Measure Theoretic Probability 1

Professor Suprio Bhar
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture 19

Construction of RVs with a specified law

Welcome to this lecture. Before we continue with the topics of this lecture, let us first quickly recall what we have been doing in the past week, and this week. So, using the ideas about measure theory from Weeks 1 to 3, we have identified the law of a random variable or a random vector. So, what we have done, given a Borel measurable function X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we put the measurable structure of X and probability measure $\mathbb{P}$ together and obtained a probability measure on the range side.

So, for the case for a random variable, we got the real line with the Borel σ-field and for the case of $\mathbb{R}^d$ valued random vectors, we got the $\mathbb{R}^d$ with the Borel σ-field of $\mathbb{R}^d$. So, on top of that we have constructed with probability measure which we denoted by $\mathbb{P} \circ X^{-1}$ and that was called as the law or distribution of the random variable or the random vector X.

And after identifying its properties and using these properties, we have gone to the distribution functions and we have found that the basic properties of distribution functions that we studied in basic probability courses, all of that can be recovered. So, from the properties of the probability measure $\mathbb{P} \circ X^{-1}$, we can prove the properties of the distribution functions.

In particular, in the previous lecture, we have spent a lot of time in identifying contribution from the jumps and the remaining part and using that we have gone back to the standard decomposition of a distribution function into the contribution from the jumps that is discrete part and the remaining part, that continuous part.

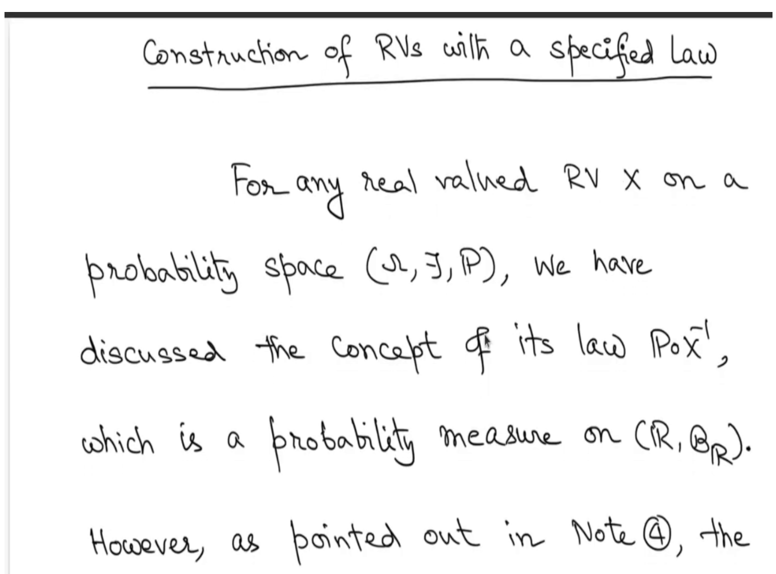
And then we have also identified something about the discrete random variables and we have talked about their law. So, we have connected these discrete random variables with discrete distribution functions. And we have also mentioned that for the case of continuous random variables and corresponding continuous distribution functions, we will require some more measure theoretic setup, which we will come back to in Week 8.

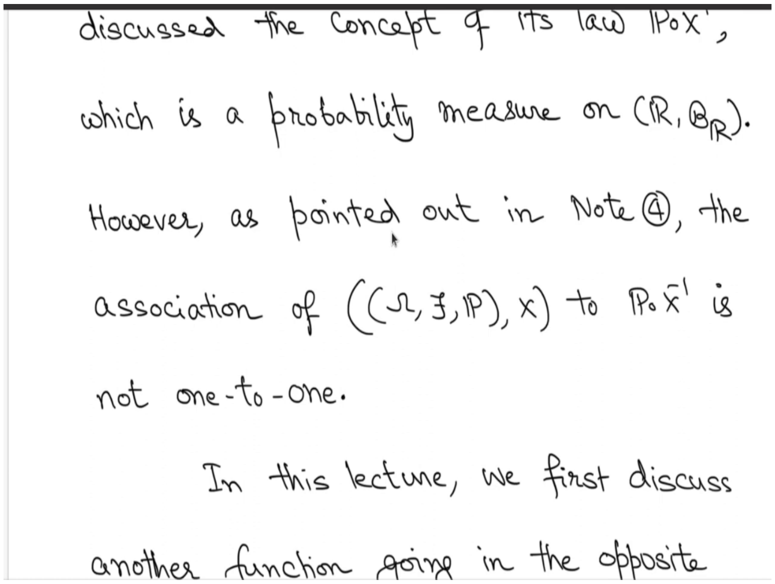
So, for the moment, we continue with the ideas about general random variables or if required, we will restrict our attention to discrete random variables. But then, one of the important questions that we have left out in between was this. So, given a random variable or a random vector on a probability space, we can construct a probability measure on the range side.

