data f_1, f_2, f_k which is denoting that, the first interval e_1 to e_2, whose midpoint is x_1, has frequency f_1, e_2 to e_3 which is having the frequency f_2 and so on, and the sum of all this frequency is denoted by here n. So, I can see here that the midpoints x_1, x_2, x_k they have got the frequency f_1, f_2, f_k respectively.

Refer Slide Time: (14:04)



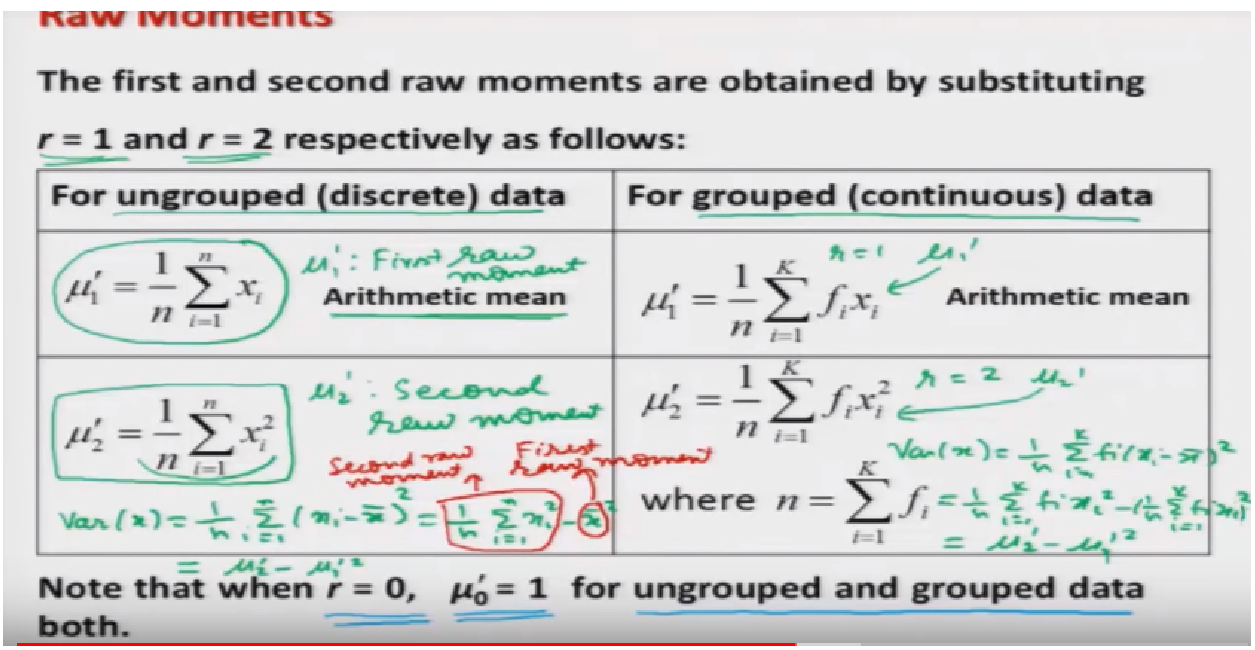
Now, after this I first try to define the moments about any arbitrary point here, here. So, let me define here the general r^{th} moment as we have just discussed, so the r^{th} moment of a variable x about any arbitrary point A, based on the observation x_1, x_2, x_n is defined as follows, for the case of ungrouped data, discrete data, this is simply the sum of the r^{th} power of deviations, and these deviations are obtained, you can see here x_i minus A, can see here A, and then the arithmetic mean of these deviations has been obtained, and this is denoted as mu r prime, this symbol here, I see mu r, and this symbol here, this is called here prime. So, this is the basic definition of the r^{th} moment for the ungrouped data around any arbitrary point A. Now, similarly if you have group data on a continuous variable, the same can be similarly defined, that now I am trying to find out the weighted mean of the other power of the deviations, and weights are given by the frequency, Right! and this is also converted as mu r prime. Although, I agree that these two the newer prime for the grouped and ungrouped data, they should have different type of notation, but we are using here same notation, because at a time we are going to handle either the discrete case, or the continuous case. So, this is how we try to define the r^{th} moment of a continuous variable x, around any arbitrary point A, and here this n is going to be the sum of all the frequencies. So, you can see here, this is the same function that we have just developed.

