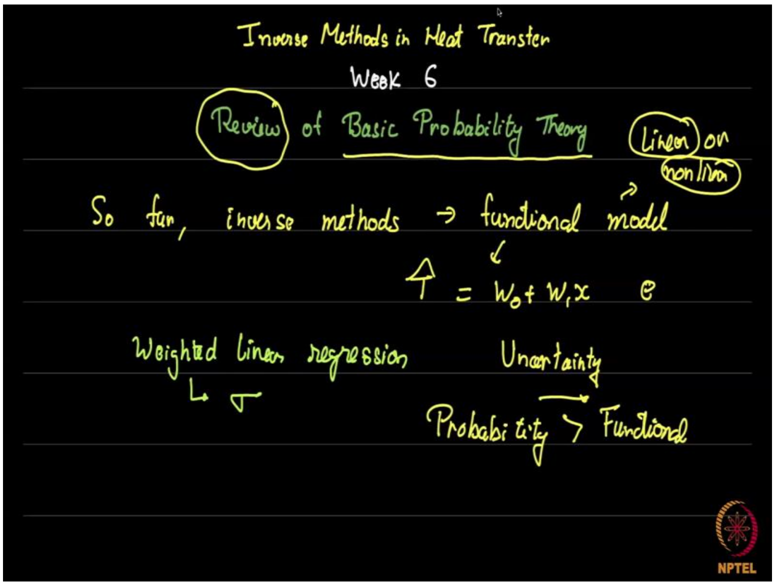
Inverse Methods in Heat Transfer

Prof. Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology – Madras

Lecture – 30
Introduction to Probability for Inverse Methods

(Refer Slide Time: 00:21)



Welcome back. This is week 6 of inverse methods in heat transfer we are going to review basic probability Theory in the series of videos in week six and we will apply this basic probability Theory to some probabilistic methods in week seven. Now so far, the methods that we saw, were methods that were inverse methods that utilized some functional fit. So, a functional model the functional model is something like $\hat{T} = w_0 + w_1 x$ or quadratic or exponential etcetera.

So, these models were either linear or non-linear and we saw that we needed different approaches to solve for linear or non-linear models. Now within that you would have seen a subtle thing, you might remember that we had our weighted linear regression where we saw that not all data points of which we were doing the inverse were weighted equally. Now here I had introduced some term σ which was basically the uncertainty in measuring the temperature of the thermocouples point.

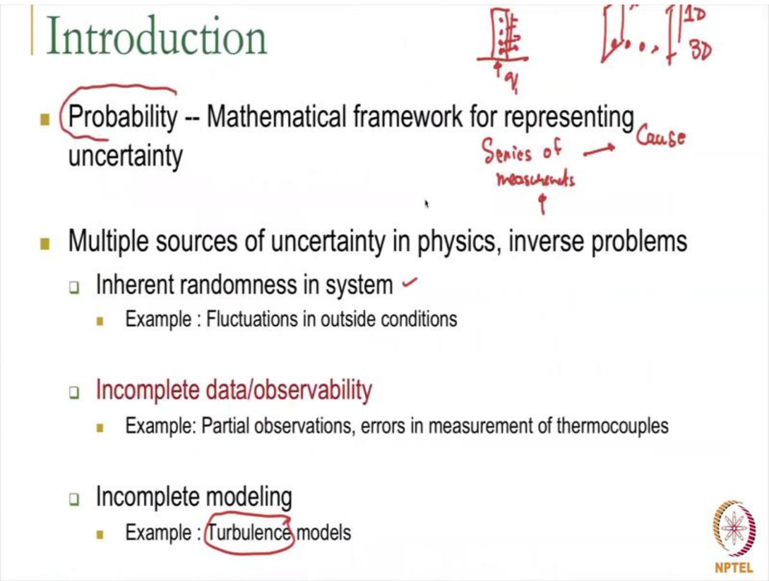
But I am measuring the uncertainty in the measurement made by the thermocouples. Now this is what is going to be extended within this week and the next week. The idea is we actually account for the fact that there is uncertainty in measurement. So, this I am going to talk about a little bit further, but there are different sources of uncertainty and it turns out that this is a powerful series of methods which arise from probability that handle this uncertainty.

In fact, we will see that the probability method is actually a superset. So, probability methods in some senses are more powerful than the functional methods. We will see next week that all the methods we derived so far can actually be derived from the viewpoint of probability. of course, this is a really deep subject as is of course probability theory. So, really speaking we should spend about 40 to 50, 60 hours just reviewing Probability Theory and Basic Statistics.

We do not have the time to do that. we will just spend an hour or two I am assuming that you are already familiar with basic counting probability, though I will review it and we will have one or two such review questions in the assignment. But the purpose here is just a review and not really a full-blown theory I will try to make this compact. So, that you can utilize these insights within the next week when we come to some Bayesian techniques and Markov chain Monte Carlo etcetera.

Few inverse techniques that have been that have proven to be extremely powerful in practice. This is just a basic introduction once again I am just re-emphasizing this the purpose of this in case you found this find this a little bit question or a doubtful or a little bit confusing, please ask questions within the Forum and will be happy to answer it. So, let us go forward and just review some basic probability theory for this week.

