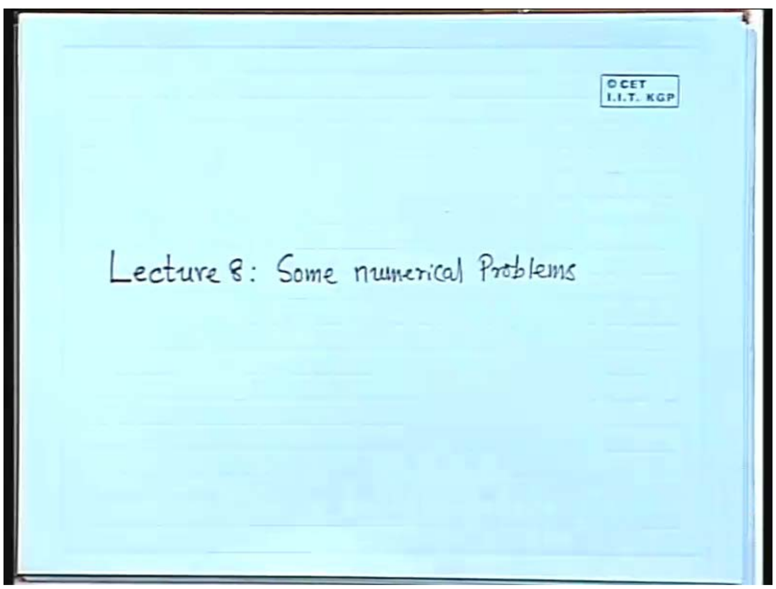
Estimation of Signals and Systems

Prof. S. Mukhopadhyay
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 08
Some Numerical Problems

For the last few classes, we were having I mean theoretical discussions, that I wanted to progress. And anyway we are having a tutorial in the evening, so I thought that maybe we can start some numerical.

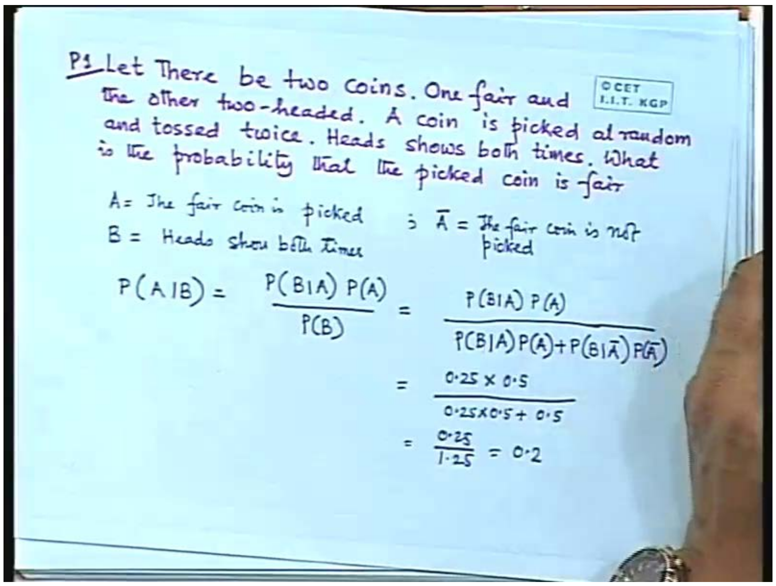
(Refer Slide Time: 00:55)



That also you know saves me, the time to prepare another set of notes, so easy I mean I had to prepare for the tutorial probably more than what you have to prepare for the tutorial. So okay so I just I did, the same thing as probably what you are doing. So I picked up some problems; some of them I could solve, some of them I found difficult to solve. So I said so I just like everybody else; I thought okay i will get back to these hard problems later, just like everybody else does. So I have solved some, so I thought that I can share what I have solved with you, right.

So the first problem, this problem was somewhat simple. This is I mean just see whether you can make out, what is happening here. So the problem says, that, let there be two coins, right.

(Refer Slide Time: 02:08)



One is fair, fair means that the probability of getting a head and the probability of getting a tail is same and is equal to point five, let's call it a fair coin. And the other is two headed; it is like that Sholay movie coin, you know so both sides are head, right. So a coin is picked at random and tossed twice; so there are two events here, one is one is picking up the coin, among the two and the other is tossing it twice, right. Heads shows both times. So when you toss them twice in a given experiment, you got heads both times. Given that head showed both times; find the probability that that the picked coin is fair, it is the other way, right.

