Statistical Inference

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

Lecture No. # 19

Bayes and Minimax Estimation – II

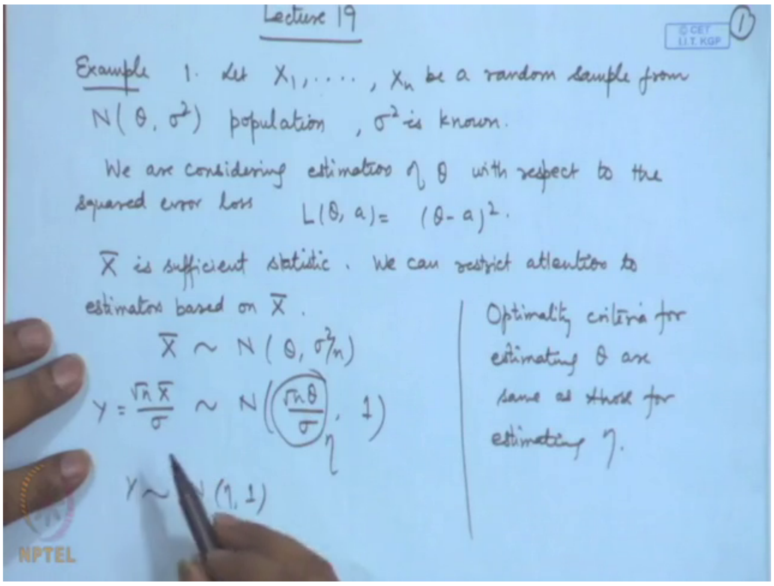
In the last lecture I have introduced two criteria for determining optimal estimators. As we discussed earlier, that in an estimation problem it is not possible to find the best estimator or which is optimal in every sense or which is having the smallest risk function throughout, except in the trivial or in consequential situations. So, what we did? We introduced, there are two methods of resolving these issues; one is that we can restrict attention to a smaller class of estimators. We can consider criteria such as unbiasedness or invariance and then find out the best in that class, that is the best unbiased estimator, which we call as uniformly minimum variance unbiased estimator or a best equivariant estimator, that is having the smallest risk in the class of equivariant estimators if we can introduce certain group of transformations under which the estimation problem is invariant.

Another method is that we can introduce a new criterion for ordering the decision rules, one is through averaging and this is called the Bayesian estimation. Here, what we do, we consider the parameter also, which is appearing in the distribution as a random variable and we take a distribution, which we call prior distribution. Now, we take the average of the risk function with respect to this prior distribution. So, the risk function is reducing to a number and now you order, that is the one, which is having the smallest Bayes risk, this is called Bayes estimator. And another criterion is, consider the worst possible scenario for every estimator. So, we take the maximum risk and then we minimize, that is known as the minimax estimator.

So, we introduced Bayesian and minimax estimation. Now, related to the Bayesian estimation, I introduce some additional concept, such as limit obvious rules and epsilon Bayes rule, extended Bayes rule or generalized Bayes rule.

So, today we will start with the elaboration of each of these concepts and then we will try to look at some relationship, how to find out minimax estimation given a sequence of Bayes estimators, etcetera.

(Refer Slide Time: 02:57)



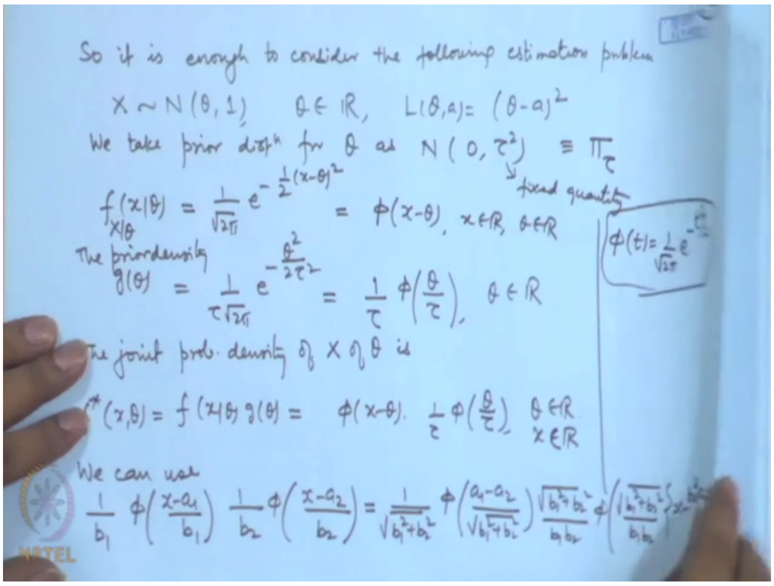
So, let us take, say X 1, X 2, X n be a random sample from normal, say theta sigma square population and sigma square is known. Now, if sigma square is known, in this case we are considering estimation of theta with respect to say, squared error loss function, that is, L theta a is equal to theta minus a whole square.

Now, note here, that when sigma square is known, X bar is sufficient statistic and we can restrict attention to the, we can restrict attention to estimators based on X bar. Now, what is the distribution of X bar? X bar follows normal theta sigma square by n. See, we can do some shifting, we can write root n X bar by sigma that will follow root n theta by sigma 1. Now, let us use some other notation, say eta here, and we call this y, then actually, y is following normal eta 1. Estimation of theta is same as estimation of eta because it is only a scalar multiple of theta here.