Refer Slide Time: (16:32)



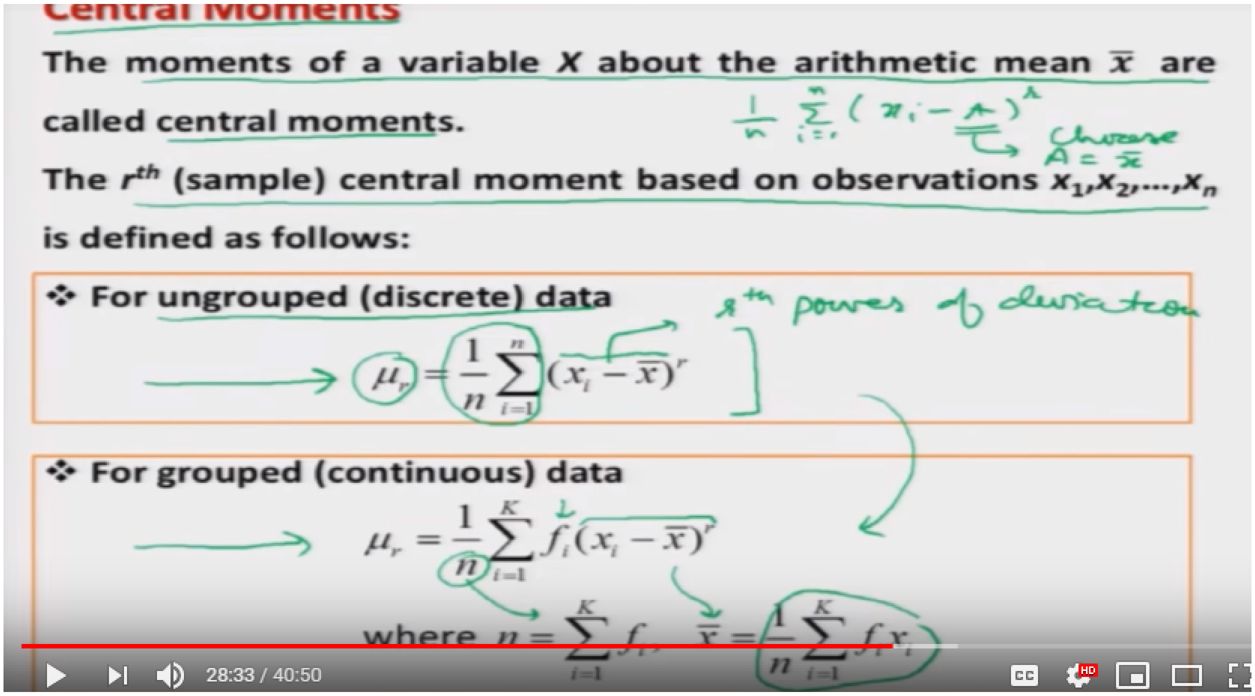
Now, I try to give you the idea of what is called here raw moments, now you will see that, most of the things they are simply the particular case of what we have obtained. So, as you had seen in the case of arithmetic mean, we are trying to take the mean of say x_i. So, similarly instead of taking here x_i, I am trying to take the r_{th} power of x_i, and this is going to give us the value of the r_{th} moment, or this is called as the r_{th} raw moment. So, if you try to see this is nothing but in the earlier case if you try to substitute A equal to 0, you get the same thing. So, the r_{th} sample moment around the origin A is called as raw moment, and it's defined here like this, in case the data is discrete, and similarly in case, if you have the group data on a continuous variable, the definition of discrete can be extended to continuous case, where you are simply trying to find out the average of the group data, say f_i, x_i, and then you try to convert x_i, into x_i is power of here r, and here n is going to be the sum of all the frequencies. So, this is how we define the r_{th} raw moment in this of continuous data, one thing you would like to notice here, that in the first line of the definition I am using here a word sample, you can see my here pen, why I'm trying to write down here the sample, and that is inside the bracket, what does this mean? Now, you can see here, I will try to explain this concept in this color, so you can notice it, suppose I try to take the first case, where I have defined the r^{th} moment for the ungrouped data, you can see here, that I am trying to take here the sum over the observed data, what does this mean, that I have taken a sample of data and I am trying to find out the average over the sample values, but if you try to see, there are two counter parts one is the sample and another is the population, what is the sample, this is only a small fraction from the population. So, as we have computed this quantity on the basis of a sample, the similar quantity can be estimated for the entire population units, and essentially in statistics, we are interested in knowing the population value, but since this is unknown to us, so we try to take a sample and we try to work on the basis of sample. So, as we have defined the moments, or say r_{th} moment or raw moment or other types of moment that we are going to define, we are going to define on the basis of sample, but surely there exists a value in the population. So, the counterpart of the sample is the population, and when we try to compute the same moment, on the basis of all the units in the population, that would be called as the population moment, or the moment based on the population values. So, that will be defined here, for example, if I try to define this mu r prime, for the population values, and suppose if I say that there are suppose, capital N values in the population, then in this case this is going to be i goes from 1 to capital N, instead of a small n, and this value is going to be here, x_i this power of here r. So, N is here the total number of units in the population. Now, there are two counter parts, one for the population and, say another for the sample but obviously our interest here, is to compute the values on the basis of sample. So, in this lecture and in the previous lecture, I will try to define these values, and you have to just be careful that whatever we are doing here, they are on the basis of sample. So, in practice we always do not call it as a sample moment, but we simply call it moment, but the interpretation is that, that the moments have been computed on the basis of given sample of data. So, this is what you always have to keep in mind, Okay? So, now we try to take a particular example,