So.. so this is this is the case, where you have, what is the probability? What is the general probability, that the picked coin is fair? Which is the half? Because, you are picking up one of the two coins and you are picking it up in random. So the probability of picking up the fair coin is equal, to the picking probability of picking up the two headed coin. So the a priori probability of picking up the fair coin is point five, but but you are given the now you are given a fact. It is not the a priori probability that you are calculating; you are you are given a prior information, that both times it shown heads. Now what is your probability, that the picked up coin is fair? So the probability is going to be different, right. So so we will apply, what will apply base law, right. So let us first mark the events, okay. So let A be the event that, the fair coin is picked. And let B be the event that, heads show both times, all right? So what are we trying to calculate? We are trying to calculate probability of A given B; this is what we have to calculate, right? So what is probability of A given B, it is probability of B given A into probability of A divided by probability of B, right.

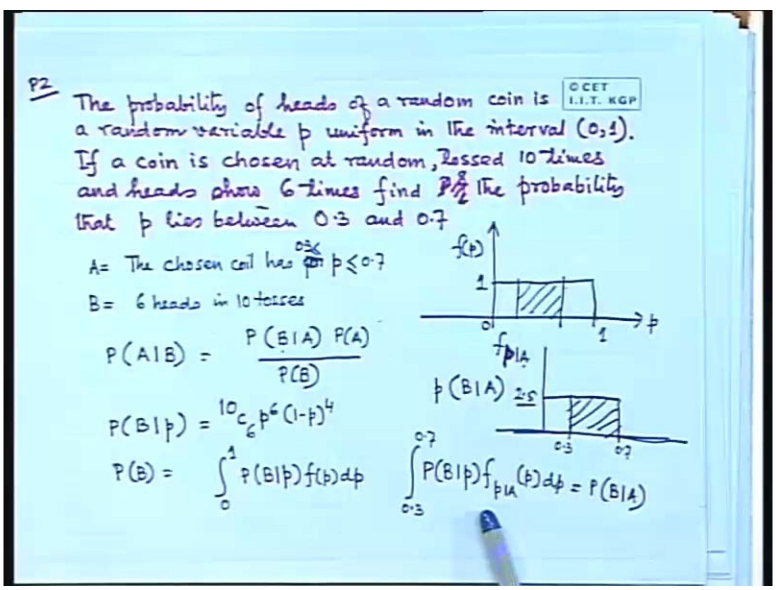
Now what is probability of B? Why? of a first probability the head comes up again second time as a that is if the fair coin is picked probability of B means yes is exclusive no no no probability of B event you have defined head show both times. Right! Head show both times, what is the probability irrespective of each coin is picked? So so irrespective of which coin is picked, head show both times probability is not half. We is is not the if the fair coin is picked; then the probability is point two five, not otherwise. So this is actually probability of B given, let us say suppose A bar is the event that the fair coin is not picked, I am having difficulties writing, so please excuse as such my writing is not great into probability of A plus probability of B given A bar into probability of A bar. This is the over all probability of B, irrespective of which one is chosen. So either A is chosen, either A occurs or A bar occurs. So if A occurs, then what is the probability? And then if A bar occurs then what is the probability? Taking them together, you will get the probability that either A or A bar occurs.

And then then what is the then, top remains the same, right. Now you could value, so this is this is what you are saying point two five. Probability of B given A; that is point two five into, what is probability of A? Half divided by this is this this is the same, so it is point two five into point five plus. What is probability of B given a bar? It is one. So it is just point five. So it is zero point two five by one point two five is equal to zero point two. So what you are effectively saying is that; if you see both heads coming up then there is a much higher probability that, the coin the you have picked is actually the two headed one. Because the probability that, you have picked a fair coin and you got two heads is actually rather low, right?

So this is one of the this is the see otherwise, even if you do not apply base law then this problem appears somewhat you know confusing, that then how do you you you asking kind of the the reverse question. You know we are more use to the fact that, if the coin is fair what is the probability of getting two heads? We are we are generally more use to this, that is that is more simple to imagine, but here we are asking the reverse question. Given two heads, what is the probability that the coin is fair? And using base law, now you have inverted the conditionality, right? So I have this is clear? This probably you could anyway do it, it was a simple problem.

Now let me complicate this problem, a little farther, okay? This this also nice problem, which uses the same kind of base rule, but for a different application. This says that, the probability of heads of a random coin is a random variable p? Now we are having many coins, you know some of them are unfair coins.