So, optimality criteria for estimating theta are same as those for estimating eta. For example, if I say delta is a minimax estimator for theta, then root n by sigma delta, that will be minimax for eta. If root, if delta is Bayes with respect to for theta, then correspondently, for eta we can obtain root n by sigma delta as Bayes estimator. Similarly, if I say, delta is u m v u E for theta, then root n by sigma delta will be u m v u

E for eta. Therefore, the whole problem has been then reduced to, you can consider the model in a simplified…

(Refer Slide Time: 06:15)



So, it is enough to consider the following estimation problem. We can consider x follows normal theta, 1; theta is any real number L, theta a is equal to theta minus a square. Now, this is a hallmark of all such problems. In any inference problem, we reduce the problem to the simpler model by means of sufficiency. Now, in this particular case, this is extremely straight forward because the loss function is squared error.

In fact, the first thing, that we can always say, that whenever there is a general inferential procedure and it is not based on sufficient statistics, then we can always consider the class of all inferential procedures based on sufficient statistics. They will be more optimal correspondent to or in some sense, they will be better than the original procedures. If the loss function is convex, then we can restrict attention to further non-randomized estimators. See, randomized estimators means that we may choose estimator delta 1 with certain probability, estimator delta 2 with certain probability, etcetera, or we can have a probability distribution over the class of usual estimators.

Now, if the loss function is convex, we can restrict further and we can consider only non-randomized estimators. Same thing happens in the Bayesian estimation also, where we can always restrict attention to the non-randomized Bayes estimators. So, the initial problem of estimating the parameter theta in a sampling from a normal distribution has been reduced to writing, that x follows normal theta, 1; theta belongs to r and L, theta a is equal to theta minus a square.

Now, let us consider here Bayesian estimation. We consider prior distribution for theta as, say normal 0 tau square, so this is our notation; I call it pi tau as the prior distribution. This is normal 0 tau square, so tau square is some fixed quantity because we are assuming some prior distribution. So, this parameter of the prior distribution will be known. Now, in order to calculate the Bayes estimator, we need to calculate the posterior distribution, that is, the conditional distribution of theta given x.

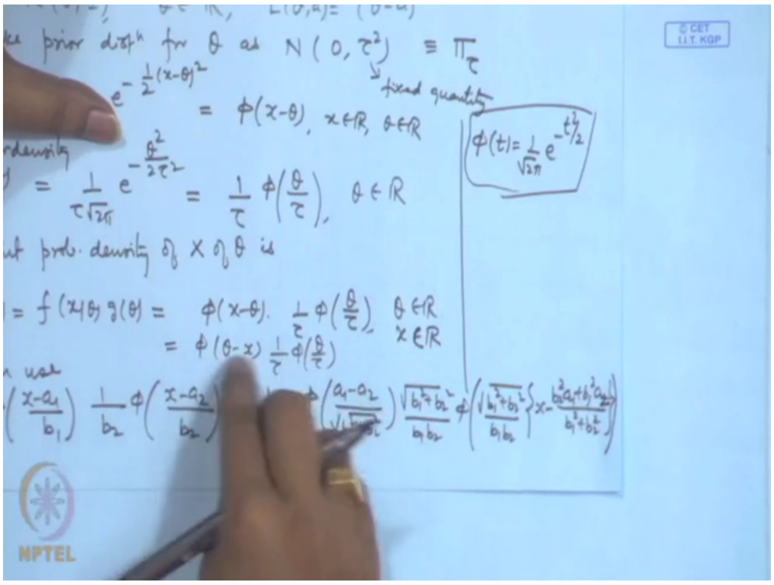
So, we write down all the models here. The distribution of x will be considered as the conditional distribution of x given theta. So, the probability density is 1 by root 2 pi E to the power minus 1 by 2 x minus theta square, we will use the phi notation here, phi of x minus theta, where phi denotes the probability density function of a standard normal random variable. So, we will use this notation, this notation will simplify these calculations. Here, x is any real number and theta is any real number. So, this is treated as the conditional distribution of x given theta, which is actually the original distribution of x, the density of theta, the prior density that is equal to 1 by tau root 2 pi E to the power minus theta square by 2 tau square. Once again we will use the phi notation, then we can write it as 1 by tau phi of theta by tau theta belonging to R. Now, we write down, this is the conditional distribution of x given theta; this is the prior distribution or the marginal distribution of theta.

So, we write down the joint probability, both are continuous and therefore, this is density. We can talk about the joint probability density of x and theta is f star x theta, that is, the product of the conditional distribution into the marginal distribution, so that is equal to phi of x minus theta 1 by tau phi of theta by tau. Here, theta and x, both are on the real line. Now, to calculate the conditional distribution of theta given x, we need the marginal distribution of x and the marginal distribution of x is obtained by integrating this term with respect to theta.

Now, you note this thing, here theta is involved at both the places, therefore we need to resolve this into a single term involved in theta. So, we use the following formula, we can use the following formula something like 1 by b 1 phi of x minus a 1 by b 1 1 by b 2 phi of x minus a 2 by b 2. So, this is 1 by root 2 pi b 1 E to the power minus half x minus a 1 by b 1 square and this is 1 by root 2 pi b 2 E to the power minus half x minus a 2 by b 2 square.

