Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 68
Applications of N-P Lemma – I

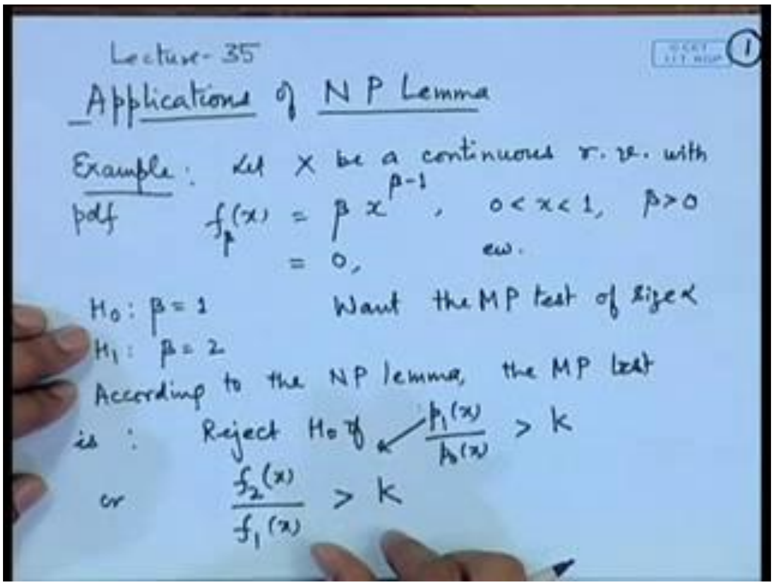
In the previous lecture I had introduce of basic concepts of testing of hypothesis. So, let me review the basic terminology- a test of a statistical hypothesis is testing about the probability distribution of a certain population. We may be able to know that what is the proper probability distribution and then we may test about the parameters of the distribution. We have a null hypothesis and the alternative hypothesis.

So, the test is to decide on the basics of a random sample whether to accept or reject a null hypothesis. If the sample supports the hypothesis; that means it is in favor of the hypothesis then we say that we cannot reject the hypothesis or we say we accept a null hypothesis, otherwise we say we that reject the null hypothesis. We have classification of the hypothesis as a simple hypothesis and a composite hypothesis. So, a simple hypothesis is the when the hypothesis statement completely specifies the probability distribution otherwise we call it a composite hypothesis.

When we conduct a test of hypothesis; that means the decision is based on a sample then we may commit two types of errors which we called as type I error and type II error. That is we may reject a true hypothesis or we may accept a false hypothesis. We have seen that it is not possible to minimize the probabilities of both types of errors to a minimum. So, a practical approach is to fix the highest level for one type of error usually we fix for the type I error and find out a test of hypothesis for which the other type of error is minimized or 1 minus that is maximized which we call the power of the test. That gave the concept of the most powerful test.

In the last lecture I explained that there is a result known as Neyman-Pearson fundamental lemma which for simple hypothesis versus a simple hypothesis problem gives a most powerful test. So, now let me go for the application of this Neyman-Pearson Lemma.

(Refer Slide Time: 03:03)



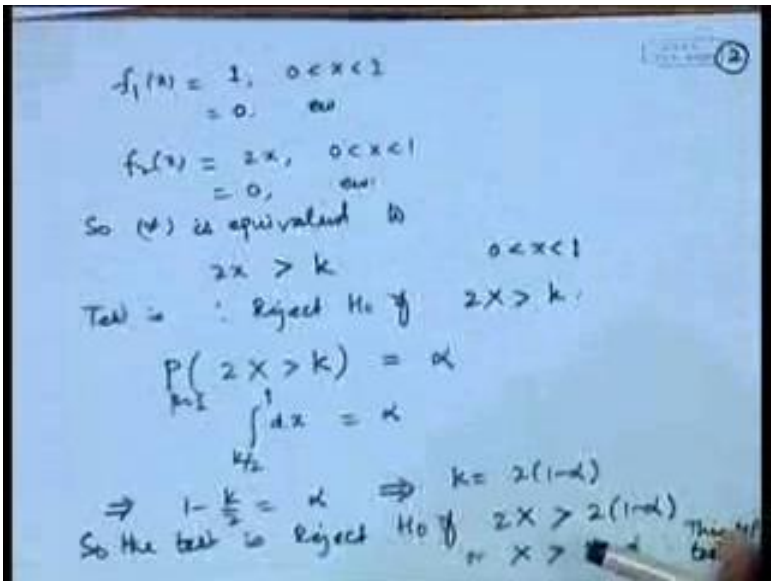
Let me start with a following example: let x be a continuous random variable with probability density function given by f x is equal to say beta x to the power beta minus 1 for 0 less than x less than 1 where beta is a positive parameter and the density is 0 elsewhere. We want to test say hypothesis beta is equal to 1 against say H 1 beta is equal to 2.