(Refer Slide Time: 09:48)



What is mean by unfair coin? That means, the the probability of getting one event is actually larger than the other, right. Some sometimes use by you know gamblers; gamblers will use this kind of thing, so that they will change the mass distribution of something that and they will toss it in a clever way, such that always as the probability of a tail coming up will be higher, okay. Such a coin is called, an unfair coin. So in general; you can imagine it is very difficult to have coins which will have point nine and point one probabilities and still look like, a coin that is rather difficult to construct probably. But for the sake of you know mathematics, let us imagine that there are there could be a coin, in which this probability of having a head can can vary anywhere between zero and one. So it is a random variable, you you pick a coin you get a probability, that sort of thing. If a coin is chosen at random, now you are choosing a coin at random and toss ten times, and heads show six times. Find the probability that this value p; for the chosen coin lies between point three and point seven, that is a little more complicated, is not it?

So so so how do you do it? Again will will will we again so, what is the what is the again first mark events? Okay. So first mark events events is A is the chosen coin has p, this is the first event. And the second, event is six heads in ten tosses, ..right. Now, what are we supposed to find out? We are supposed to find out, probability of A given B. We got we are suppose to find out, in the previous problem this p function was were actually discrete. It was either one or point five, p cannot take any other value in the previous problem. Now in this problem p is at a p itself is a random variable variable p is have point three to point seven what is the probability that, both are equal both are less than equal to one we should do in only quantity node p will be p will lie between, point three and point seven. Suppose you have made ten tosses; you got six heads now for looking at this event, what can you say what is the probability that that that the coin that you have tossed with has the probability between point three and point seven? That is the question in which is asked of course because if, because you are choosing a random coin and its probability can vary between zero and one.

So we are we are suppose to find out probability of A given B, now what is that? Obviously it will be it will be the same thing; it will be probability of this this will apply, will be a probability of B given A into probability of A divide by probability of B, there is no doubt about that, right. Now the question is how do you evaluate these things? The evaluation is not that simple. Now so first of all let us find out; what is the probability of B given p, if we are told if we are told that this coin that you have chosen has probability p, then what is the probability of B? It is Burghley trials, so it is ten c six p to the power six one minus p to the power four, right? This is the probability that that will be six heads; probability of six heads is p to the power six, they are all independent tosses. And probability of rest four is one minus p to the power four and you have ten c six number of them, that is the heads and tails can occur in any order. So so how many such arrangements are possible? standard.

So this is probability of B given p; so from this we can say that probability of B, you can you can immediately calculate. Probability of B is nothing but zero to one, probability of zero to one because, f p d p for all values of p now you are taking average, correct. See what are you doing, you have to you have to imagine, so probability is like you know we have a we have we have we have frequency interpretation of probability. That that is that is the interpretation which people use to give till you know, this major theory came up. So people use to say that; if I do n number of tosses, how many times I get heads, how many times I get tail?, ratio is probability. That is to be the interpretation, the people is to give. So if you think; if you think in that way, what are you doing? You want to find out the probability that that suppose, there are n number of coins heap together. So you are taking you are randomly taking one coin; doing very large number of tosses, then taking another coin doing very large number of tosses, doing another coin. In this way for each, you are picking up a coin and you are doing a large number of tosses with that coin.

So if you see just the probability of this tosses for each coin, then you get what is what you get here. Pick up any one and do a large number of tosses; and see the ratio of heads and tails, that ratio is given by this. Now if you sum up, what when you are saying probability of B, you are saying that now I will not sum up for a given coin, I will sum up over all coins. So then I have to so then I will sum up over all the experiments, that I of of coin picking also, right. Suppose, if you do thousands of for each coin; if you do thousand tosses and if you pick thousand coins, so now you are you will calculate this this number of heads and tails, over ten to the power six numbers of tosses. So obviously now you let to know; if you when you when you pick them up thousand times, how many times I mean what will be the frequency of a given p coming up that will be given by this f p d p naturally. That is why you you have to integrate, are you getting the idea?

So this is probability of B, now the question is what is probability of B given A? Probability of A is known, probability of probability of A is probability of the chosen coin has, see this p has a is a random variable p uniform in the interval, so so the probability of choosing a coin with this where p lies between this is simply point four. What is f p is this this is fp, this is p, zero and one. So suppose this is one, so this is f p. Uniform in the interval zero to one, this is zero, this is f p. So this whole area is one. So then between point three and point seven; say point three is somewhere here let us say, and point seven is somewhere here let us say, so the between point three and point seven, it is going to be point four naturally. So p A is point four, there is no doubt about it. p B also now we can calculate, we can simply put this here that is reasonably simple to do. Question is tricky part is what is p probability of B given A? This is the tricky part.