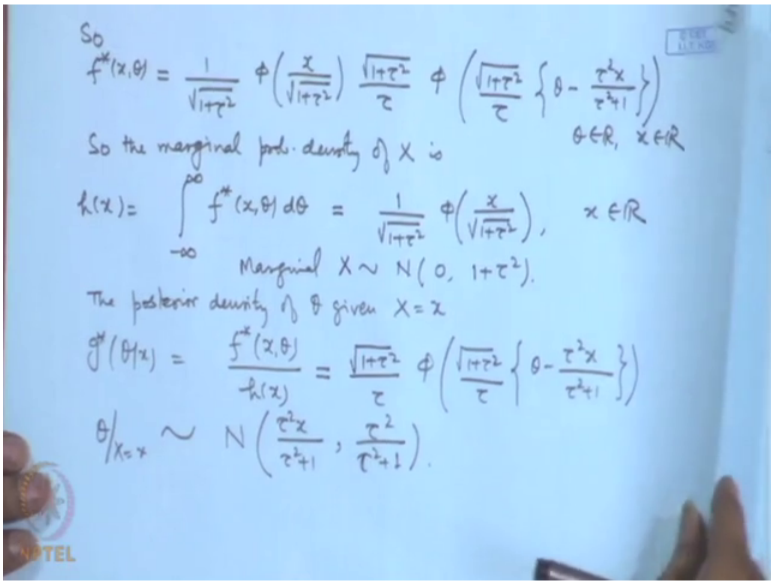
In this one what we do? We collect and make a perfect square for x again and if we resolve this, this turns out to be 1 by root b 1 square plus b 2 square phi of a 1 minus a 2 by root b 1 square plus b 2 square into root b 1 square plus b 2 square by b 1 b 2 phi of root b 1 square plus b 2 square by b 1 b 2 x minus b 2 square a 1 plus b 1 square a 2 by b 1 square plus b 2 square, this is b 1 square a 2.

(Refer Slide Time: 13:10)



Now, you notice here, in this expression x is appearing in both of them here, x is appearing here and this term is free from x. So, if we apply this formula to this one, we can write, this is phi of theta minus X 1 by tau phi of theta by tau. If we compare this with this here, here x is playing the role of theta here and x is a 1, b 1 is 1, then b 2 is tau, a 2 is 0 and b 2 is tau here.

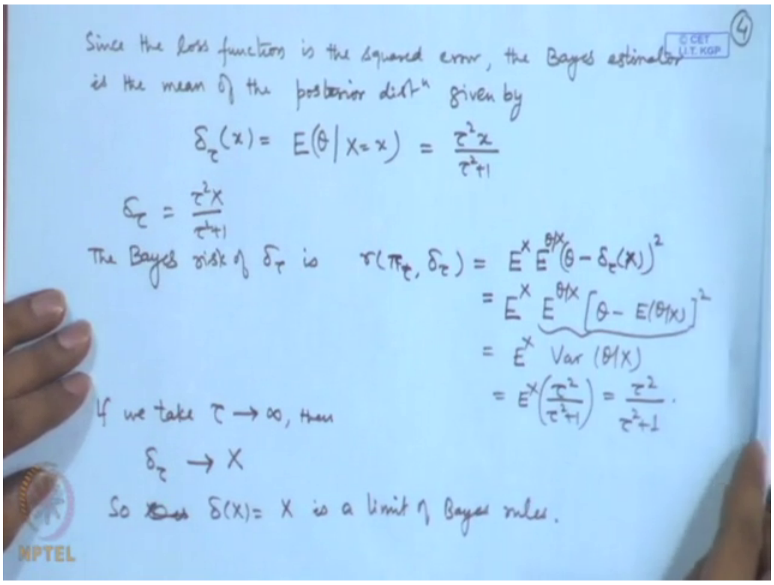
(Refer Slide Time: 13:53)



So, using this we can write, f star x theta can be written as 1 by root 1 plus tau square phi of x divided by root 1 plus tau square root 1 plus tau square by tau phi of root 1 plus tau square by tau theta minus tau square x by tau square plus 1. Here, theta is a real number, x is any real number. So, the marginal probability density of x, let me use a notation, say h x here, that is, integral f star x theta d theta from minus infinity to infinity. So, here if we integrate this term with respect to theta, note here, this is term free from theta and in the 2nd term is nothing, but a probability density of theta with mean mu given by tau square x by tau square plus 1 and sigma given by tau by square root 1 plus tau square. So, integral of this quantity will be 1 and you will be left with 1 by root 1 plus tau square phi of x divided by root 1 plus tau square for x belonging to r, that means, actually, the marginal distribution x is normal with mean 0 and variance 1 plus tau square.

Now, if we consider the conditional distribution, that is the posterior density of theta given x, let me use the notation, g star theta given x, that is, f star x theta divided by h x. If we consider this term divided by this, then you can see easily this density part cancels out and we are left with the remaining density, which is again a normal density. So, what we are saying, that theta given x follows a normal distribution with mean tau square x plus tau square plus 1 and the variance is tau square by tau square plus 1. Now, the loss function is squared error, therefore the Bayes estimator will be nothing, but the mean of the prior distribution.

(Refer Slide Time: 17:18)



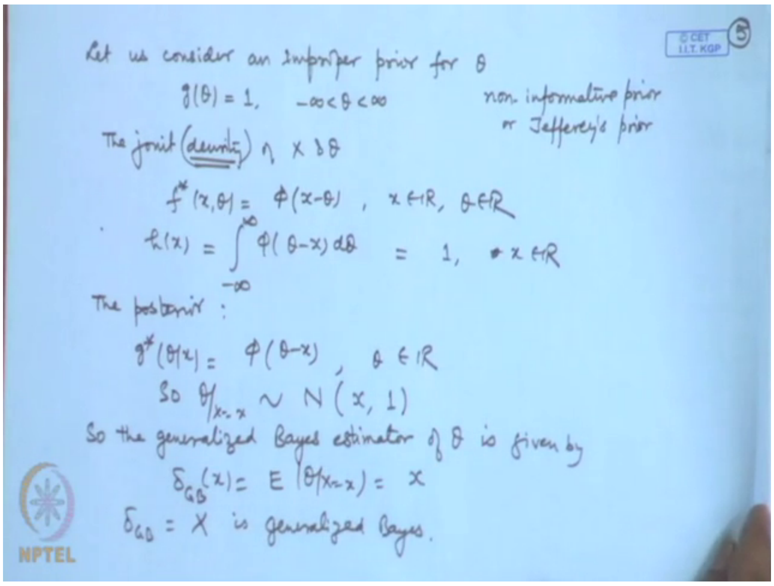
So, since the loss function is the squared error, the Bayes estimator is the mean of the posterior distribution, that is, so let me use the notation delta tau, so that is equal to expectation of theta given x. So, here the mean is tau square x by tau square plus 1. So, the Bayes estimator is simply tau square x by tau square plus 1 delta tau form is given like this and also, we can talk about the Bayes risk. The Bayes risk of delta tau is, r tau, r pi tau delta tau that is equal to expectation of theta minus delta tau. Here, what we can consider? Firstly, I can consider with respect to the posterior distribution and then with respect to the marginal distribution of x.