(Refer Slide Time: 03:57)



So, what is probability? Probability as all of you know is just a mathematical framework for representing uncertainty. as I said a little bit before the reason, we are looking for probability within inverse methods is, it actually gives us multiple ways of handling the inherent uncertainty in physics, especially in inverse problems. Remember that inverse problems deal with a series of measurements and we want to find out the inherent cause for these measurements.

Again, I will go back to the example we looked at, let us say we have a fin we have a whole bunch of thermocouples we measure the temperature and we want to find out what was the heat transfer here. So, this is the cause and this is the effect. Now of course there are a whole bunch of sources of this uncertainty in physics. for example, you could start right at inherent randomness in the system.

For example, you could start with Quantum fluctuations or fluctuations in outside temperature T Infinity. So, that would be some a little bit of Randomness in the system in the boundary conditions. So, that is one place. Another Place Another source of uncertainty is you could have incomplete data. So, for example here we actually do not have temperature measurements everywhere we only have them at a few places and even within that.

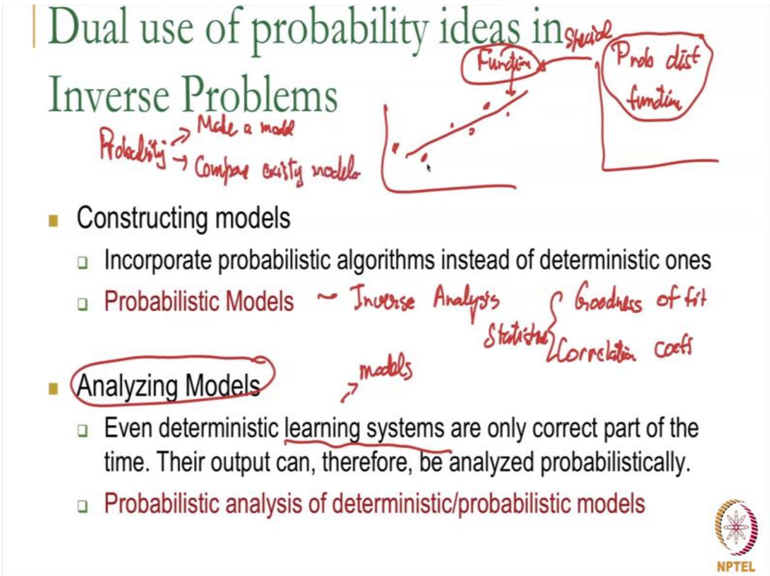
So, your main partial observations or you might have made errors in the measurement of the thermocouples. So, for example any observation so, if you are dealing with MRI there could be noise in your detector you could have a whole bunch of series of problems in how you made the measurement itself. Then finally you could be modeling the problem incompletely. For example, we have been modelling our convenient slab problem as if it is 1D but really speaking the problem is three dimensional.

Again, we assume that all the dimensions are the same but this could be wrong or more seriously the modeling itself is incomplete in that even if you apply a full 3D law and you put geometric certainties and you do not have any inherent Randomness in the system. you still cannot model it because uh you do not have sufficient resolution. So, that is where things like turbulence models when you have turbulent heat transfer etcetera all these things come in.

So, there are multiple sources of uncertainty and we are going to club those together within this one huge thing of saying, there is some uncertainty in the system let us put that together and talk about probabilities and also remember this is extremely useful for ill-posed problems because as you know the same measurements as we saw within our functions could be drawn or could be created by multiple infinite models.

So, we saw that briefly during overfitting also. So, all that which model it is, how accurate the model is, how accurate is the data, all these uncertainties put together come into this one big picture of probability and um so, this is where probability theory is useful in inverse methods.

(Refer Slide Time: 07:30)



Now there are two ways in which we use probability ideas within inverse problems. This is also true of machine learning as we will see later in the next uh three four weeks from week eight to week 10 or week 8 to 11, we will be looking at machine learning methods. So, you will see that there is a very close relationship between inverse problems and machine learning which is why we have included that within this course.

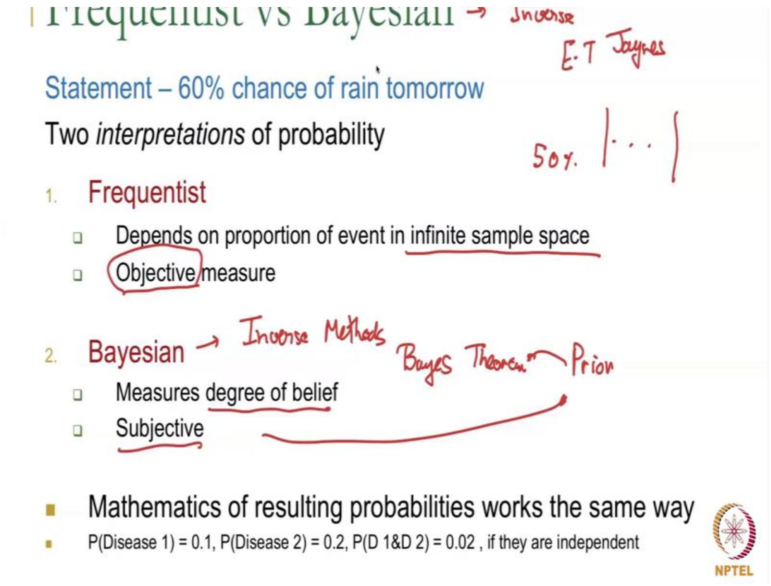
So, the first use of probability ideas is within constructing models. So, the idea is remembering the models that we constructed So far, were well I will assume there is a function which fits this data in a least square sense. So, this is the functional approach. Now instead of that we are going to do something else we are going to say there is a probability distribution function I will explain what this is in case you are not familiar with it I will explain what this is within this video itself at least we will take a first step.