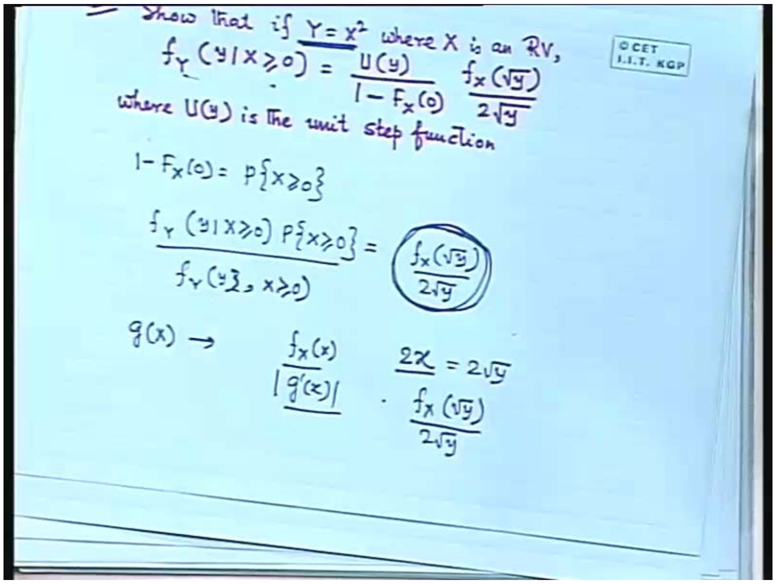
Now the probability of B given A is not just this, between point three and point seven. That is not correct, that is it do not do not just change the integral limits here, because now you are given that the coin has probability between point three and point seven. So now imagine that you have, somebody has in the form of big lump of coins, which you had the big heap, you have separated out; somebody has separated out only those coins, which had point three and point seven and then made a made a smaller heap. And then he is asking you that, now you why do not you do your tosses from this small heap and then find out what is probability of B? That is still uncertainty about p; because all coins in the heap does not have the same p but, it is now given to you that it is between point three and point seven. You have to choose from the smaller heap now.

So if you have to do that now, obviously your probability density function will change. And what will it be now? Your probability of probability density function, that is the f of B given A is now different or rather; this is not f of B given A, this is.. f of p given A. what is the probability density function of p; but given that p, lies between point three and point seven. This is only the ordinary probability density function of p, nothing is given, it could be anywhere. But now, you are what is the probability density function of p given that, they all lie between point three and point seven. That is if you can you characterize, the probability density function of that small heap. So now it will be everywhere; so now it will exist only between point three and point seven of p, everywhere else it will be zero. Because you do not have any other coin and since the whole thing is uniform, so so so here also it will be uniform, in any interval it will be uniform. So this so now this is going to be; again like this but this level will now be two point five, because the area must come to one, correct. So this is your probability density function of p, given A, right.

So now if you have if you if you want to find out, probability of B given A, then you have to find out simply, probability of B given p f of p given A p d p. This integral can obviously since; this is zero everywhere else, so this integral you have to integrate between point three and point seven. This is equal to probability of B given A. Remember that, when we did that longevity example, what did we find out? We said that what is the the probability that, if if somebody‘s age is greater than sixty five, what is the expected value of his leaving? right. So you find the found that, the expected value now lies between the remaining intervals.

People leave between zero to hundred; so somebody has already leaved sixty five, so his probability of leaving is again uniformly distributed, between sixty five and and an an hundred. So the mean came at came in the middle as, eighty two point five, that that is what we did in the class. So here also the same thing is happening. So now this is probability of B given A now, you can do the problem. Probability of B given A is given, this one you know, this one is nothing but two point five. This is the this you know and you have to integrate between point three and point seven; and probability of A is what is probability of A....? Point four. So now we are know, you know all the terms, so you can find out. Is this clear? This is the way you applied, base rule all right? Now you have another one, this is simple, simpler! Let us say, show that if Y is x square that is now we are defining a function of a random variable, where x is an RV.