Now, here if you see, delta tau is nothing, but the mean of this, so we can express it as expectation of theta given x theta minus expectation theta given x, but this is nothing, but the variance of this distribution. So, it is variance of theta given X. Now, variance of theta given x is tau square by tau square plus 1. So, this value turns out to be tau square by tau square plus 1. Now, this is constant, therefore this value will remain as it is.

Now, we note certain facts here. In this, if I consider tau tending to infinity, if we take, say tau tending to infinity, then we observe, that delta tau converges to x. So, x, which is equivalent to X bar, in the original problem if I consider x 1, x 2, x n following normal theta 1, then that is actually a limit of Bayes rule. In the earlier case we have discussed the u m b u E maximum likelihood estimator etcetera, so we have seen is x is actually a maximum likelihood estimator, it is method of moments estimator, it was also shown as the minimum variance unbiased estimator.

Here, what we are showing is, that actually it is a limit of Bayes rules, here let me use a notation, delta x or we can call it, yeah, delta x is a limit of Bayes rules because this convergence is for all x. Now, let me also illustrate the concept of generalized Bayes rules and extended Bayes rule, etcetera.

(Refer Slide Time: 21:20)



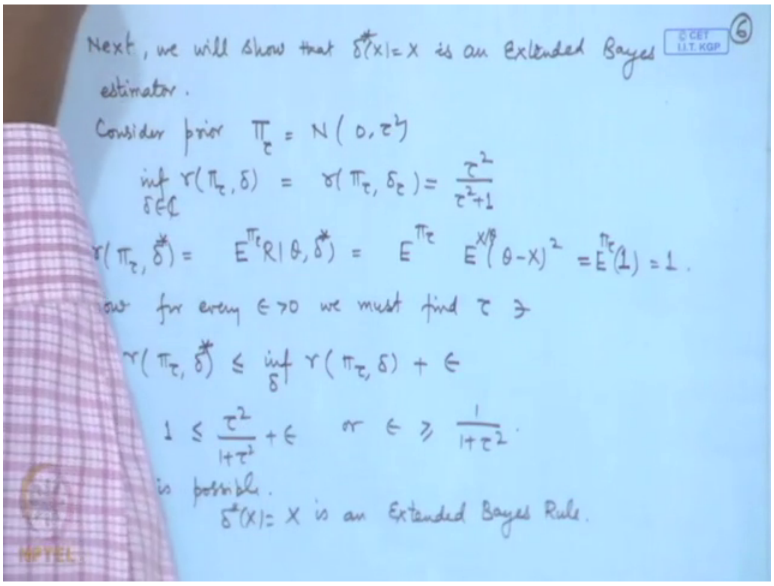
Let us consider an improper prior for theta and we consider g theta is equal to 1 minus infinity less than theta less than infinity. So, if this is nothing, but the Lebesgue measure actually and what this is actually denoting? This is, actually in popular terminology, called a non-informative prior or it is, in this particular case, it is a Jefferey's prior named after Jefferey. And Jefferey proposed that non-informative prior should be taken as a proportional to the Fisher's information.

Now, in this particular case we have calculated the Fisher information. The Fisher information was n r 1, so that was actually a constant, if you remember, for the normal distribution n by sigma square. And then, when we take the reciprocal, we get sigma square by n as the (( )) lower bound for the variance of the unbiased estimator of theta. Here, n is 1 and sigma square is also 1, that we have taken in the model. So, the Fisher information will be actually equal to 1. So, we are taking simply that constant. We can put c here also, it does not matter, this is called non-informative prior and here it is improper prior.

Let us calculate the, apply the same procedure for calculating the base estimators. So, the joint density, which is not actually density, it is improper of x and theta. So, we use a notation, say f star x theta, so f x theta was phi of x minus theta and g theta is 1 here, x and theta, so h x, that is, the marginal distribution of x, that we obtain by integrating this with respect to theta. So, if you integrate this with respect to theta, you get 1 and again this is not a distribution. Surprisingly, if you calculate the posterior, this is a density, let me use a notation g star theta given x, then it is phi of theta minus x divided by 1, here theta is a real number. So, theta given x equal to x, that is a proper distribution, it is normal distribution with mean x and variance 1.

Therefore, we will be able to calculate the Bayes estimator, which will be called the generalized Bayes estimator here. So, the generalized Bayes estimator of theta is, let me call it delta GB, that is equal to x, that is the mean of this x. So, delta GB is equal to x, is generalized Bayes.

(Refer Slide Time: 25:18)



Next, we consider extended Bayes estimation. So, we will show, that delta x is equal to, x is an extended Bayes estimator. Now, for extended Bayes estimator what you have to show, that for every epsilon there exists a prior, for every epsilon there exists a prior, let us call it pi epsilon such that this estimator delta is epsilon based with respect to pi epsilon.

Now, in this, our problem, we already have taken a sequence of prior distributions with respect to which we showed actually, that x is a limit of Bayes rules. So, we will take the same sequence here and let us consider here. So, consider here, prior pi tau as normal 0 tau square and with respect to this infimum Bayes risk, overall estimators is nothing, but the Bayes risk of the Bayes estimator, which we have calculated as tau square by tau square plus 1.