Now this is a slightly different approach. what turns out is the functional approach is a special case of the probability distribution function. in fact, you can show that in some sense the mean of this probabilistic distribution is what appears as the function in what we looked at in the last four five weeks. So, one way in which we use probability theory is to construct or to make such models and then use them for inverse analysis.

The other way in which you can use probability ideas is even if you use deterministic learning systems or deterministic models, they are only correct part of the time. So, for example you might remember goodness of fit, how good is the fit? we will see something else called correlation coefficient later on this week. So, all these ideas are statistical ideas. So, they are useful in analyzing models.

So, even when we come to machine learning we will see that to analyze to tell you how good our model is, we still need a probabilistic hiding. So, that some other model could have fit this data even if it is a deterministic model. So, the model could be, So, probability can be used to make a model or to compare existing models. So, we will be using this theory in both these ways. So, please note is this in making models as well as analysis. So, these are the dual uses of probability which will be using within this course.

(Refer Slide Time: 10:30)



Within probability Theory, there are two different philosophies. I am not going to go too much into the philosophy in fact we are going to almost eliminate it other than this brief discussion. However, since it is a frequent discussion, it makes sense to just address it here. there is something called a frequentist approach and there is of course a Bayesian approach. Bayesian approach is what we are going to be using and it is particularly useful in fact for inverse problems.

You will see it is almost impossible not to take a Bayesian approach uh to inverse problems which is why I am addressing this. Frequentist is usually the method that we have seen. So, for example if you say something like there is a 60 chance of rain tomorrow, then there are two interpretations the Frequentist approach is that you assume an infinite sample space I will Define what a sample space is uh later in this very video.

But what you assume is let us say you have a coin and you are tossing a coin and you assume that if it tosses the coin infinite number of times, then half of those will be heads and half of those would be Tails. So, the next time I toss the coin and say there is a 50 chance of getting a head. So, this talks about some objective measure but the objective measure of course depends on something which is not really realizable that you can never do it in practice.

But you can see it on the computer that is one advantage of doing coding which we will see next week. that you can actually start seeing it slowly converts to 0.5 the number of hits. But if I toss a coin 10 times, I might not exactly see five heads or five tails what I will usually see is something like three heads and seven Tails or seven heads and three tails and we will see how to deal with this inherent uncertainty within I mean this is sort of a meta uncertainty, uncertainty within uncertainty that we will see later.

But what the frequentness approach says is that whenever I say 60 chances of rain, it means assume you are going to live in a universe which splits or in a multiverse and it splits into Infinite universes and out of all those universes in sixty percent of those universes, they are going to have rain and in this is sort of an objective measure. Now there are strong people who believe in this. in fact, the whole school of quantum mechanics in fact some of you might have heard of quantum computation that strongly depends on this frequentist approach.

It strongly depends on assuming that there is some such objective measure of how things happen. And we actually have technology that uses this. on the other hand, we have the parallel assumption of what is known as a Bayesian. So, this measures a degree of belief. So, in the sense that it is a degree of belief it is objective. So, you might say something like you know I did well in the exam. So, I think there is a 90 chance that I will get more than 15 marks.

But nobody has actually done the experiment infinite number of times this is just your degree of belief or when you say 60 chances of rain, it is a rough estimate of saying well I am not 50-50 sure I think there is more chance of rain than not. So, I am kind of estimating. Now you might think that this is vague but it turns out this is very powerful because we have an objective theorem coming out of this subjective idea which is known as Bayesian theorem.

We will prove that we would have anyway seen that in school and purling in college, but we will prove that later on this, because this is as I said before this is almost indispensable for inverse methods why is that? Because when I say something like find out the thermal conductivity given a whole bunch of temperature measurements and if I give a particular material, I actually have an idea of what the thermal conductivity will look like.

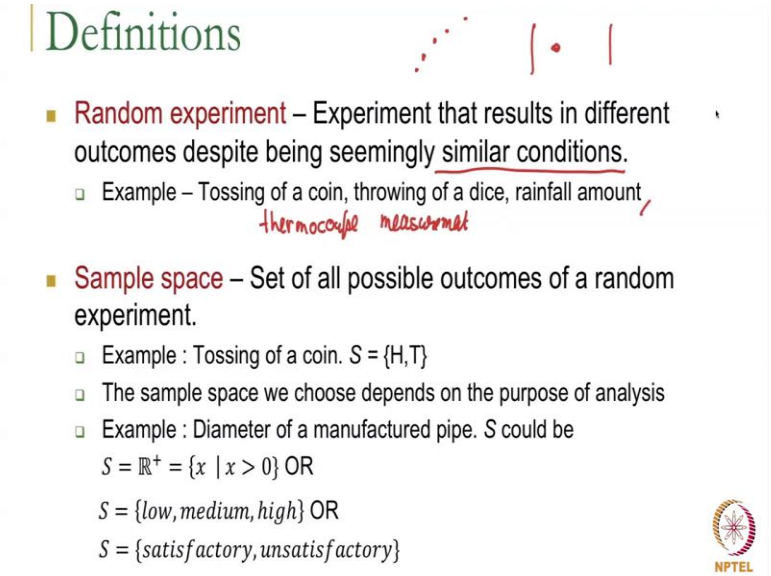
Similarly, if I ask you to estimate, you know how much will be the rainfall in Chennai you have an idea that it is not going to be thousand kilometers. So, you have a subjective idea of how much the range is going to be, in such cases this is called a prior and you can incorporate that the subject to believe can be incorporated back and there has been a lot of debate over the centuries and especially over the last century about which view of probability is correct.