(Refer Slide Time: 24:21)



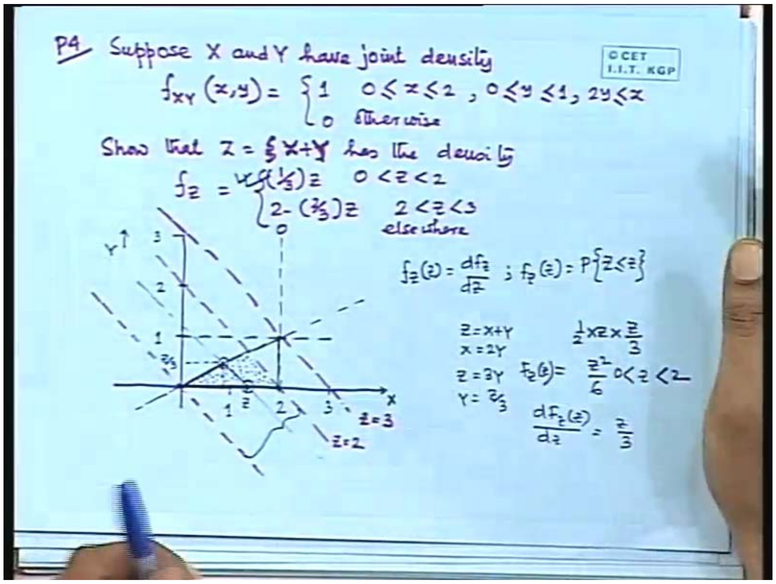
F Y such that, f Y of y given, x is greater than equal to zero, is this. What is U y; it is the unit step function that is, the function is zero for all negative values and it is one for all positive values, right? So, how do you do this? The certain things that you have to know; first of all you have to find out that, with see some some some parts whether we can recognize them, why they have come? Firstly, why why this U y has come? Why this unit step function should be there? Because; because y cannot be negative, so there is there is no question of when you define as y is equal to x square, there is no question of getting a negative y for real x, correct? So therefore, the the the probability density of y must be must be zero for all y less than zero, so that is immediately achieved when you have U y. So you write it only for positive values of y; without indicating as such, and then you put this U y so whatever even if you put a y is equal to negative here.

See if you put y equal to negative will get abstract results but suppose, this was this even this was some other function, you can put y equal to negative, does not matter this will still make it zero. So you do not have to specifically write that, y positive or anything like that that, it just for that reason. So so that is why they put U y because y cannot be negative. Why why this? What is f x zero? F x zero is a probability that that, x is less than zero. We are not bothering about those, you know those finicky I mean continuity condition whether less than or less than equal to, let’s ignore them. They will their so, what what is one minus f x zero? It is a probability that x is greater than or equal to zero. So this part; we should recognize this, one minus F x zero is nothing but probability, that x is greater than or equal to zero.

So what are you having here; so if you multiplied this side, you are having f Y, y given x greater than or equal to zero into probability that x is greater than or equal to zero.... If you multiply that, it is it is equivalent to prove that, that same thing is equal to f x of root y by two root y. U y I am for the time being ignored. It is the same of same same, what is this? This is nothing but, f y into y. This part will go now; once you have multiplied it by probability then, it will be f y, is it not? No. It will be f y and x greater than or equal to zero, right, right or not...right. Now anyway you have to consider x greater than or equal to zero only because, x less than or equal to zero will anyway go out. So we are considering only positive values of x, so this this condition is automatically satisfied. Now, what is this? Remember, that we had if we have a function g x, then its probability density function was what? It was f x by g dashed x, remember? Is it not? So so what is g dashed x here, g dashed x is two y in other two x is equal to two of root y.

And what is f x f x of x is equal to f x of just; this is this is expressed in terms of y since, on the left hand side you have Y, the same thing is expressed in terms of y rather than in terms of x. So what is f x of x? It is written as f x of root y and at the bottom you have twice x, that is written as twice of root y. Just simple application of the formula, only little bit twisted because of this pon terms, understood? Okay? So this is also not. Now we have another one.

(Refer Slide Time: 30:38)



First read the problem. Suppose x and y have join density, f x y, x y is characterized like this, so what is how does it look like? Looks like this. So suppose this is one, this is two, this is three. Similarly this is one and just this is two and this is three. This is the x y plane, okay. This is y, this is x. So zero x and the x is between zero and two, so x is in this region. y is between zero and one, so y is in this region. And twice y and x is less than or equal to twice y, so what is the x equal to twice y line? x equal to twice y line is here, is it not? So, and what is x is greater than or equal to twice y, that is means it is in this region. So the area in which f x y is non-zero, turns out to be this area. This is the area in which f x y is non-zero, inside this triangle, right. First, agree to this.