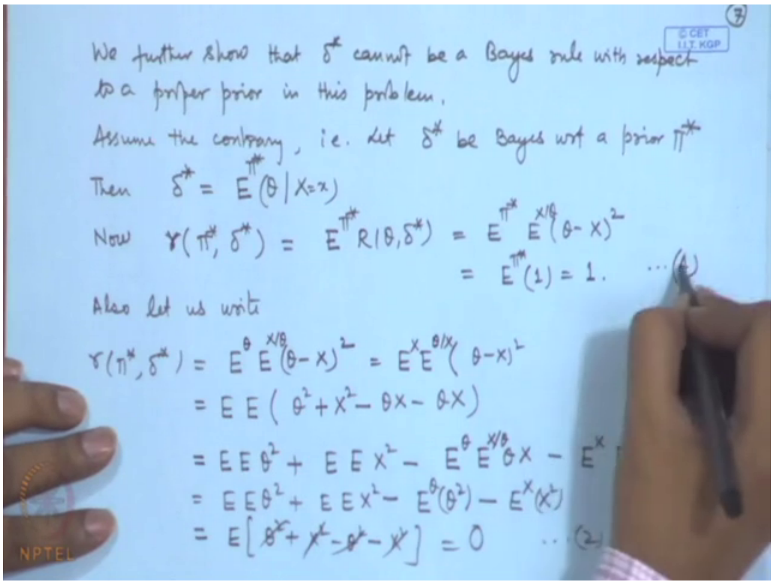
Let us also look at what is the Bayes risk of delta, that is nothing, but expectation of r theta delta with respect to pi tau, but this quantity r theta delta is nothing, but expectation of theta minus x square with respect to the distribution of x, but this is nothing, but the variance of x, that is 1, so it is equal to 1.

Now, if we consider for every epsilon greater than 0, we must find tau such that r pi tau delta is less than or equal to infimum of r pi tau delta plus, so let me use another notation here, say delta star for this. Let me put here this as delta star, this is delta star, this is delta star plus epsilon. Now, this statement left hand side is 1, so 1 less than or equal to tau square by 1 plus tau square plus epsilon or epsilon is greater than or equal to 1 by 1 plus tau square. Now, here we are taking limit as tau tending to infinity, so this is possible, that for every epsilon greater than 0 we can choose a tau such that epsilon is greater than or equal to 1 plus tau square. So, delta star x is equal to x, is an extended Bayes rule.

Now, one important observation, that we should make, we considered the sequence of prior distributions with respect to which we could obtain the Bayes estimators. The form of the Bayes estimators was tau square x by tau square plus 1 here, which is actually a scalar multiple of, you can write it as say c x and here, this c x is always between 0 and 1 because tau square denotes the variance term. So, this will be strictly between 0 to 1; that means, c equal to 1 is not a possibility here.

When we, so the question is, is there any other sequence or any other prior with respect to which x could be Bayes? Actually, we have obtained x as a generalized Bayes estimator, so is there a proper prior also with respect to which x is Bayes? The answer is no. Now, this observation is an important observation and this can be extended to other distributions also or other problems also.

(Refer Slide Time: 30:35)



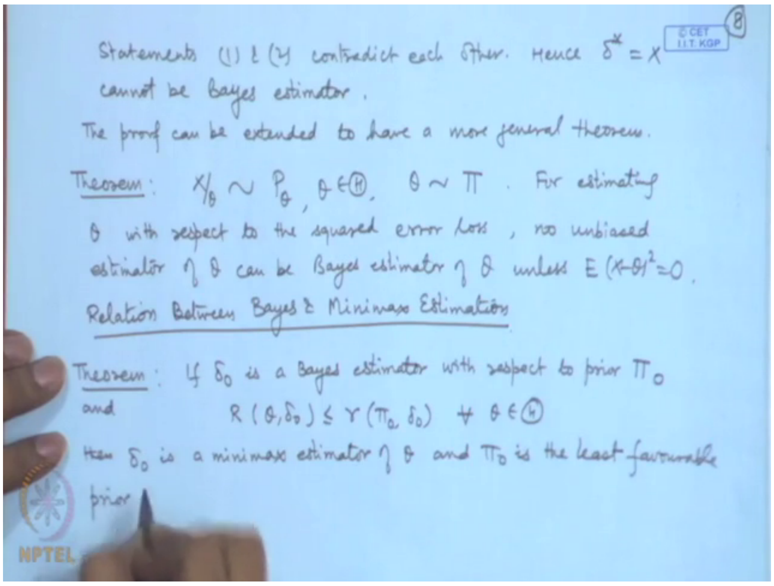
So, let me, firstly show this thing. We further show, that delta star cannot be a Bayes rule with respect to a proper prior in this problem. Let us take, say, assume the contrary, that is, let delta star be Bayes with respect to a prior, say pi star. If that is so, then in the squared error loss function you are getting delta x is equal to expectation of theta given x with respect to this particular prior.

So, let us write down r pi delta star. Of course, this is equal to expectation of r theta delta star with respect to pi star. Now, this is expectation of, expectation of theta minus x square, this is with respect to the conditional distribution of x given theta, that is, expectation pi star. This is nothing, but the variance of x, that is, 1.

Let us also consider r pi star delta star as expectation expectation theta minus x square, this you are writing as expectation of expectation. Now, when I write two expectations, then it can be in any order, like it can be x given theta, then with respect to theta, or it could be theta given x and then with respect to x, it can be in either form. Now, let us write this, so this is equal to expectation of expectation theta square plus x square minus theta x minus theta x. I am writing this term two times, so the first one is expectation of expectation theta square plus expectation of expectation x square. Once again I can take any order, when we, I take the repeated expectations minus expectation of expectation theta X. Now, here what I do? I firstly take with respect to x, then with respect to theta and in the 2nd case I reverse this process, I take x here and I take theta given x.