So, here if we see this is f beta the density is dependent upon the parameter beta. So, we are interested to test that whether beta is equal to 1 or beta is equal to 2, now you can see here that both of these are simple hypothesis. So if we want to find out, if we want the most powerful test then we can make use of the Neyman-Pearson lemma. So, we want the most powerful test of size alpha or level alpha. So, according to the Neyman-Pearson lemma the most powerful test is reject H naught if p 1 x by p naught x is greater than k.

So, now this is the quantity which we can analyze here. What is p 1 and what is p naught here? This is corresponding to; that is p 1 is corresponding to the value of the probability distribution or density when the alternative hypothesis is true. So, here it will be f 2 x divided by, p naught is the value of the hypothesis value of the probability distribution when the null hypothesis is true here beta is equal to 1; that means, it will become f 1 x this is greater than k. Where, the constant k is chosen in such a ways that the probability of the type I error is equal to alpha.

So, now first of all let us look at when are we actually going to reject. So, this statement is equivalent to.

(Refer Slide Time: 05:59)

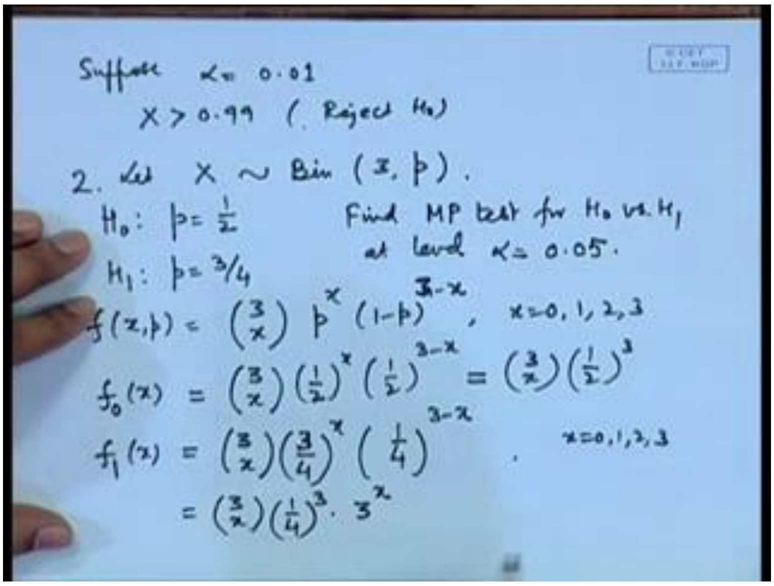


So, we have to consider the value here. What is f 1 x? F 1 x will be obtained by substituting beta is equal to 1 here which gives us simply 1; that means, the uniform distribution on the interval 0 to 1. In a similar way f 2 x if I put beta is equal to 2 here I will get 2 x. So, by statement f 2 x by f 1 x greater than k this is equivalent to, so this statement let me call it as star this is equivalent to 2 x divided by 1 is greater than k; of course here you are taking 0 less than 1.

Now, we want that probability of type I error must be 1. So, the test is reject H naught if x is greater than k or you can say 2 x is greater than k. Now we want probability of 2 x greater than k when it is true; that means, when beta is equal to 1 this probability to be equal to alpha. Now, when beta is equal to 1 we have written here the density is uniform distribution. So, this value can be calculated this is probability of x greater than k by 2. So, this becomes integral of say dx from k by 2 to 1 this is equal to alpha, or you can say 1 minus k by 2 is equal to alpha which is implying k is equal to twice 1 minus alpha.

So, the test is in theoretical terms we can write reject H naught if 2 x is greater than twice 1 minus alpha which is equivalent to x is greater than 1 minus alpha. So, a most powerful test of size alpha is reject H naught when x is greater than 1 minus alpha; so this is the most powerful test. So, this is most powerful test. So, you can see here, now the decision making process is quite simple we observe a random variable from this population and we see its value.