Because in some sense towards the end it has some subtle differences on how we calculate probabilities but as far as we are concerned within this course the probability the way in which the mathematics of probabilities work works the same way within our limited context, we are not doing quantum mechanics. So, if we have something like the probability of disease one is point 2 another the probability of disease 2 is 0.2 then probability of disease 1 and disease 2 is the multiplication of these two probabilities if they are independent regardless of whether you take a Frequentist approach or a Bayesian approach.

So, we will stop talking about these philosophies and primarily take as we go ahead especially during the next week a Bayesian approach and that will be subtle it will be behind the scenes it is not going to be something that we explicitly talk about. Of course, we use base theorem but that is not only the deal with Bayesian probability. in case you are interested in this kind of debate I would recommend that you read some basic books on that.

There is an excellent book which supports the Bayesian approach by E.T James it is an old book it is a really good book if you want to think about such philosophy. So, we are going to end the philosophical debate here and a move on to actual definitions of what probability means and what the various quantities that we are interested in are.

(Refer Slide Time: 16:22)



So, the definitions we are interested in again these are just so, that we have some simple terms to talk about. the first one and the most important one when we deal with probability is what is called a random experiment. So, literally the experiment that we run in heat transfer would be the example of a random experiment. So, I take so, for example the unsteady case that we looked at we make multiple temperatures and measurements in time.

Now if I do the same experiment multiple times with the same condition. so, for example, the simple examples I have given here you take a coin it looks like I have kept it at the same place in your thumb and you are tossing it up Suddenly sometimes it results in head sometimes in tails. Now you might argue this is due to micro conditions or something like that but that apart from a macroscopic perspective it looks like we have set as similar situations as possible and we are getting different results throwing of a dice rainfall amount or thermocouple measurement.

So, I keep the same slab and I put a thermocouple at a particular point. This is not going to give you the same measurement each time you would have done this in school use the Vermeer to measure the diameter of a marble something of that sort and measure it multiple times each time you get slightly different results. Such an experiment which seems to result in slightly different answers each time or different answers each time is called a random experiment. So, please remember this term random experiment.

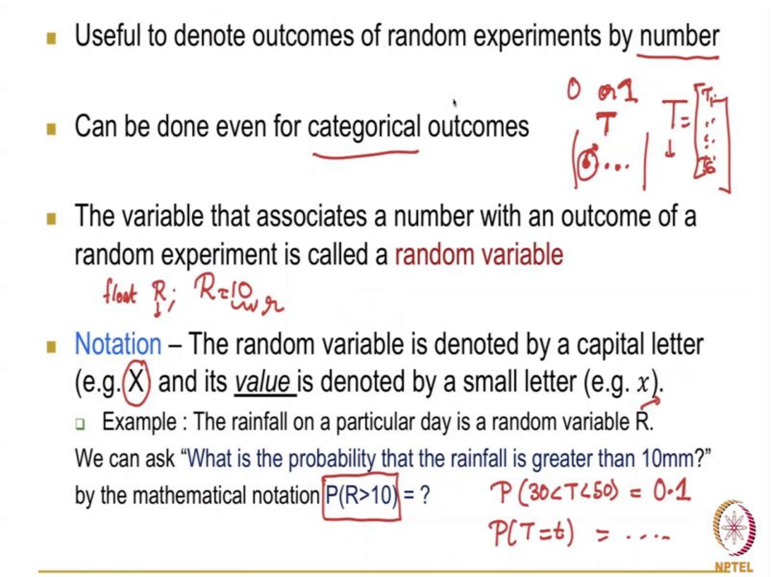
Now the second important term is what is known as a sample space. So, this is basically the set of all possible outcomes of a random experiment. So, in some sense the experiment is sampling out of this entire space, a space has a whole bunch of taste, a whole bunch of possibilities, imagine this is a bag you are pulling stuff out. So, for example the tossing bag has a sample space which has a head or a tail and you are putting your hand in and picking out either head or a tail.

The thermocouple bag is a continuous back. It has a whole bunch of possibilities but even here there will be a range. So, the temperature within a slab with some let us say 10 degrees on 150 degrees on another end is not going to turn out to be 3000 degrees Celsius. It is not going to happen somewhere on Earth in a simple situation. Similarly on the sun you do not expect if the sample space will have high temperatures and it will not have only minus 20 degrees Centigrade.

So, the sample space is where we pick out the possible outcomes of a random experiment, these could be discrete or these could be continuous. so, diameter of a manufactured price. So, how we choose the sample space depends on the purpose of the analysis. this is not particularly relevant here but let us say you are choosing a something like the diameter of a manufactured pipe you will say the sample space is the set of all positive integers or you could say that if you are only interested in finding out whether it is a low or medium or a high diameter type.