Now, if I look at this term here, this term expectation of x given theta, this will give me theta x, so this will become theta square. And if I look at this one here, expectation of x, then expectation of theta given x of theta, that is x, so that is becoming x square. Now, in this expression, you can consider this as simply expectation of theta square plus x square minus theta square minus x square. So, these terms vanish, that is equal to 0, so this 1 and 2 are, this 1 and 2 are in contradiction.

(Refer Slide Time: 35:07)



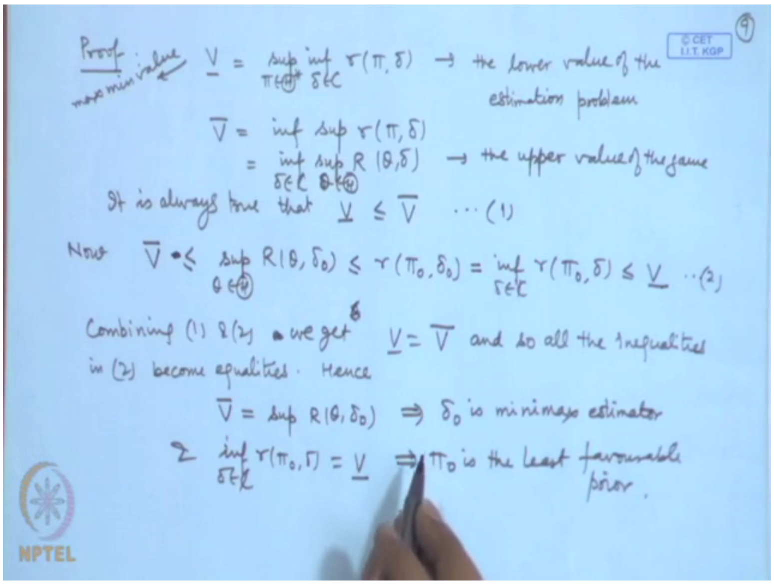
Statements 1 and 2 contradict each other, hence delta star is equal to x cannot be Bayes. Now, what is the property that we have used in arriving at a contradiction? The property that we have used in arriving this contradiction is that x is unbiased for theta and the loss function is squared error. Therefore, the Bayes estimator is the mean of the posterior distribution. Now, whenever this property is true, this becomes a general result.

So, we state the following general result can be extended to have a more general theorem. So, x given theta, that follows p theta, theta has a prior distribution. Then, for estimating, say theta with respect to the squared error loss, no unbiased estimator of theta can be Bayes estimator of theta, unless of course, expectation of x minus theta square is 0, which is of course, impossible.

Next, we establish a connection between Bayesian and minimax estimators. In fact, the Bayes estimators are used to prove the minimaxity of certain estimators, so Bayesian approach is the approach that is being used here.

We have the following theorem here, if delta naught is a Bayes estimator with respect to, say, prior pi naught and r theta delta naught is less than or equal to r pi naught delta naught for all theta, then delta naught is a minimax estimator of theta and pi naught is the least favorable prior.

(Refer Slide Time: 39:16)



Let us look at the proof of this statement. So, you remember, that in the previous lecture I introduced the minimax value or the upper value of the game and maxmin value, that is, the lower value of the game. The way we had defined the lower value and the upper value, the lower value was defined as supremum infimum r pi delta, where infimum is over all the estimators and supremum is over all the prior distributions, that was the lower value of the estimation problem.

This comes from the game theory, so generally we use the terminology lower value of the game. However, since we are dealing with the estimation here, I will write it as the lower value of the estimation problem. And similarly, the and this is also called the maxmin value and similarly, we have the upper value of the game, that is, infimum supremum r pi delta, it is also equal to infimum supremum r theta delta, where the supremum is over all the theta and infimum is over all the estimators here. This is called the upper value of the game and it is always true, that v lower bar is less than or equal to V upper bar. Now, we consider here, say V upper bar, so V upper bar will be certainly less than or equal to supremum of r theta delta naught because this is the minimax value.

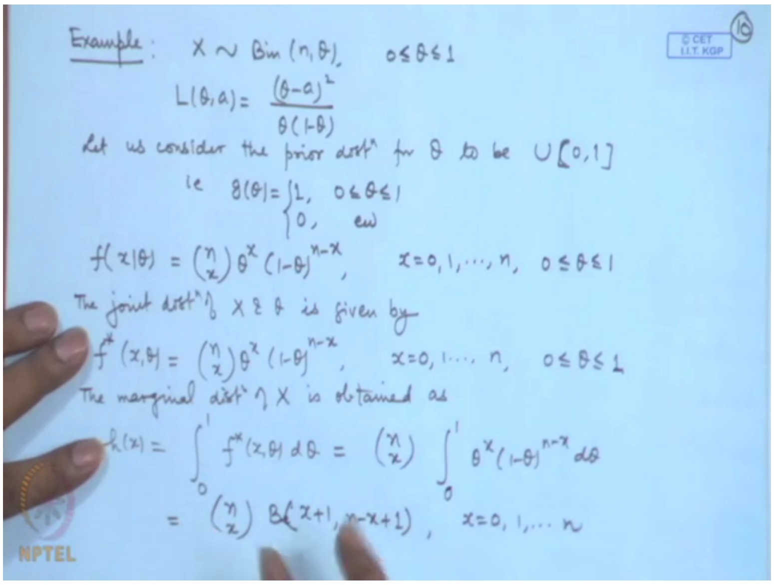
So, this is for a particular estimator delta naught, so this is going to be more than or equal to V upper bar.

Now, we have assumed here, that r theta delta naught is always less than or equal to r pi naught delta naught, so the supremum is also going to be less than or equal to that, that it is less than or equal to r pi naught delta naught. Now, here delta naught is a Bayes estimator with respect to pi naught, so this is equal to infimum of r pi naught delta over all the estimators. Now, this is less than or equal to the lower value of the game and therefore, what we get here? V upper bar is less than or equal to v lower bar.