(Refer Slide Time: 08:57)



So, suppose I say that alpha is equal to say suppose alpha is equal to say 0.01 then I should observe x to be greater than 0.99, then only you will reject H naught. On the other hand if you observe x to be between say less than 0.99 or less than or equal to 0.99 you have no reason to reject H naught. So, this is the test function for the most powerful test.

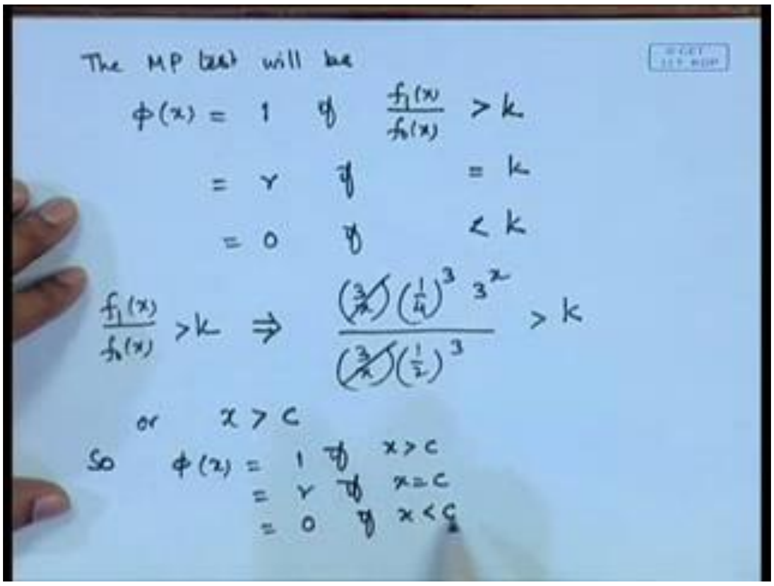
Another point which you should observe here that when we wrote Neyman-Pearson lemma- we wrote acceptance region to be when this is less than k and there was a probability gamma of rejecting when this is equal to k. But since this is a continuous distribution we do not have to look at that region, because we are able to achieve the exact level alpha by this test here. So, we can simply state in the form that when we are rejecting or when we are accepting.

The point x equal to 1 minus alpha does not make any difference, because that as probability 0. In the case of discrete distribution we may have to take some randomization which is explained through the following example. Let me take this example here: let x be a binomial random variable with parameter say 3 that is n is equal to 3 and probability of head is say p. We want to test say H naught p is equal to half against H 1 say p is equal to 3 by 4. So, find most powerful test for H naught against H 1 at level say alpha is equal to say 0.05.

Now, this is again a case of simple versus simple hypothesis, because p is equal to half or p is equal to 3 by 4 completely specifies this probability distribution. Therefore, we will consider the application of the Neyman-Pearson lemma here. So, let us write down the distribution first. So, f x p that is equal to 3 c x that is m c x p to the power x 1 minus p to the power n minus x. We will need the values of f naught x and f 1 x. So, f naught x is the value when p is equal to half which is reducing to 3 c x half to the power x into half to the power 3 minus x; this is 3 here x is equal to 0, 1, 2, 3.

So, naturally this is simply equal to 3 c x half cube, whereas f 1 x is the density when the alternative hypothesis is true that is p is equal to 3 by 4. So, the value is 3 c x 3 by 4 to the power x 1 by 4 to the power 3 minus x for x is equal to 0, 1, 2, and 3. Now this can also be simplified little bit we can write it as 3 c x 1 by 4 to the power 3 and 3 to the power x.