You will choose the sample space as discrete. this is a continuous sample space and this is a discrete sample space or a satisfactory or unsatisfactory. Similarly for temperature you could say all positive values, in case we are measuring the temperature in Kelvin or we could say low temperature, medium temperature, high temperature or too cool or too hot or satisfactory thermal conditions or unsatisfactory thermal conditions in case you are doing some HVAC design.

(Refer Slide Time: 20:16)



Now let us come to what constitutes the main thing in the sample space and this is what we are interested in these are random variables. within inverse heat translation generally in engineering and science we deal with numbers. So, random variable is instead of saying heads or tails we would actually want to denote outcomes with an actual number. For example, if it is heads or tails that is what is categorical outcome means two categories, we can call it zero or one.

The variable that Associates a number with an outcome of a random experiment, so, I take a slab I make a measurement and the variable that is going to associate a number is temperature. So, the temperature is now a random variable. Now if I could also take a series of six temperatures and call it some capital T which is made up of T_1 through T_6, this is also a random variable, of course it is a six dimensional or six component random variable.

This will seem like a simple idea but the notation often whenever you read inverse methods literature, it can get a little bit messy especially if you are not used to it. we are going to be a little bit casual about notation but at least once during the course I should mention it because if you do a more advanced course on inverse methods this will often become a bottleneck. So, one of the purposes of this course was to make a somewhat accessible introduction to inverse methods if you see textbooks, they often are very heavily filled with probability jargon and they make it less accessible.

So, hopefully after you go through this course if you are actually entering any formal inverse methods course in heat transfer or in any other subject maybe a little bit more accessible. So, let me make this a little bit clearer. So, the random variable is typically denoted by a capital letter and its value is denoted by a small letter. So, I am going to take a simple example and then I will give you a heat transfer example also.

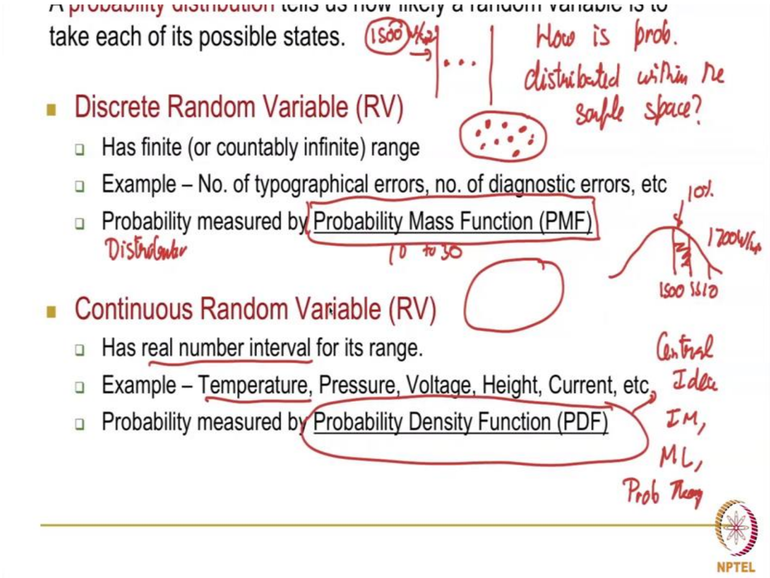
So, for example I would say that the rainfall on a particular day is a random variable R. So, R is the name of the variable so remember just like in a code we have variables. So, if you use C you will say something like in or float or double R something of that sort you would say. And r itself will take values so, for example you will say r equal to 10. this I will denote by small r we may say that is already capital R, r is the set of all possible values of small r.

So, that is what you should remember again but do not get too hung up on the notation, this is just. So, that in some of the slides I am going to show right. Now I will be a little bit particular about the rotation and you might get confused on what capital and small r our capital is the larger set. So, to think and small is the actual value it takes. So, for example if I ask what is the probability that the rainfall is greater than 10 mm, I will denote it by this notation $P(R \geq 10) =$?.

So, for example I could also ask what is the probability that the temperature at this point lies between 30 degrees and 50 degrees. So, let us say this is point one I would give something of this sign I could also say probability that the temperature is equal to this is actually wrong because you cannot get a specific temperature but temperature equal to let us say small t is something. We will come to this a little bit more carefully when we come to probability density functions.

So, I hope the idea itself is clear Capital denotes the set of variables and small denotes the actual value that variable takes.

(Refer Slide Time: 24:28)



Now here is probably the most important idea that you should get at a conceptual at a gut level when we deal with inverse problems, that is the idea of probability distributions. Now the important idea here is that probability distributions tell us how likely a random variable is to take each of its possible states. Now this might seem way but let me explain why this is important usually we talk about probabilities of events.

Individual probabilities are for events or outcomes from sample space on the other hand the probability distribution is for the entire sample space, how is it different? I might ask what is the probability of a head and you might say it is 0.4 let us say it is a biased coin and you might say it is 0.4 that tells me something but, in this case, because there are only two outcomes, I will immediately know that the probability of tails is 0.6.

So, when I talk about probability distribution, I am talking about the entire sample space. So, imagine there is a sample space all the possible events are here and all of them sum up to one, right all possible events sum up to one, what probability distribution tells you is how is probability distributed within the sample space. Now so, remember probability distribution corresponds to sample space whereas probability corresponds to one event or a few events whatever you are interested in.