The value is given as one that means, it is uniformly distributed over this interval for this area. And what is the area of the triangle? It is one, half base into altitude. So it is one, if it was anything then, there was corresponding scale factor because, over all probability must be one. So this is tallying, so this is where you haves effects. Now you are defining a function, called z is equal to z is equal to x plus y. So what is that function? That function is this line and the z is equal to this z is equal to x plus y line. This is actually; the z is equal to two line, you can have various kinds of lines. For example, z is equal to zero line, will be this, this is the z is equal to zero line. Similarly, this is the z is equal to three line, right. Read. So now the question is, what is we have to find out what is f z? right. So what is f z; f z is d f z by d z, derivative of the distribution function. What is the derivative of the, what is f z? Probability that z is less than or equal to z, right. So for example, if you are given small z is equal to two, what is that probability? The probability will be this line, this area. Because on this side z is less than or equal to two; now x and y can take values only in this triangle, they do not take values, elsewhere. So the values of x and y which will which which are possible to occur and which will make z less than two, can you see this dots? Is it visible on the TV?

So so this is the area; for example, if you chose that is equal to two, so is so f z z will be the area of this triangle, right. Now you have to question is that, you have to find out this area for an arbitrary z, that is that is that is the thing. So obviously you will find that, if you now you obviously, if z is less than zero; probability is going to be zero because, it exists only in this triangle. That that is why they have said that, z must be with either between zero and two and between two and three; that means z can be only between zero and three, it is zero elsewhere. So zero zero there is no probability, beyond three also there is no probability. Rather rather beyond three a f z is zero, but if you integrate capital F Z will become one, correct. So now you have to find out that, as you swap z from zero to three, how does this density change? That is, how the the area inside the triangle swapping is changed? that what is the rate of swapping of the area inside the triangle, that is f z, right.

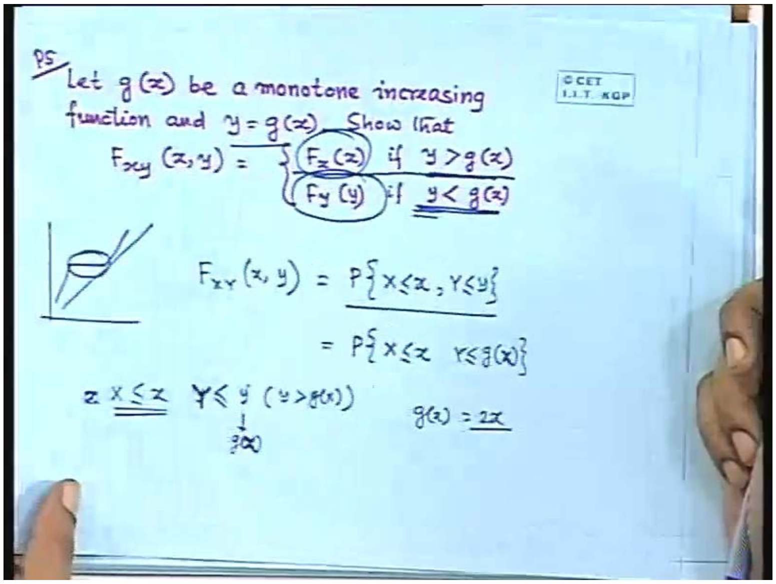
Step by step we have to visualize, in fact I mean nothing, I mean I I get the in yesterday evening. So I am just sort of loudly telling you, what was my thinking process. This problems are not solved in the book. So so now what you do, so now you find out, now now you discovered; that the way the area changes from zero to two, as it shifts from zero to two is..that is in this zone of z, the area increases in a given way. And in the from two to three, you can see it from the picture itself. The area changes in a different way, can you see that? That is why you have you have different f z's, in in the region zero to two and two to three. Now it is not difficult to find out, how this area is swiped. That is very simple. Now for example, we can do do you want to do one of them; we can do it after all suppose, we take from point here, an arbitrary point this one, okay. Find out what is this point? What is this point? It is in the intersection of this point, what is what is the coordinate of x? It is z because, z is equal to x plus y and y is zero. So so this point is z, what is this point? This point is z minus one; how z minus one? It is z is equal to x plus y is is it is intersects and and x is equal to twice y. So z is equal to three y, so y is equal to z by three. So this is y, so this is z by three. So this is z this is z by three, right.