(Refer Slide Time: 12:54)

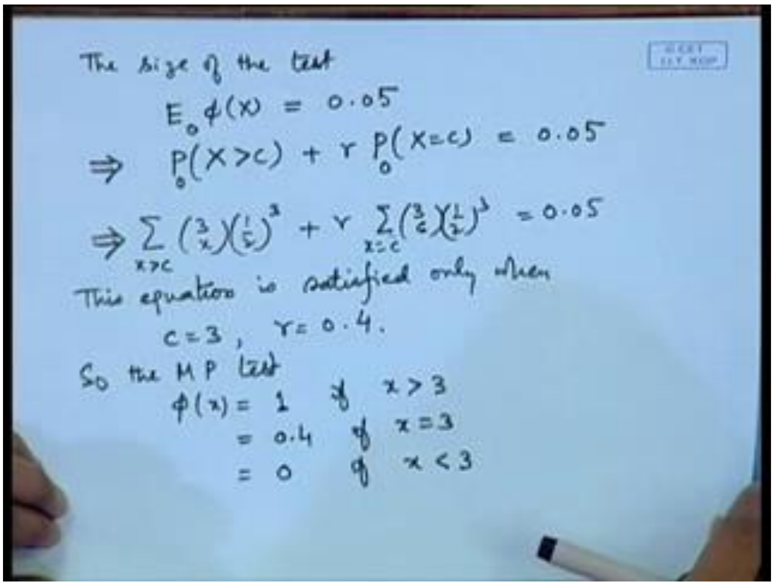


So, the most powerful test form we can write; the most powerful test will be- so since here randomization may be required we write the test function.

So, phi x is equal to 1 if f 1 x by f naught x is greater than k, it is equal to gamma if this is equal to k, it is equal to 0 if this is less than k. So, this condition that f 1 x by f naught x is greater than k let us write down this condition here 3 c x 1 by 4 cube 3 to the power x divided by 3 c x half cube greater than k. So this term cancels out, this is some constant and if I take logarithm here then this will become x log 3 greater than some constant. So, we can say x is greater than some c.

So, phi x function can be written to be 1 if x is greater than c, it is equal to gamma if x is equal to c, it is equal to 0 if x is less than c. This is the test function that we will be getting. That means, rejecting when x is greater than c, accepting when x is less than c, and rejecting with probability gamma when x is equal to c. This is the randomization part here; here it may be required as we will see now. Now the size of this test must be equal to alpha that is 0.05.

(Refer Slide Time: 14:46)



So, if we put that condition; the size of the test that is expectation of phi x when null hypothesis is true that is equal to 0.05. So, this value is equal to probability x greater than c when p is equal to half plus gamma times probability x is equal to c when p is equal to half that is equal to 0.05.

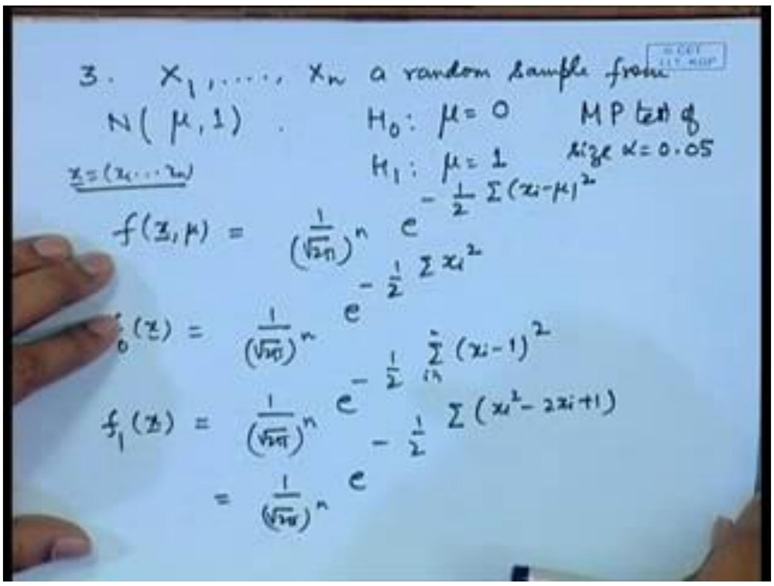
Now you can see here, when the null hypothesis is true the density function is written as 3 c x half q. So, this becomes that we have to consider the values of x for which it is greater than c and the probability distribution has to be added up is this; that is 3 c x half cube summation when x is greater than c plus gamma into; well this is becoming 3 c c half cube when x is actually equal to c. So, basically there is a point here which will be satisfied for integer's values only. So, we will see that when is it satisfied; 0.05. This equation is satisfied only when; so you will substitute the values of c is equal to 0, 1, 2, and 3 we get here c is equal to 3 and gamma is equal to 0.4.