Why is this important? this is probably in my opinion the most central important idea as far as inverse methods are concerned, the idea of a distribution. Because it is not just enough for me in an inverse case to tell you, heat transfer is 1500 what is per meter Square. If I ask you how sure are you then it actually makes sense for you to give the entire distribution. why is that? So, you want to know not only what the possible heat transfer is or what the most likely heat transfers you also want to know the entire range.

Let us say the likelihood of lying between 1500 and 1510 is something like 10 percent but there is actually a chance that you might hit 1700 watt per meter Square. you want to know what the chance is? Is it 10 percent? is it five percent? is it less than one in a million because when you design equipment you want to know not only one or you want to know the maximum temperature you will say what is the maximum temperature within the slab you might say 70 degrees Centigrade but maybe the there is a small chance that it might hit 90 and it might conch off.

Usually when design goes wrong it goes wrong because we did not take care of the distribution and we looked only at the averages. So, here is where the idea of a probability distribution becomes very powerful. You do not look at only one specific event you look at all events in the sample space and you tell them or you tell the person who is interested in finding out something about the system the entire distribution.

So, remember this word distribution people will keep on talking about it whether you come to inverse or whether you come to machine learning especially people will keep on. So, what is the distribution, what is the distribution people want to know not only the specific event that you are interested in it but all the possibilities of all the events. So, usually when we report we report the mean what was your average score.

So, just imagine you are going to a college for admission and you tell them that my entire GPA or my percentage was 85 percent, but they want specific courses they want to see what your full marksheet your full grade sheet. And if you go within further detail how much did you score in each exam within each grade. So, probably distribution has the entire information. So, if you want to quantify the uncertainty of any event or any sort of experiment all the uncertainties the distribution has the entire information about it.

So, rather than asking for specific probabilities you ask for the entire information if it is available. So, sometimes it is not available and we will see how to reconstruct it next week that is our primary task as far as next week is concerned. But remember this if nothing else from this week please do ingrain this within your ideas that probability distribution is very important. Now once you understand the idea of probability distribution you can immediately see that there are two possible ways in which this distribution could be done.

One is it is discrete that is the sample space has discrete separate points. So, for example this could be finite or it could be countably infinite I will explain what countably infinite is. So, it is not particularly relevant. Finite is a number of typographical errors in a page for example how many errors I might have made some spelling errors grammar errors within this page how many are those how many errors did a doctor make?

How many times did it rain last month all these are discrete random variables all of you understand what discrete is? What I mean by countably infinite is, how many stars are there in the universe it could still appear like a random variable because we do not quite know the number but it could be really large there is more specific limit to it but nonetheless it is discrete it is not 3.1.

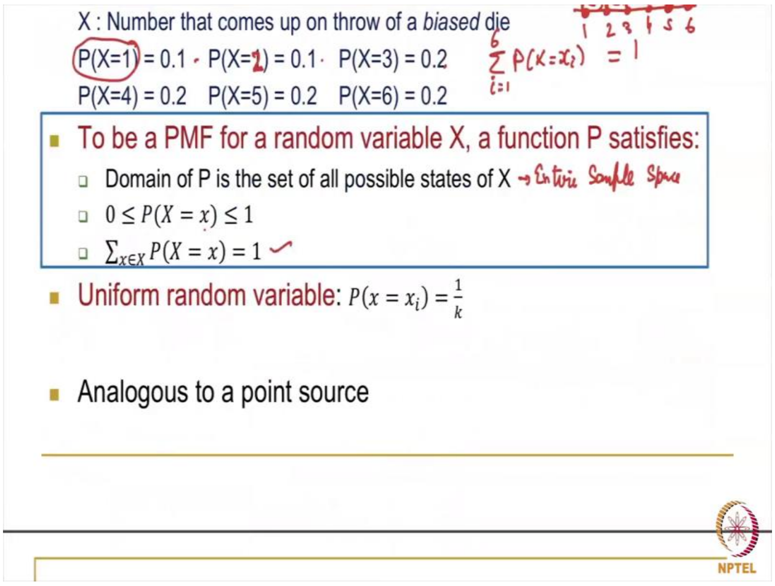
So, there is continuous random variables is where the sample space is like temperature. it has a real number interval. So, for example if I want the heat transfer or if I want the temperature, I want the heat flux, I want to pressure all these are continuous random variables you could say temperature will vary somewhere between 10 to 30 degrees Centigrade, but it will vary continuously for example within Chennai it will vary somewhere between in very rare occasions like seven degrees centigrade to let us say 48 to 50 it will not go above 55 it will not drop below 7.

So, that could be a range but nonetheless it is a continuous image. So, in the case of a discrete random variable, you will measure probability distribution really as what is known as a probability Mass function, whereas in the case of a continuous random variable you will measure a probability by a probability density function. Now we will not consider this too much though I will introduce you to this in the next slide just for completion's sake but this thing PDF.

So, PDF is not our document PDF is a probability density function, it is a central idea I cannot tell you how important it is it is a central idea within inverse methods as well as machine learning's and in machine learning as well as in general probability Theory because most of the time especially for our applications we tend to deal with continuous variables. So, and usually for some reason all of us find it a little bit confusing.

So, we will deal with this and you will see some deal of this confusion when we come to next week also, but hopefully we would have done enough background this week that will be clear when we come to this next week. So, discrete random variable versus continuous random variable and if nothing else remember this continuous random variable and probability density function.