So, combining 1 and 2, we get v lower bar is equal to V upper bar, but that would mean all the inequalities, all the inequalities in 2 becomes equalities. Hence, now if you put V upper bar, that is equal to supremum of r theta delta naught, this implies delta naught is minimax estimator. And similarly, you will get infimum of r pi naught delta is equal to v lower bar, which means that pi naught is the least favorable prior.

Let me explain this theorem with respect to one example here. So, what we need in order to apply the theorem here, that we need a prior and we have a Bayes estimator with respect to that and then, we have the risk relationship between the risk of delta naught with the Bayes risk here. So, if we have that, then the estimator becomes minimax and the corresponding prior will become the least favorable prior.

(Refer Slide Time: 44:09)



So, let me explain it through one example here. Let us take, say x follows binomial n theta, here the parameter space is the interval 0 to 1 and we take the loss function to be theta minus a square divided by theta into 1 minus theta. Now, let us consider the prior distribution for theta to be uniform 0 1, uniform distribution on the interval 0 to 1, that is, g theta is equal to 1 for 0 less than or equal to theta less than or equal to 1 and it is 0 otherwise.

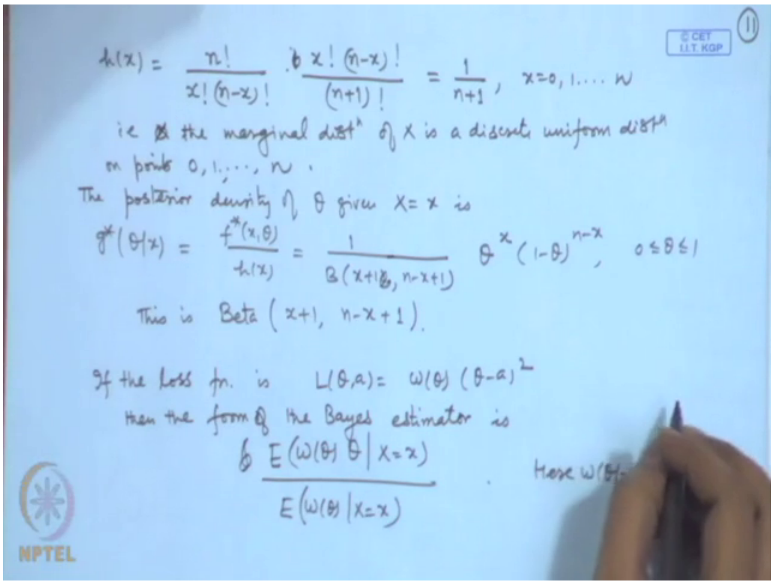
Now, here the distribution of x, that is, n c x theta to the power x into 1 minus theta to the power n minus x, where x can take values 0, 1 to n and theta is any value between 0 and 1. So, now in this case you notice, here the distribution of x is discreet and the distribution of theta is continuous. So, here it is probability mass function and this is probability density function.

Nevertheless, we write in the same way, the joint distribution of x and theta is given by, so I will not call it a density function because this part is not a density function, f star x theta, that is equal to n c x theta to the power x into 1 minus theta to the power x into 1 minus theta to the power n minus x for x is equal to 0, 1 to n and 0 less than or equal to theta less than or equal to 1.

Note here, that if I sum with respect to x, then I will get it as 1 and then, if I integrate this respect to theta, I will get 1. So, this is the proper probability distribution, but this is known as a mixture kind of thing because here x is a discreet variable, theta is a continuous variable. So, the marginal distribution of X, now since it is continuous in theta we will integrate it with respect to theta, so it is integral of f star x theta d theta from 0 to 1.

Now, in the integration you see, here n c x will come out, you get integral from 0 to 1 theta to the power X 1 minus theta to the power n minus x d theta, but this is nothing, but a beta function. So, we get the value as n c x beta of x plus 1 n minus x plus 1 for x is equal to 0, 1 to n. Actually, it is a proper probability mass function, if you want to see we can actually simplify this and check it.

(Refer Slide Time: 47:48)



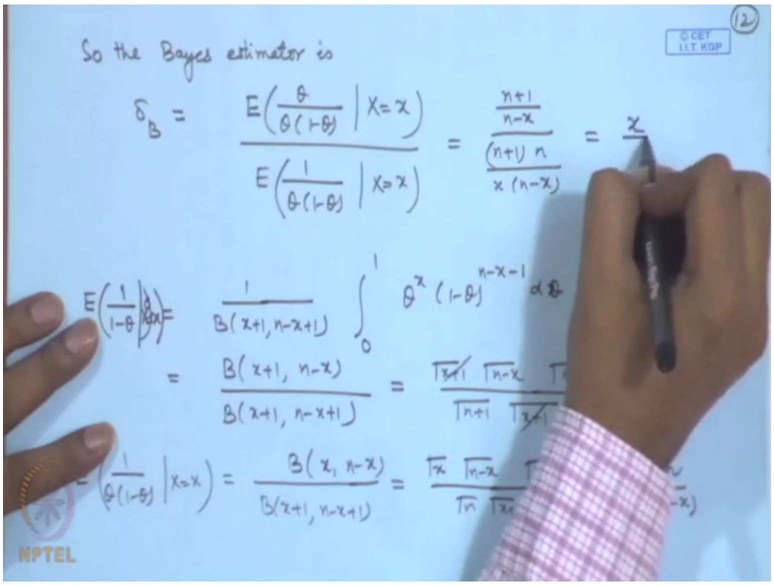
h x, if you simplify, n c x is n factorial divided by x factorial n minus x factorial. Then, beta term, it is gamma x plus 1, that is x factorial, then gamma n minus x plus 1, that is, n minus x factorial divided by gamma n plus 2, that is, n plus 1 factorial. So, this after simplification becomes simply 1 by n plus 1 x is equal to 0, 1 to n, that is, actually you are getting x. The marginal distribution of x is a discreet uniform distribution on points 0, 1 to n, that is, each one is having equal probability 1 by n plus 1.