So, the MP test is phi x is equal to 1 if x is greater than 3, it is equal to 0.4 if x is equal to 3, it is equal to 0 if x is less than 3. So, let us look at the interpretation of this. The interpretation of this test is because x is taking values 0, 1, 2, 3 only; that mean this test it is never rejecting with probability 1. It is rejecting only with probability; that means when we conduct the experiment and if I observe x is equal to 3 then we will reject with probability 0.4 and accept with probability 0.4. In all other cases we accept the null hypothesis; that means, if x is equal to 0, 1, 2 then we do not reject H naught.

So, this may look surprising, but if we see carefully our problem; the problem was to test that whether the coins is fair against whether it is biased in favor of head. So, biased in favor of head we are accepting if x is greater than 3 only and which is not possible. Even if x is equal to 3 we are only partially agreeing; that means, we are accepting in favor of H 1 only that means you are rejecting only with probability 0.4. That means, there is a hypothesis is heavily biased in favor of H naught here, because x is equal to 0, 1, 2, 3, so only you are having the rejection for x equal to 3 that to with a probability 0.4 here.

So, we can see here that by application of the Neyman-Pearson fundamental lemma we are able to get the most powerful test. Of course, it is another matter that if we change this alpha to be say 0.01 or 0.1 then the test will be slightly modified here.

(Refer Slide Time: 19:11)



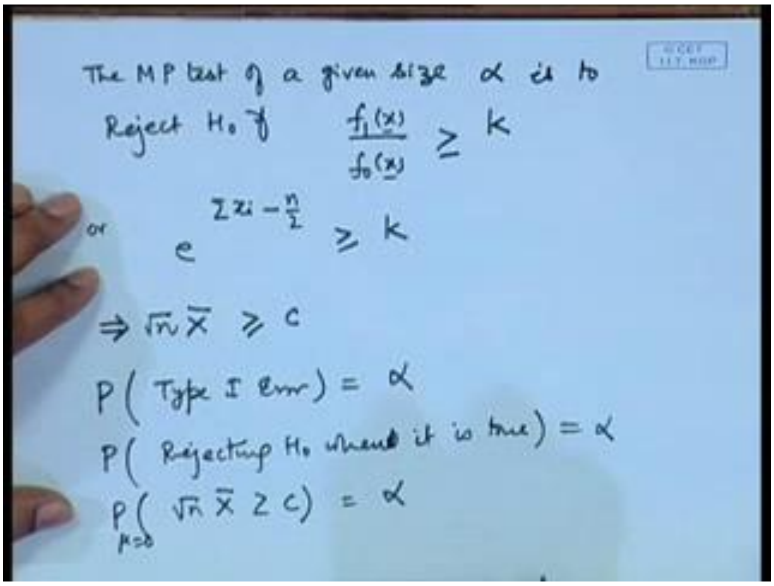
Let us look at the applications to the normal distribution here. Consider say- x 1, x 2, x n a random sample from say normal distribution with mean mu and variance say unity. We want to test the hypothesis say mu is equal to 0 against say mu is equal to 1. We want the most powerful test of a certain size say alpha is equal to 0.05 say. So, the first thing is that in application of the Neyman-Pearson lemma we need to write down the probability density function that is f x mu.

Now, here x means we are observing a sample x 1, x 2, x n, therefore we need to write down this density function as a joint density function of x 1, x 2, x n. So, this is turning out to be 1 by root 2 pi to the power n e to the power minus 1 by 2 sigma xi minus mu square. So, our f naught value that is there when mu is equal to 0 and this turns out to be 1 by root 2 pi to the power n e to the power minus 1 by 2 sigma xi square. In a similar way f 1 x is equal to 1 by root 2 pi to the power n e to the power minus 1 by 2 sigma xi minus 1 square. Now this term can be simplified little bit we get it as 1 by root 2 pi to the power n e to the power minus 1 by 2 xi square minus twice xi plus 1.

So if we take the most powerful test, because the hypothesis H naught and H 1 both are simple hypothesis so we can apply the Neyman-Pearson fundamental lemma to get the most powerful test for a given size. So, the most powerful test of a given size say alpha is to reject H naught if f 1 x by f naught x is greater than k. Since it is a continuous distribution we will be able to achieve the exact level alpha by a non randomized test itself.