Refer Slide Time: (22:03)



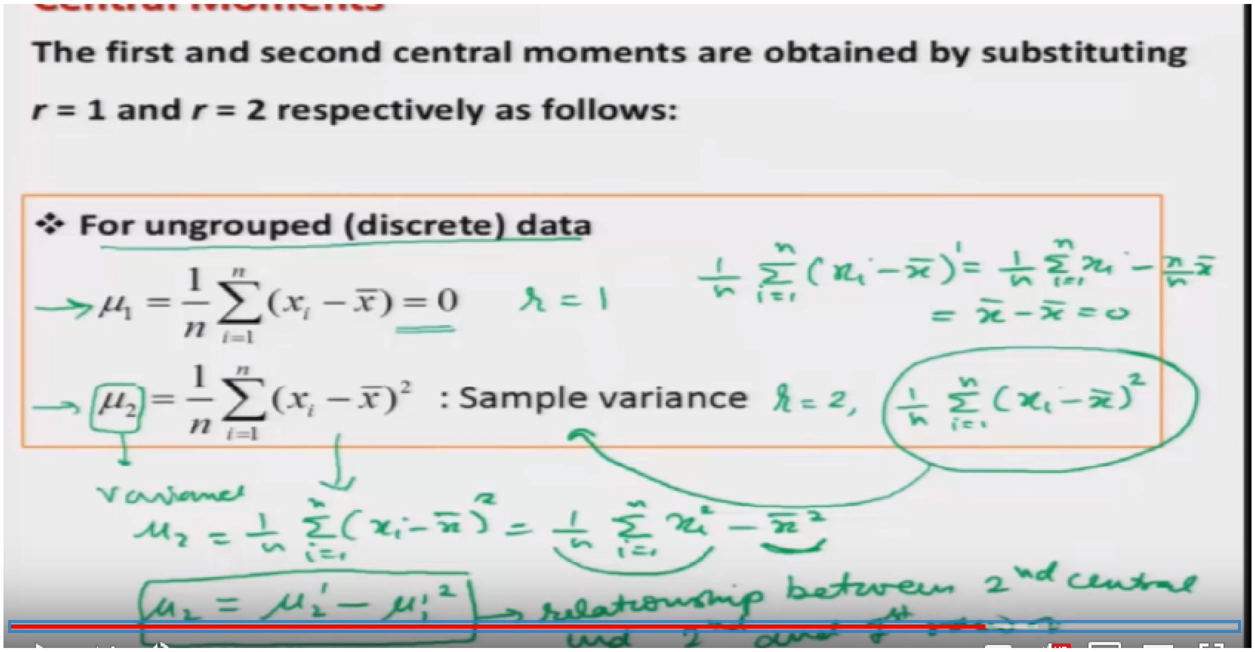
of this raw moment, and I simply try to choose two values r equal to 1, and r equal to 2, and then in the case of ungrouped data, you can see here as soon as I say, r equal to 1, I get this value, mu 1 prime is equal to 1 upon n, summation x_i, and if you try to identify what is this, this is nothing but your arithmetic mean, and this mu 1 prime, this is called as first raw moment, and similarly, if you try to put r equal to 2, then I get here mu 2 prime is equal to 1 upon n, summation i goes from 1 to n, x_i square, this quantity mu 2 prime is called as second raw moment, and if you try to recall, where this quantity was occurring, you may recall that when you wrote the expression of variance of x, this was written as 1 upon n summation i goes from 1 to n, x_i minus x bar whole Square, and which we have written as 1 upon n, i goes from 1 to n, x_i square, minus x bar square. So, you can see here, that the first term here this one, this is nothing but your second raw moment, and this quantity here x bar, this is simply the first raw moment. So, now this will give you an idea, that why this raw moments are needed, and now you can see here that this raw moments are going to help us in defining the variance of x, Right! So, in this case I can redefine the variance as say here, mu 2 prime, minus, mu 1 prime square, and similarly if you try to take the grouped data, then when I substitute r equal to 1, I get here mu 1 prime which is defined here like this, which is the arithmetic mean, once again and when I try to put r equal to 2, then I get mu 2 prime which is here defined here like this, and this also has the same interpretation, that if you try to see variance of x bar, we had defined 1 upon n, summation i goes from 1 to K, fi, x_i, minus, x bar, whole square, and we had denoted it while i goes from 1 to K, f_i, x_i, square, minus, 1 upon n, summation f_i, x_i, whole square, I was wrong! 1 to K. So, this is nothing but your mu 2 prime, minus, mu 1 prime square. Now, see here by this example, that the interpretation and the utility of the moments in case of grouped and ungrouped data, this is similar, Right! what you have to just observe here, that if I try to choose here r equal to 0, then in this case, this mu 0 prime, becomes 1, and this is true for both the cases, for ungrouped data and for grouped data. Now, after discussing the concept of raw moment,