(Refer Slide Time: 33:03)

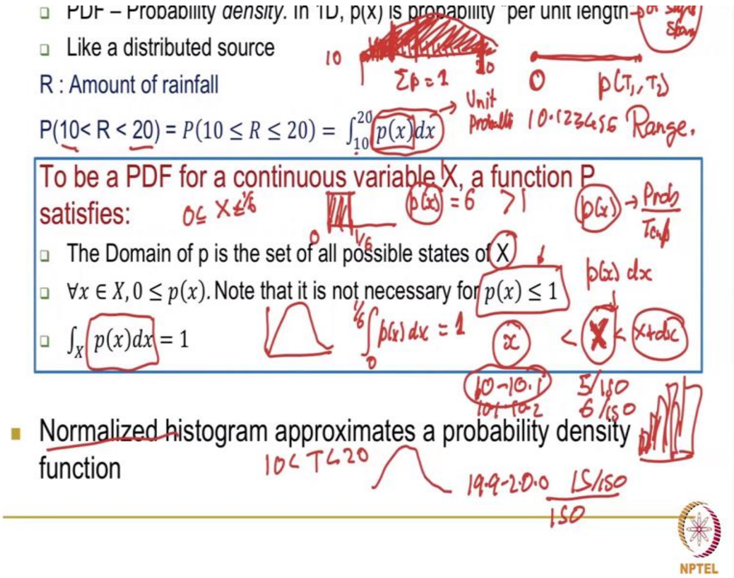


So, let us quickly come to the probability Mass function, which is a little bit more intuitive and then we will just use this intuition to figure out what a probability density function is. So, probability Mass function is just the list of all possible values along with their probability. So, this is a simple idea. As I told you, you take the sample space the sample space has a few events in this case let us say we are throwing up a dice which has been manipulated and not all sides are equally likely.

And you just list all their probabilities this should have been probability of x equal to 2. So, it throws up either a one two three four five six and the probabilities are written here. So, to be a probability Mass function for a random variable x, you need to satisfy a few things, of course remember the entire sample space has to be written down. And each individual probabilities for example if it reaches 1, 2, 3, 4, 5, 6, all of these probabilities have to be less than one that is kind of obvious.

And the summation of all those possible probabilities should be exactly one. So, which is the case here 0.6 plus 0.8 Plus 0.1 plus 0.1 this is $\sum P(X = x_i)$, this is the simple idea of a probability Mass function you could also draw it like this. So, for example x equal to 1, 2, 3, 4, 5, 6 and it will look like this it will look like a point load 0.1, .1, .2, .2, 2, .2. So, this is the diagram of a probability Mass function notice it cannot take any value between one and 2, 2 and 3, 3 and 4 this is where it differs from the idea of a probability density function.

(Refer Slide Time: 35:25)



So, the probability density function is what we will see next but as I mentioned here our probability Mass function is like a point source. So, you could think of if you have done solid mechanics, you would have something like a point load or within something like heat transfer you could just have q at a single point all these are idealizations but it works in a very similar way in that we simply sum up these things one by one and find out the net probability.

Now when we come to probability distributions as I said the more important or the more practical quantity for inverse methods in heat transfer is this idea of a probability density function. So, this probability density function can work in multiple dimensions which will come to actually next week not really this week. But let us just assume this is in 1D, this is somewhat like probability per unit length, unit length of what, unit length of the sample space.

So, like I said this is like a distributed source of probability it is useful for you to imagine that there is the space and a whole bunch of infinite events are happening continuous infinite units. So, for example this could be a space with temperature and pressure or you could have two thermocouples, this could be simply a sampling between T_1 and T_2, XY coordinates would mean T_1 and T_2 and each one of these points is a possible value of T_1 versus a possible value of T_2. So, in that case we are giving something like $P(T_1, T_2)$.

Now there is a small problem here which is what we try to address through this probability density function rather than a probability Mass function. So, the problem is this; the problem is let us just take a 1D case. So, let us say there is some I have given the example of rainfall but let us take the example of temperature. let us say temperature varies between 10 and 20. And we have to give a probability for each event, remember we are trying to distribute this entire probability which sums up to one between these values.

Now there are infinite points here right. So, obviously each of them regardless of what you do each of them is likely. So, you will basically get zero when you divide 1 divided by Infinity you are going to get 0 as the likelihood of each point. So, you cannot ever say what is the likelihood of getting a temperature of 10.123456 etcetera. What however we can do is give probability for a range or we can give the density of probability.

Just like if I want the mass of a point within a solid body, I cannot give it but I can give density of a solid point because that is mass per unit volume. Now you can think of probability density as a non-uniformly dense body, it starts here ends here and it tells you in each Place how dense how much more likely it is. So, I will show this via histogram later on but the reason why we define probability this way.

Now notice $P(x)dx$ is actually the unit probability what it means is instead of giving probability for a number we give probability for a range. as an example, let us take a finite range and then we will reduce it to infinitesimal range. So, for example I will take a finite range and say what is the probability that rainfall lies between 10 and 20 mm or I can say what is the probability the temperature lies between 10 and 20 degrees centigrade. Now for that you actually integrate in this entire area and you tell me the net probability.