So, we need not put here gamma for f 1 x by f naught is equal to k, we may just consider f 1 by f naught is greater than k or greater than or equal to k it does not make any difference here, because the probability of the equality will be equal to 0 for the case of a continuous random variable.

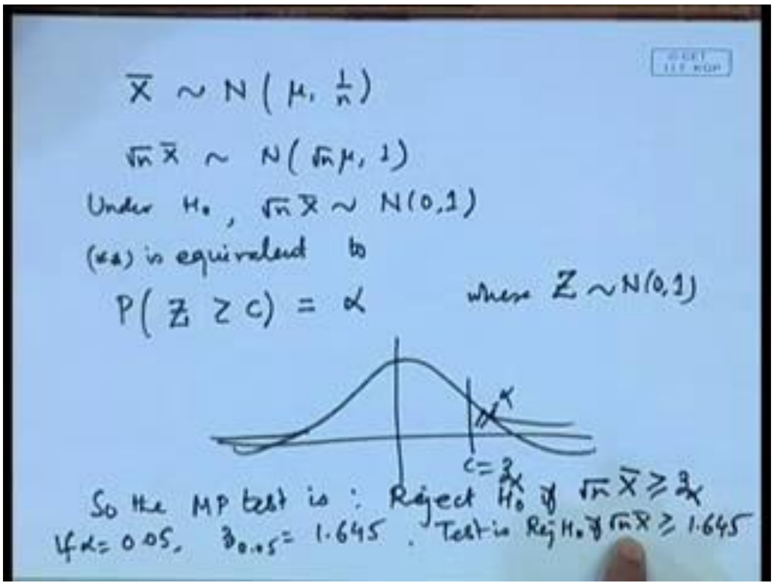
(Refer Slide Time: 21:43)



So, if we write these functions here now we are getting the f 1 and f naught term here both had the same factor so when we write the ratio this coefficient cancels out and also e to the power minus 1 by 2 sigma xi square will also cancel out. So, we will be left with e to the power sigma xi and then there will be a constant term minus n by 2 is greater than or equal to k. If we take the logarithm then this is reducing to x bar greater than or equal to say c term and we may multiply by root n here to get a proper form of the distribution. Why that is useful, because we want probability of type I error equal to alpha. So, that is probability of rejecting H naught when it is true that is equal to alpha. So, probability of root n x bar greater than or equal to c when mu is equal to 0 is equal to alpha.

Now you see the distribution of x bar.

(Refer Slide Time: 24:15)



Since, x 1, x 2, x n is a random sample from normal distribution x bar follows normal mu 1 by n. So, root n x bar follows normal root n mu 1. So, under H naught root n x bar follows normal 0, 1. So, from here the statement that probability of mu is equal to 0, so that this statement let me call it say statement double star this is equivalent to probability of z greater than or equal to c is equal to alpha, where z is a standard normal random variable.

That means if we are considering a standard normal probability density function then c is the point such that the probability beyond this is alpha so this is equal to z alpha. So, the most powerful test is reject H naught if root n x bar is greater than or equal to z alpha. So if I am considering alpha is equal to 0.05, then z 0.05 we know from the tables of normal distribution it is 1.645. So, the test is reject H naught if root n x bar is greater than or equal to 1.645.

So, you can see Neyman-Pearson lemma gives us a precise test for taking the decision to accept or reject a null hypothesis in a given situation. Now let us also look at the interpretation of this. We were testing the hypothesis whether mu is equal to 0 against mu is equal to 1. So, you can see here we want that whether the value of mean is less or more, because we may consider here this mu 0 is value which is less than 1. So, naturally here you can see that the as a Lehman you would have made a decision that for a larger value of x bar you will tend to favor H 1 and for a smaller value of H naught; for a smaller value of x bar you will tend to favor H naught.

But, how much value of x bar is considered to be larger or a smaller that is dependent upon the probability of type I error. And therefore, we are now able to formulate in the terms of that this decision making process as that root n x bar should be greater than; that means, x bar is greater than 1.645 by root n. Of course, if n is large then value will become much smaller. That means, even for a smaller value of x bar you will consider it to be little larger, but that much distinction is permissible because on a absolute scale we cannot compare 0 and 1.

One may say that the difference between 0 and 1 is 1, but what is a scale here. So, if we are having a pretty large value of n then that difference may be still considered to be large, whereas for a very small value of n that difference may not be considered to be large.