Refer Slide Time: (25:51)



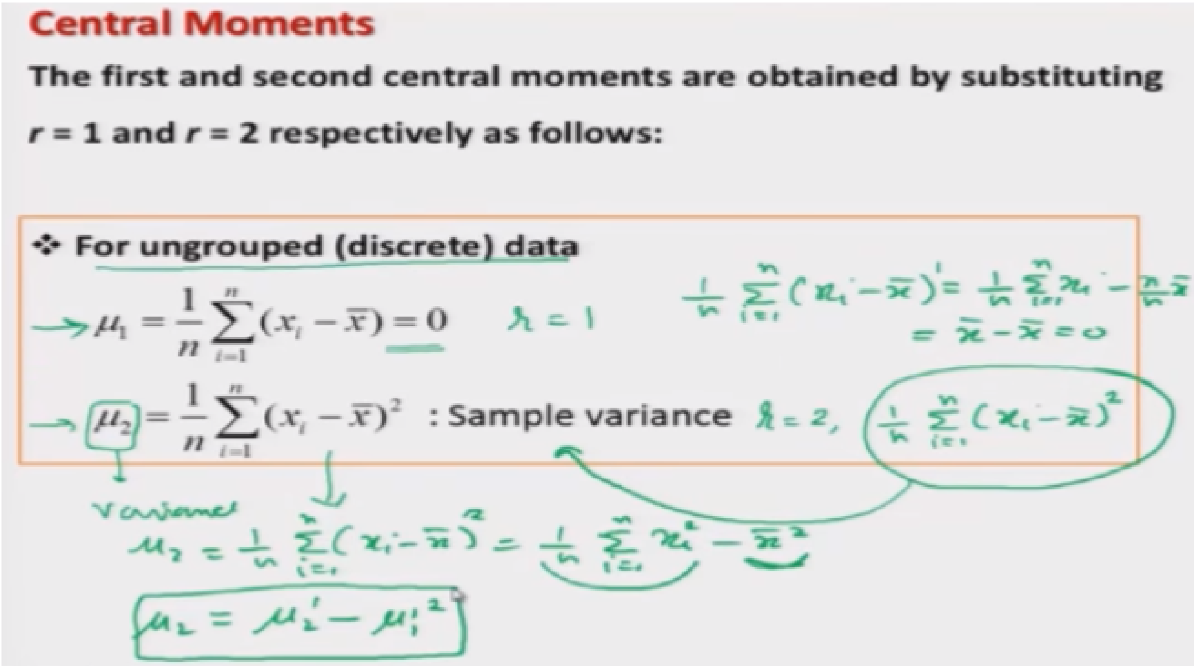
let me define here what is called here as central moments. So, we had discussed the r^{th} moment, around any arbitrary point, see A, x_i minus summation of x_i minus A, raise to power here r, and in this case, if you simply try to choose A, to be the sample mean x bar, then whatever is the moment, that is called a central moment. So, the moments of a variable X about the arithmetic mean of the data, they are called as central moments. So, now if I want to define the r^{th} central moment based on the sample data x_1, x_2, x_n, then in the case of ungrouped data, I simply have to replace A by x bar. So, this quantity becomes here like this, so if you try to see here this quantity here is nothing, but the r^{th} power of deviations, and then arithmetic mean of these are the power of deviations have been taken, and this quantity is defined here as say mu r, there is no prime in this case. So, that is the standard notation, that the raw moments are defined by mu r prime,, and the central moments, they are defined by mu r. So, the r^{th} central movement in case of ungrouped data is given by this mu r, is equal to 1 upon n, summation i goes from 1 to n, x_i minus x bar raised to the power of r, and similar definition can also be extended to the case of group data, and in this case, this is the r^{th} power of the deviation, and they are multiplied by the frequency f_i, and then the arithmetic mean of this quantities have been taken where, n is equal to the sum of all frequencies, and what you have to just notice, that in this case you are going to compute the arithmetic mean by this expression, that is 1 upon n summation, i goes from 1 to K, f_i, x_i and whereas in the case of discrete data, you try to compute the x bar, simply as a summation x_i upon n. So, this quantity is called as the r^{th} central moment of our variable X. Now,