But then, there was this question that given a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, the Borel σ-field on $\mathbb{R}$, real line with the Borel σ-field on $\mathbb{R}$ or the $\mathbb{R}^d$, with the Borel σ-field of $\mathbb{R}^d$, if you are looking at such measurable spaces, and if you are given a probability measure on top of it, can you construct a appropriate probability space? Basically, can you reconstruct the domain side? If you are given the range side, can you reconstruct the domain side? So that is basically the question.

So can you find probability spaces and random variables defined on top of it or random vectors defined on top of it such that $\mathbb{P} \circ X^{-1}$ turns out to be the exact specified probability measure that you have started of with. So, we are going to answer this question in this lecture.

(Refer Slide Time: 04:00)



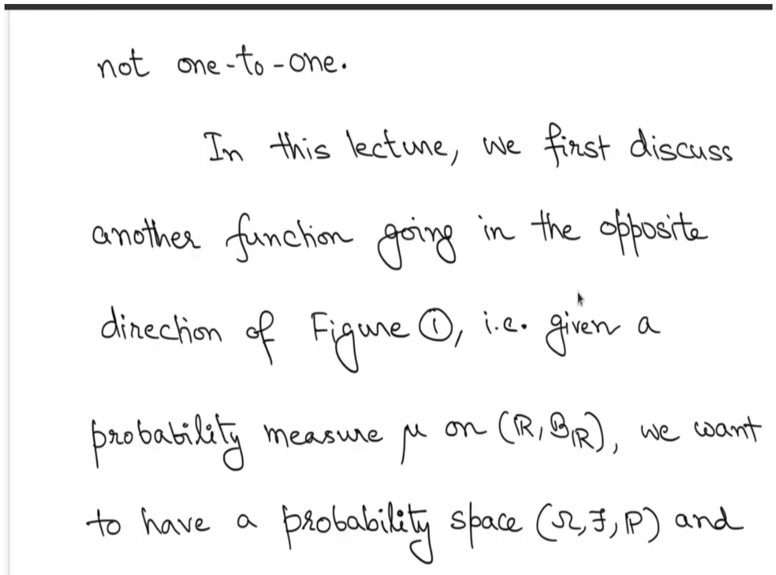

discussed the concept of its law $\mathbb{P} \circ X^{-1}$, which is a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. However, as pointed out in Note ④, the association of $((\Omega, \mathcal{F}, \mathbb{P}), X)$ to $\mathbb{P} \circ X^{-1}$ is not one-to-one.

In this lecture, we first discuss another function going in the opposite

So, for any real valued random variable, let us look at this situation that if it is defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we get that the law is defined on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, it is a probability measure. But then, remember, this association, that whenever you are starting off with the probability space and the random variable and when you go to its law which is a probability measure, that association, it is not one to one because if you consider any probability space and specify the constant or degenerate random variable, then you will get the Dirac distribution as its law.

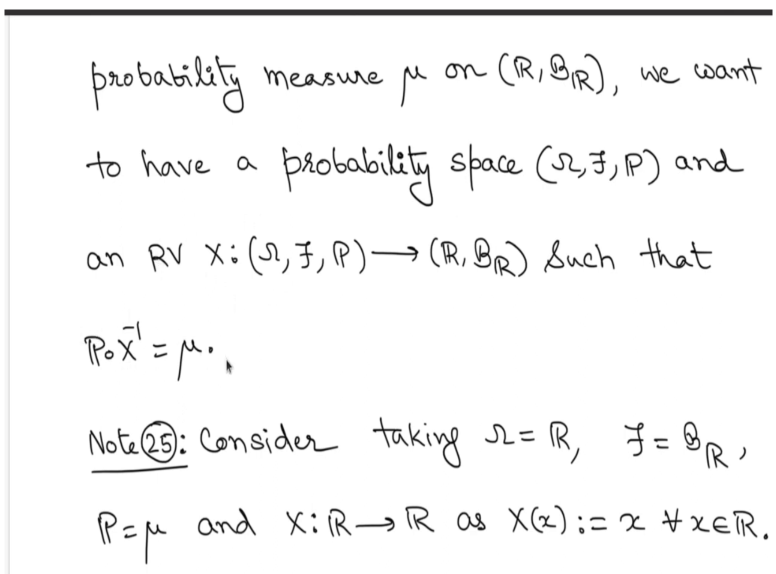
So, therefore this association is not one-to-one. So, there may be many possible different probability spaces and random variables defined there such that its law are the same. So, this association is not one-to-one.