Now, we utilize this to calculate the posterior distribution of theta given X. Now, once again notice, here in theta the random variable is continuous, therefore the posterior distribution will be actually a continuous distribution. So, we will have a density function, the posterior density of theta given x, then that is obtained as g star of theta given x, that is equal to f star x theta divided by h x, that is equal to simply 1 by beta of x plus 1 n minus x plus 1 theta to the power x into 1 minus theta to the power n minus x.

Actually, this is nothing, but a beta distribution, this is beta distribution with parameters x plus 1 and n minus x plus 1. Now, the loss function, we have considered a weighted squared error loss function, if you remember, in the previous lecture I have given the formula, if the loss function is say w theta into theta minus a square, then the form of the Bayes estimator is expectation of w theta theta given x divided by expectation of w theta given x is equal to x.

Now, in this case w theta is, in this particular case w theta is 1 by theta into 1 minus theta 1 by theta into 1 minus theta.

(Refer Slide Time: 51:39)



So, if we calculate, now so the Bayes estimator, let me call it delta b, that is equal to expectation of theta divided by theta into 1 minus theta given x divided by expectation of 1 by theta into 1 minus theta given x is equal to x, that is expectation of 1 by 1 minus theta.

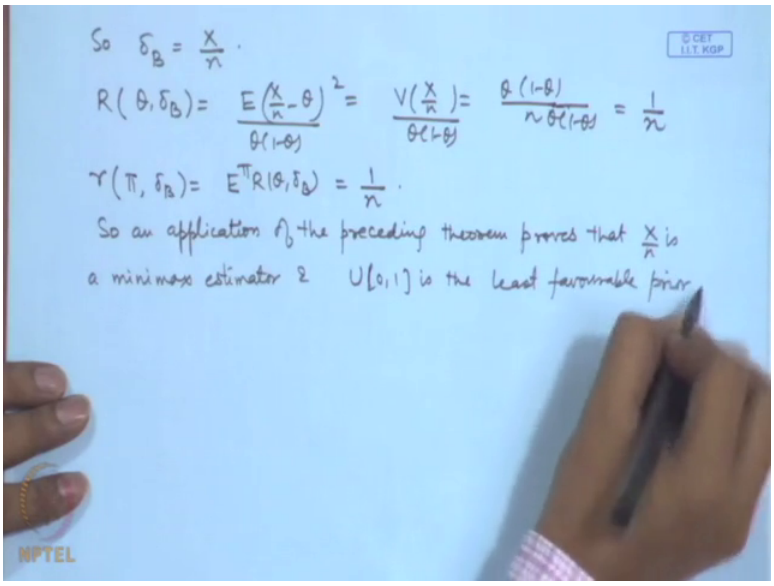
Now, these quantities we can evaluate the posterior distribution is obtained like this. So, from here if we calculate expectation of 1 by 1 minus theta, that is equal to 1 by beta x plus 1 n minus x plus 1 integral theta to the power x 1 minus theta to the power n minus x minus 1 d theta from 0 to 1. This is conditional expectation with respect to X. Now, this is nothing, but a beta function, so get beta x plus 1 n minus x divided by beta x plus 1 n minus x plus 1.

So, this can be simplified, gamma x plus 1 gamma n minus x divided by gamma n plus 1 and in the denominator you will get gamma x plus 1 gamma n minus x plus 1 and gamma n plus 2. So, these terms get cancelled out and we get here n plus 1 divided by n minus x. In a similar way, if we calculate expectation of 1 by theta into 1 minus theta given x is equal to x, then that will be equal to 1 by beta x plus 1 n minus x plus 1 and the numerator will get beta X n minus x.

So, after simplification this will become equal to gamma x gamma n minus x by gamma n, then gamma x plus 1 gamma n minus x plus 1 gamma n plus 2. Once again, if you simplify these terms we get n plus 1 into n divided by x and n minus x. So, if we substitute these quantities here, we get n plus 1 by n minus x divided by n plus 1 into n divided by x into n minus x that is equal to x by n.

So, our minimum variance and biased estimator or the maximum likelihood estimator is actually the Bayes estimator with respect to the uniform prior, if the loss function is taken to be theta minus a square by theta into 1 minus theta.

(Refer Slide Time: 55:17)



So, so the Bayes estimator is actually x by n. Now, let us look at the risk of this, so risk of the Bayes estimator, that is, expectation of x by n minus theta square, that is nothing, but variance of x by n, that is equal to theta into 1 minus theta by n. And here, we have to divide by theta into 1 minus theta divided by theta into 1 minus theta. So, this turns out to be simply 1 by n.

So, if I look at r pi that prior, which we are using uniform 0 theta, I am using a notation pi say delta b, that will be equal to expectation of r theta delta b with respect to pi. So, it is again equal to 1 by n.

So, if we look at the theorem, that I gave here in this result, we are having r theta delta naught here equal to 1 by n r pi naught delta naught is also equal to 1 by n. So, this proves that. So, an application of the preceding theorem proves that x by n is a minimax estimator and uniform 0, 1 is the least favorable prior.

In the next lectures I will be discussing further applications of the Bayesian estimation, in fact, how to use a sequence to prove the minimaxity. In this case we could use only one prior and we could prove the minimaxity. In the next one we will also introduce an equalizer estimator, etcetera and how to use that to prove the minimaxity of estimators, so that I will be taking up in the next class.