Refer Slide Time: (28:34)



let me try to choose here, r equal to 1, and r equal to 2, and see what happens to the first and second central moments. So, first I try to take the case of ungrouped data, and this case if I try to substitute r equal to 1, you can see here what I will get 1 upon n, summation i goes from 1 to n, say x_i, minus, x bar is the power of here r, which is here 1. So, this is, I can write down here summation x_i, i goes from 1 to n, minus, say n upon n, x bar. So, this quantity becomes is x bar, minus, X bar is equal to 0. Now, I can say that here, that the first central moment, in case of discrete data will always be equal to 0, and in the next slide, I will show you, that the same is 2 in the case of continuous data also. So, in general I can say, that the first central moment is always 0, similarly when I try to substitute r equal to 2, then what I get here, 1 upon n, summation i goes from 1 to n, x_i, minus, x bar, raise the power of here r, r is equal to 2. Now, can you identify this thing, what is this thing? This is nothing but your sample variance, what we have defined in the earlier lectures. Now, I can say, that the second central moment which is denoted by mu 2 is representing the variance, or the sample variance. One thing I would where I would like to have your attention is the following, you can see here the second central moment is going to represent the variance, and first central moment is always 0. Whereas, the first raw moment that was denoting the arithmetic mean. So, when you are trying to interpret these moments you have to be careful while making an interpretation for the arithmetic mean. Arithmetic mean is represented by the first raw moment, and in the case of first central moment, this is simply de noting the sum or the averages of the deviation around mean which is always 0.