(Refer Slide Time: 04:56)



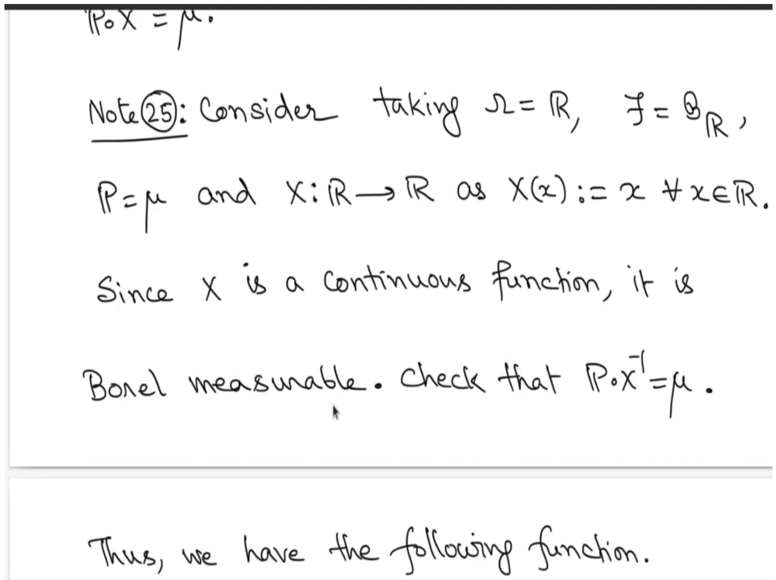
But what we would like to do is to construct a function, or a association going in the opposite direction.

(Refer Slide Time: 05:01)



So given a specified probability measure μ, we would like to construct the domain side, meaning, we would like to construct the probability space and the random variable X such that the law of $\mathcal{B}$ is exactly the, the given probability measure μ. Great.

(Refer Slide Time: 05:19)



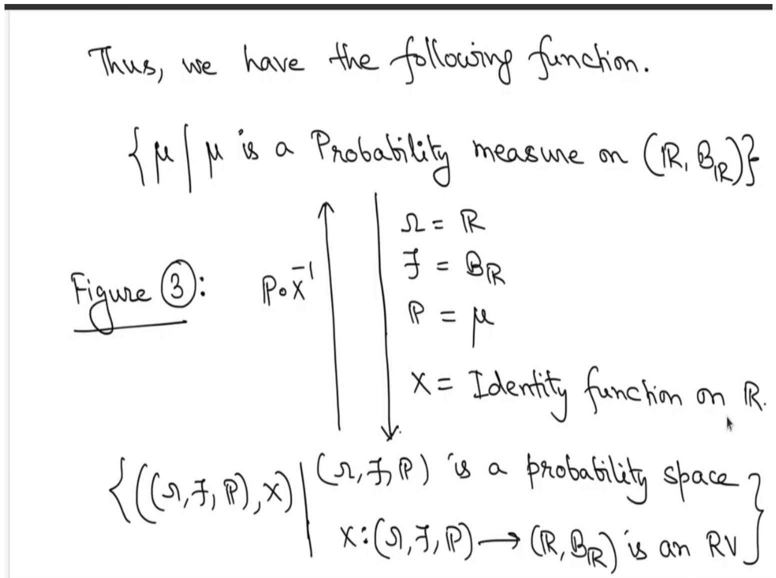
So, how do you do this? The construction is pretty simple. So, this is just a explicit construction of such a probability space and the random variable. So, what do you do? Remember, you are given the range side, range side probability space. Range side, you have the real line, Borel σ-field and the specified probability measure μ. What do you do? You take the exact same things on the domain side.

You put Ω to be the real line, take the σ-field $\mathcal{F}$ to be a Borel σ-field on ℝ and take the probability measure as the probability measure μ. Great. So, you now have the same probability space on the domain side. And now, look at this continuous function. So, this is the identity function that sends any real number X to itself. So, this is the identity function which is a continuous function.