So then, what is the what is the area of this triangle? This triangle? It is half into z into z by three. That is equal to z square by six. Now that is equal to f z of z when when, z lies between zero and two. So now, what is d F z by d z? z by three, true? It is not all that difficult, once you nicely draw a diagram and see what is happening. And I am trying to encourage you to to try myself. Actually you know, my trying I mean it has given me a lot of pleasure, yesterday evening. I really enjoyed doing this but, but we we do not get that much chance, to do these things nowadays. So I I had an really enjoyable evening; but what I am with that is of that no concern to you. I am trying to encourage you to to to try, yourself and I mean not be intermediated by this. You know I mean the moment you see this, I mean it it hits you I mean you you feel that above all, what is this? It is not all that difficult, okay?

So now probably you can try the other one; which is the little more tricky, nothing tricky I mean the expression is the little more. I did that that expression also, it is coming it is tallying with what is shown. So two x out but may be you can try yourself, okay. This is the this is very simple, now it is a it is a it is a one line are actually, if you if you realize, what is happening. So read the problem. So it says that, g x is a monotone increasing function. This this this monotone increasing function; as we discussed, yesterday we were discussing with some students and this this this monotone increasing function is is important, because that means that for every value of y you have a value of x and and then the function can be inverted, etcetera. If you have a monotone I mean the moment you have a function such that, many values of many values of x give one value of y, you have problems, because you cannot invert the function, you understand?

So so so they are all problems so; but when you have a monotone increasing function, then you have one value of x or one value of y, because because it is it is all the time increasing.

(Refer Slide Time: 42:48)

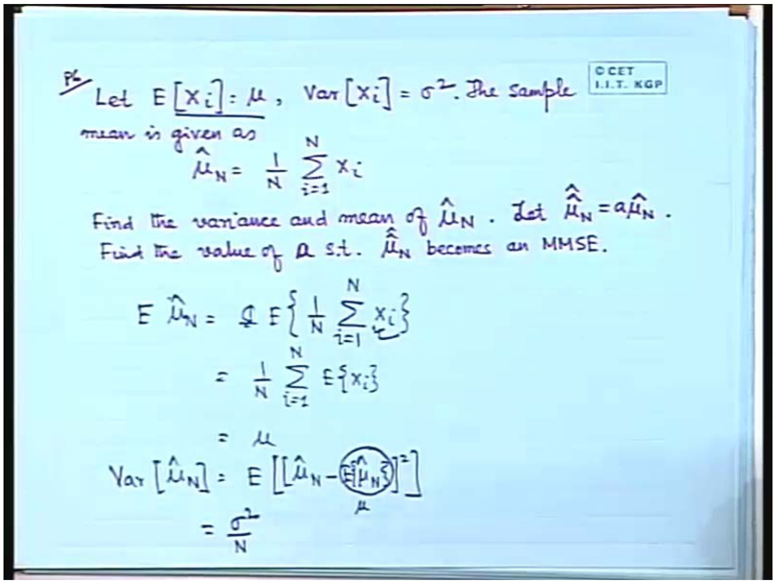


If the function x to y was like this, then here you have a problem, because for many values of x you have one value of y. You cannot invert this function, but if it is going like this, then for every value of x you have every one value of y. So I mean that way monotone increasing and monotone decreasing functions, do not pose property problem. So this is told just to you know, so that some technical difficulties are avoided; for you can forget about it, just imagine that y equal to g x. Now the question is, now now now read this. What is F x y? You have to always go back to definitions, when when confusion, is what what? It is the probability that x is less than or equal to x and y is less than or equal to y, is it not..? Now the point is that; if x is less than or equal to x, then then then necessarily y is less than or equal to g x, is it not? Because it is a monotone increasing function.

So a monotone increasing function, so as if x is below small x, then y will be below g x. If it was a decreasing function, we determine greater than or equal to. That is why this this monotone increasing function is needed, agreed. But the problem is but now the now now the question is that, now suppose; what is what is this thing this thing? This thing that; if y is greater than gx, you are chosen a y now the point is that, can it happen that x is that x is less than or equal to x and y is a y is less than or equal to y where y is greater than g x? If x is less than or equal to a x, y must be less than or equal to g x, because they are inverted like this, rather because, they related like this. So if you have supposed; g x appears to be to to be two, so you are saying that suppose just imagine that g x is equal to two x. So you are saying that, what is the probability that x is less than or equal to one and y is less than or equal to three. It cannot happen that, y is less than or equal to three, so basically y is less than or equal to two you are just artificially, putting that value three. If x is less than or equal to one, then y cannot go beyond two. So so so even if you are taking, putting a value three between two to three you will not get any value of y, it as simple as that. So that is why when y is greater than g x, only this x limit applies. So the y the so the y limit will not apply, because y is determined by x. And if you are putting a value of small y; which is beyond that then in that range between g x and y will not get any value of y, so that part you need consider. Are you following, what I am saying?