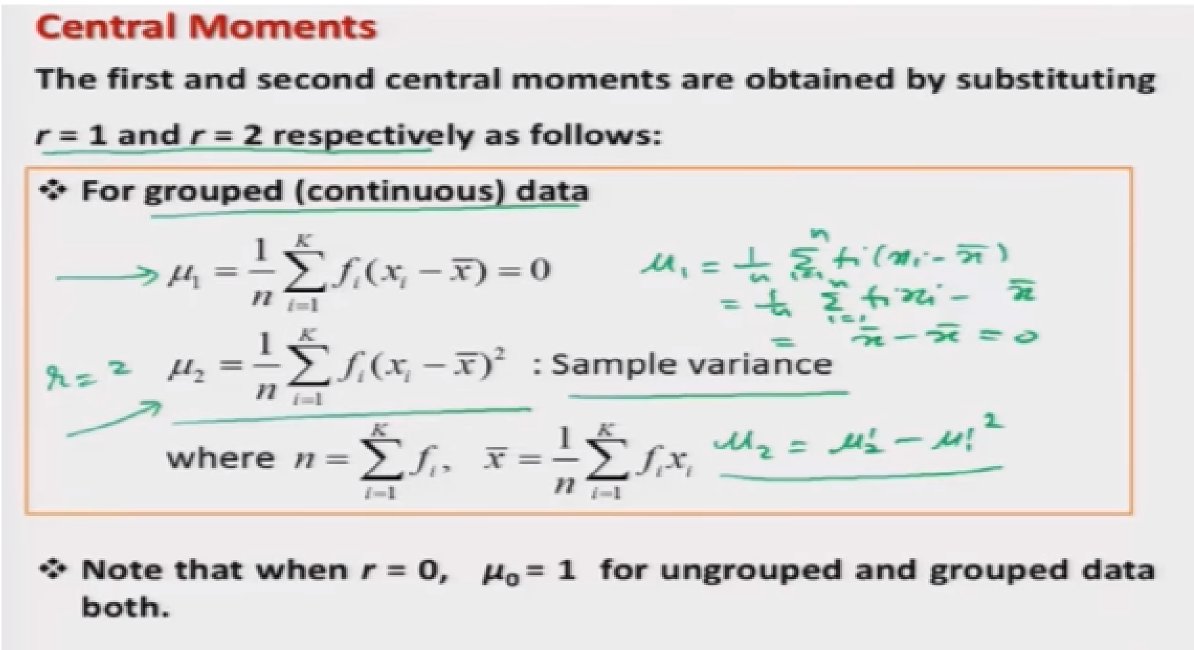
Refer Slide Time (31:21)



Now, in this is step I would also try to show you here, that when I try to write down here this expression. Then mu 2 was written has here 1 upon n summation I goes from 1 to n xi minus x bar whole square which was written as 1 upon n summation I goes from 1 to n xi square minus x bar square. So, I can write down here this quantity here, this is nothing but my mu 2 prime, and this quantity here x bar square this is mu 1 prime square. So, mu 2 becomes mu 2 prime minus mu 1 prime square. So, you can see here, that this is a relationship between second central and second and first raw moments.

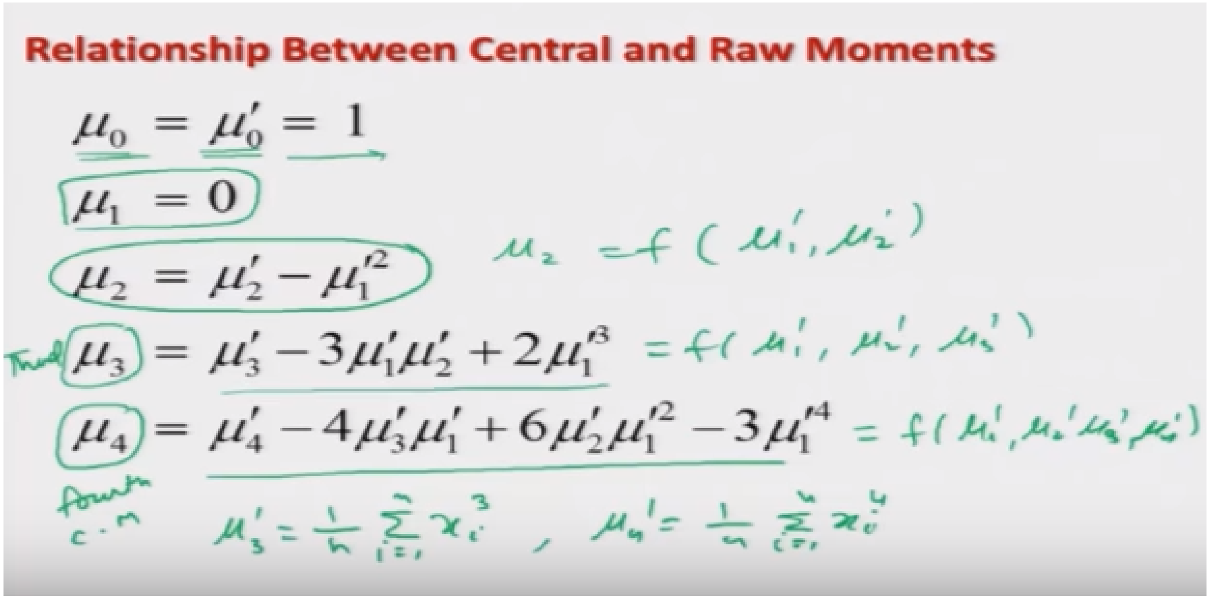
So, what I am trying to explain you here? I am trying to ensure you here, that there exists a relationship between the central and raw moments. I have shown you here that, how to express the second central moment as a function of raw moments, and I have taken here the example of r equal to 2. Similarly, if you try to take r equal to 3 r equal to 4 and so on. Then you can obtain a similar type of relationship between central and raw moment. One more important aspect which arises here, is the following usually you will see that in statistics, we are considering the 1st 2nd 3rd and 4th moments. Whether the raw moments or the central moments. Well, one can very easily compute the higher-order moments say 56 and so on. But up to now, what has happened that we have the interpretation for the first four moments? For example, first moment that's the first raw moment that is indicating the value of or the it is quantifying the information of central tendency of the data. Second central moment is giving us the information on variability. Similarly, I will try to show you in the forthcoming lectures, that third moment is giving us the information on the symmetry of the frequency curve. That is called as property of skewness, and similarly the fourth moment will give us the information about the hump of the frequency curve. The peakedness of the frequency curve and that property is called as kurtosis. But what is indicated by fifth moment sixth moment and so on? That is still an open question? So, that is the reason, that usually we are interested in finding out the moments up to order four. There also, exists a clear-cut relationship between rth central moments, as a function of raw moments. But here I am not showing you here, I am not discussing it here. But I will request you please try to have a look in the books, and there it is clearly mentioned. But here I would certainly show you, that what is the relationship of first four central moment with respect to the raw moments.