So therefore, it is a measurable function also. So, therefore, you have a measure space, basically a probability space on the domain side. You have a measurable function, Borel measurable function defined on the (())(06:26), and now, what you can check immediately that $\mathbb{P} \circ X^{-1}$, with these choices, $\mathbb{P} \circ X^{-1}$ will exactly become the measure μ.

This is just rewriting whatever information you have already been given and since you have taken the function X to be the identity function, you are not really changing anything. But this gives you a explicit construction of a random variable with the specified law μ.

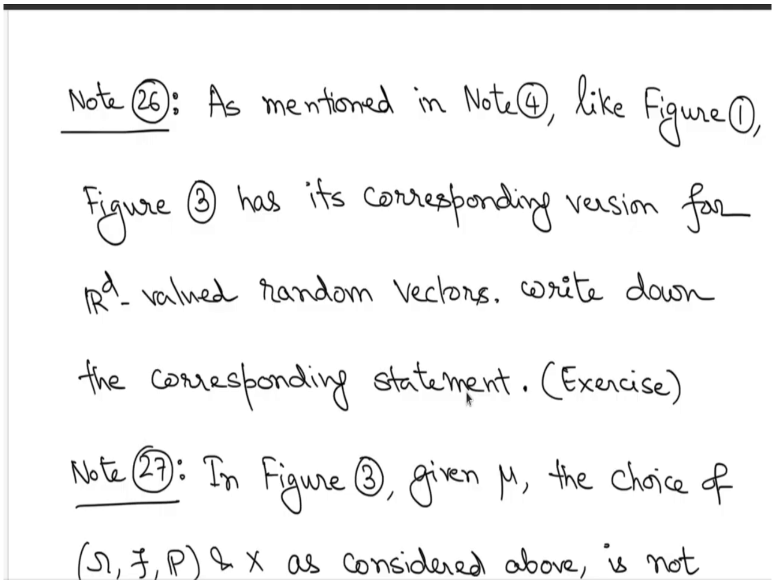
(Refer Slide Time: 06:58)



So therefore, you now have a association or a function going in he direction opposite to the structure that we have discussed earlier. So, in Figure 1, in Note 4 earlier, we had discussed this.

So, $\mathbb{P} \circ X^{-1}$ was going in this direction that given random variables and on top of certain probability spaces, you look at $\mathbb{P} \circ X^{-1}$, that landed up in (Pro) collection of probability measures. So, that was this deduction. But now, we are now constructing the other direction with these spaces for the domain side. So, therefore you get a function on the going in the other direction.

(Refer Slide Time: 07:37)



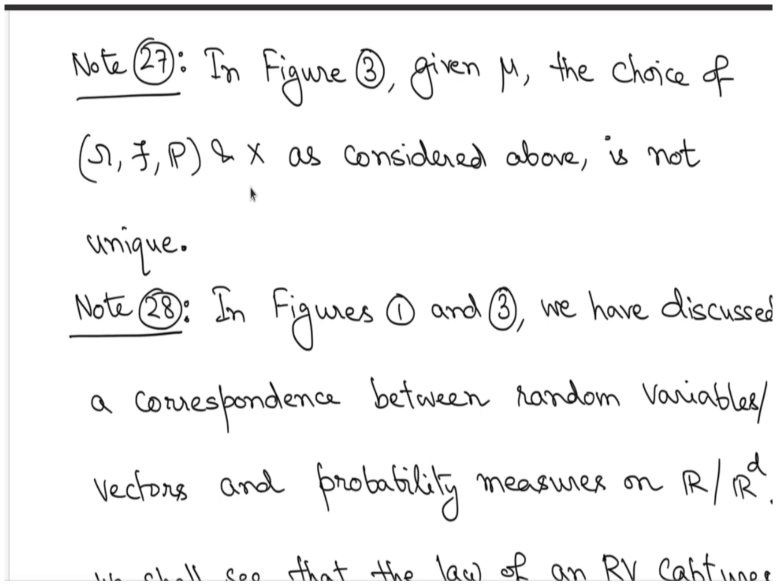
But then this was the statement for the random variables. And as stated earlier, as discussed earlier, you can now try to figure out the corresponding version for random vectors. So, take a $\mathbb{R}^d$ value random vector. Look at the corresponding law. So, it is a probability measure on $\mathbb{R}^d$ with Borel σ-field of $\mathbb{R}^d$.