Similarly, now if you have artificially chosen y less than g x; that is you you are saying that, what is the probability that x is less than or equal to one and y is less than or equal to ...one point five. So if y has to less than or equal to one point five, then x has to be less than or equal to zero point seven five. x cannot be greater than zero point seven five. So then you have to restrict the range of x and it is the range of y, which will be effective, you understand? So so if y is less than g x then, it is the probability distribution of y which will be dominating, are you following? It it it just simple, I mean there is there is there is nothing in it, I mean you do not have to do anything. It just looks you know, that is a monotone increasing; just to impressive you know, I mean it has a meaning. Not not not not just to impressive but I mean that that you know, puts you off, what is monotone increasing that sort of thing happens. So just fight it that is what I am trying to say.

(Refer Slide Time: 47:43)



Okay we have, this is a very simple problem and a very common problem too. I mean this shows that, this is so we will we will we will close with this one. So it says that suppose; you have a you have a population, you know of random variables. So somebody has performed twenty thousand experiments and have determined that, its its expectation its expected value is new and its variance is sigma square. I mean just it is a value which is now; the point is that if every day if you if you want to find out, what is the mean value and things like that, you cannot do twenty thousand experiments. So you typically, what we tend to do is we take few samples and then take an average, just simply evaluate this. So this is an this is called a sample mean, this is called a distribution mean and this is called a sample mean.

So here what we are doing is, excuse me we are taking generally taking n samples, n is the finite number ten, twenty, thirty, hundred, whatever depending on the time and money that you can spend. And averaging it, simply one by n and we are many times, using that as a mean. Now it turns out that; it is not the mean, it is also a random variable. After all you have taken capital N number of random variables and you have summed them up, so it becomes another random variable. Now the question is that is it a that is if if you consider this to be an estimate of the mean, is it a good estimate? So how do you find out whether it is a good estimate or not; at least on an average if you found out ten thousand such means, then then then took their average, then does it come to the does it converge to the actual mean? That is a measure of its estimated. If it turns out that, this estimator is will be always you know point one, away from the mean irrespective of the number of trials you make, that is not a good estimator, I mean it does not increase with effort.

So the question is does it happen? So so one way of finding that out is to find out, the variance and the mean of this random variable. It is a this is a derived random variable, so what is the mean and the and the variance of this random variable? The that is very simple to find for example, you can easily find out that, what is the expectation of mu hut n? It is nothing but expectation of one by n, just putting i is equal to one to N xi. So good expectation taking, this is the finite sum, you can easily take expectation inside. There is no question of limits not converging, etcetera. So it is one by N, sigma i is equal to one to N, expectation of xi. Expectation of xi is already given as mu, so it is mu. So it turns out that, the expectation of this random variable, at least matches with the distribution, such estimators are called unbiased estimators.

So that is one good thing, now you will ask then ok mean matching. Mean matching does not have much value, if the if the variance is very large then in a given experiment, the value can be away a head from the mean. So you not only want that, on an average it matches with the mean, you also want the variance to be small then, then it is a good estimator, right. So so it is a very legitimate so we can try to find out, what is the variance of this estimator? And it will be good if you find that, this variance decreases as we increase this N. If you take larger and larger number of averages, then this then it will it is more likely; that will come closer and closer to the mean, to the true mean of the distribution. Does it happen? It will happen, and you can you can easily just just put the formula. So what is what is the variance of mu hut N? It is expectation of mu hut N minus expectation of, just simply blindly put the formula, whole square. This is the this is the formula of variance. This you have already calculated. Expand this, put the in place of this, you put this. So it becomes an A plus, B plus, C plus, D plus, E. It becomes a sum whole square. So you know, what it is! Expand it A square plus, B square plus, C square plus, D square minus, twice AB two BC all terms open up. Then take expectation, you should get that this is equal to sigma square by N. That is good because, it is saying that if you take if you consider three samples and you will get some variance. If you take thirty samples and then do this average, see you are putting more effort. So you should get a better estimate, and it is indeed better because, this will be reduced. So if you get a thirty sample average then; you are likely to be closer to the true mean every time on an average then, if you take a three sample. That is what it means, which is true, right. So so we have seen some, my stock is over. We will again see in the evening, I hope you will come up with some other problems, but only hope that this clarified some some.