Refer Slide Time (35:26)



Now, if I try to take the r equal to 1and r equal to 2 in case of continuous data group data, then we have the simple similar interpretations. That in the case of first central moment this is going to be fi xi minus x bar I goes from 1 to n. So, this is here 1 upon n summation I goes from 1 to n fi xi minus see here, x bar. So, you can see here this is nothing but, x bar minus x bar is equal to 0. So, once again I have shown you that the first central moment in case of continuous data, is always 0, and similarly, if you substitute here r equal to 2, then this quantity is nothing but the sample variance, and in this case also, this mu 2 can be represented as mu 2 prime minus mu 1 prime c square. So, this is again the same outcome that we have obtained in the case of ungroup data. So, you can see here and notice that this relationship is not going to change in the case of group and ungroup data. Now, if I try to choose here r equal to 0, then I will always get mu 0 equal to 1 either we have a ungroup data or a group data set.

Refer Slide Time (36:50)



Now, after this I will show you what is the relationship between central moments in raw moments? You have observed, that the fourth central moment and first raw moment both are taking the value 1, and first central moment here, this is taking the value always 0. These are the two points where you have to be careful. I already have shown you this result, where I have shown you how I can express, the second central moment mu 2, as a function of first raw moment and second raw moment. Just using a similar concept I can also, express the third, and fourth central moments, as a function of raw moments as a function of mu 1 prime, mu 2 prime and mu 3 prime, and similarly the 4th central moment can be as a function of mu 1 prime, mu 2 prime, mu 3 prime and mu 4 prime and you may recall, that mu 3 prime is nothing but, 1 upon n summation I goes from 1 to n xi cube and mu 4 prime is, which is the fourth raw moment. I goes from 1 to n summation and, then X is the power of here four. So, you can see here that these are the relationship of central moment and raw moments. Well, I am trying to give you here only four relationship as I said. But using the binomial expansion you can obtain a general relationship between the r^{th} central moment as a function of the first r raw moments and that is available in in all the standards statistics books. So, I will not do it here. But I would like to stop here. I have given you the basic concepts of moments. Well, this lecture was purely theoretical. But we also know, that any development is based on the theoretical construct. Unless and until you understand the basic fundamental you will not be able to understand what the software is trying to do? and these are very important concepts in the subject statistics. So, that was my objective to give you an exposure of this concept. Now, in the next lecture, I will try to consider the concept of absolute moment, and I will try to show you that how these moments can be computed on the basis of r software, and after that I will introduce the concept of skewness and kurtosis. So, please you take a break revise the lecture try to read from the books, and I will see you in the next lecture. Till then, Good bye!