Now what do you integrate, what you integrate is the density of the probability. So, you say integral from 10 to 20 of $P(x)dx$ where $P(x)$ is probability per unit temperature that might not make much sense but it is unit length in sample space as I have written here that we take. Now, what it means in finite terms is what $P(x)$ does is if you do multiply $P(x)$ by dx it tells you the probability that $x \leq X \leq x + dx$.

So, notice this is the random variable these are the two values. So, for example P at 10 means what is the probability that the temperature lies between 10 and 10 plus some small dx or dt obviously this number will go down to 0 as dt goes down to 0 but dt, dx is a finite value what $P(x)dx$ is actually a finite value. So, what we do is we integrate between 10 to 20 not $P(x)$dx, $\int P(x)dx$ between any two x is actually a finite value is using this. And using analogy with what we did for discrete values.

We now write what a probability density function for a continuous variable is and the conditions it has to satisfy. So, the condition it has to satisfy is that the possible all possible states of x have to be included as I said here it is not necessary. So, this is where it differs from discrete cases $P(x)$ need not be less than equal to one, all we require is $\int P(x)dx$ has to be equal to one. you can imagine a function where individual values individual densities are greater than one but it integrates actually to one.

So, this is an important difference from before. So, for example we could think of a simple case something like this, let us say x is a variable capital x is a variable that varies between 0 and 1 by 6. So, x is a sample space such that 0 is less than x less than 1 by 6. and I have $P(x)$ is simply 6. So, you can notice this density is greater than one, but when I integrate $P(x)$dx between 0 and the entire sample space domain you are still going to get 1, 6 times 1 by 6 you are going to get two.

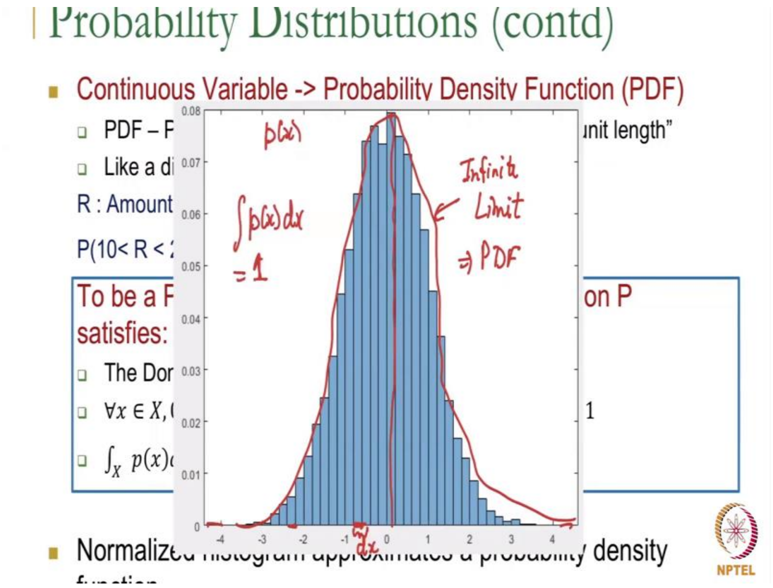
So, it is entirely possible for the probability density since it is a density it can always be greater than one all it means is it is really dense it is a relative quantity, what $P(x)$ talks about is how much more likely is it for me to find the random variable here in this space rather than somewhere else that is all it really means. So, you should not give Direct Value to $P(x)$ you can only comparatively look at $P(x)$ versus somewhere else. So, if it is higher at that place, it is more likely for you to find the random variable. this is used very powerfully with an quantum mechanics and you will see this later on also we will talk about this a little bit more when we come to the direct inverse problems in the next week.

Now a simple idea with which you can build this intuition is to make a histogram. So, for example when I say that temperature lies between 10 and 20 and temperature is a random variable um the way we would actually make measurements is we cannot measure all the infinite temperatures we will first say let me just count the number of times let us say I am taking a thermocouple and at a particular point I am measuring the temperature and I can.

Now count the number of times that the temperature light lay between 10 and 10.1. So, maybe it was five times then 10.1 and 10.2 maybe it was six times etcetera etc. then I went till 19.9 to 20 and this was let us say 15 times and now I take the total number of events, let us say this was 150 and find out the fraction 5 by 150, 6 by 150, 15 by 150 and then draw a histogram this is what is known as a normalized histogram.

So, that all of it, all of the fractions add up to one. Now of course you I could reduce this range further instead of making it point on I could make it 0.05 then I will make it a little bit thinner but it will look a little bit more continuous. And as I keep on continuing this process, this starts looking like a smooth curve. Now that is what a probability density function is. it is sort of the limit of taking infinitesimal widths and infinite measurements and that is what looks like a histogram.

(Refer Slide Time: 45:25)



So, for example something of this sort, so, here the width is let us say something like 0.2, I think I took it some variable that changes from minus infinity to plus infinity you can hardly see anything here. But it starts looking like you can imagine that the infinite limit of this histogram is what is the probability density function. So, please try to we will use the idea of probability density function in the next week especially powerfully, but please try to get it into your concepts.

That ultimately you want this gap which was point to here to go to dx as small as possible and in that case this height here would represent $P(x)$ and integral of here, I said it is treated as a discrete variable then $P(x)$, Sigma of $P(x)$ into Delta x would be 1 but ideally you want integral of $P(x)$dx is equal to 1. so, we will move on to other ideas from probability in the next video and see in the next video, thank you.