But then, if you are given a probability measure on $\mathbb{R}^d$ with the Borel σ-field of $\mathbb{R}^d$, then can you construct a random vector with that specified law? So that is basically the question. So, you can try to write down the corresponding statement by choosing the probability space and the random vector appropriately.

So, exactly try to make the same choices as done above. So, what you have to do is to take the domain side probability space to be the exactly the range side probability space and take the identity function as the random vector. Check that it is a continuous function and show that it is Borel measurable.

So again, continuous functions were Borel measurable, whenever you are looking at this $\mathbb{R}^d$ to $\mathbb{R}^m$ or this type of structures. So again, try to write down that statement. So, that is left as an exercise for you.

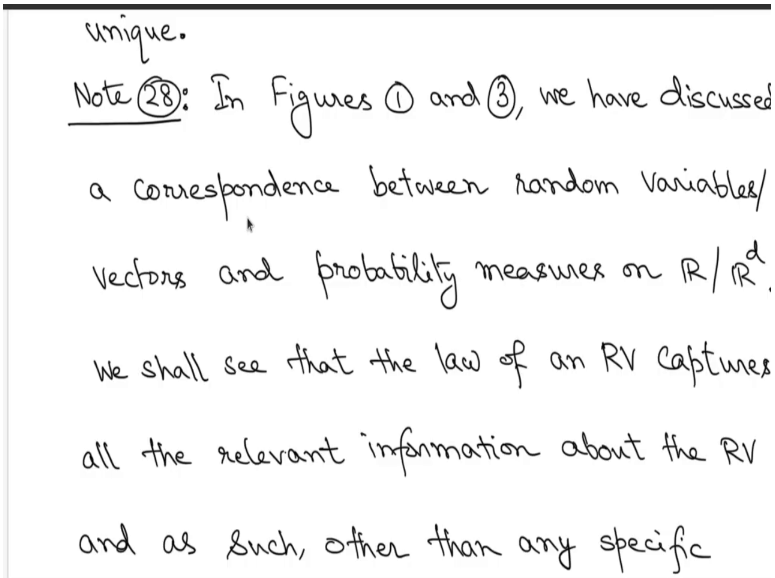
(Refer Slide Time: 08:50)



So, what we have now done, given any probability measure μ, we have constructed the probability space and the random variable and what we have figured out is that we can construct such random variables with the specified law. But is important to note that this was one of the possible choices that gave us $\mathbb{P} \circ X^{-1}$ to be μ. And the choice usually is not unique. So, you can try to check this.

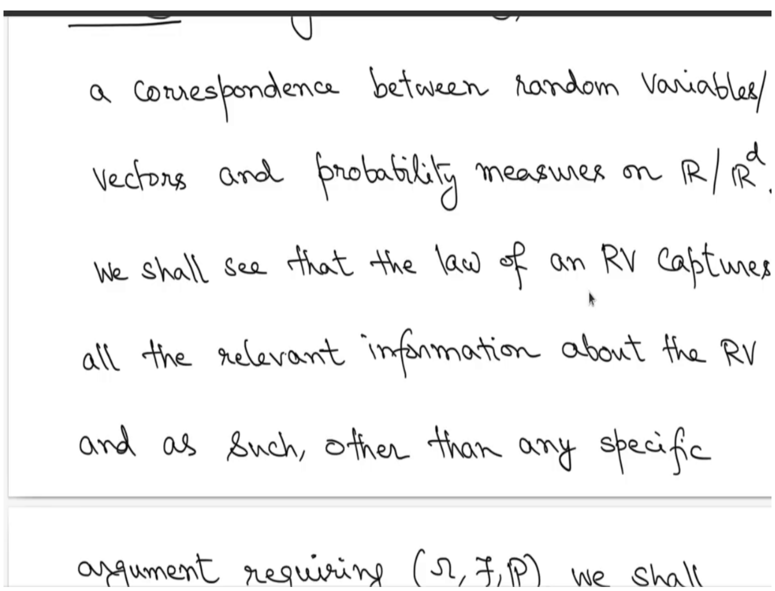
So, can you figure out other probability spaces such that you can define other random variables such that $\mathbb{P} \circ X^{-1}$ also becomes μ. So, try to work this out. Great. But a priori this choice of the probability space on the domain side and the random variable is not unique. But now put everything together.

(Refer slide Time: 09:42)



So, in Figures 1 and 3, now we have discussed this correspondence between two connections. So, on one hand you have connections of random variables on probability spaces, on another hand you have probability measures. And you have now discussed a correspondence, you have functions going in both directions. So, you have a correspondence.

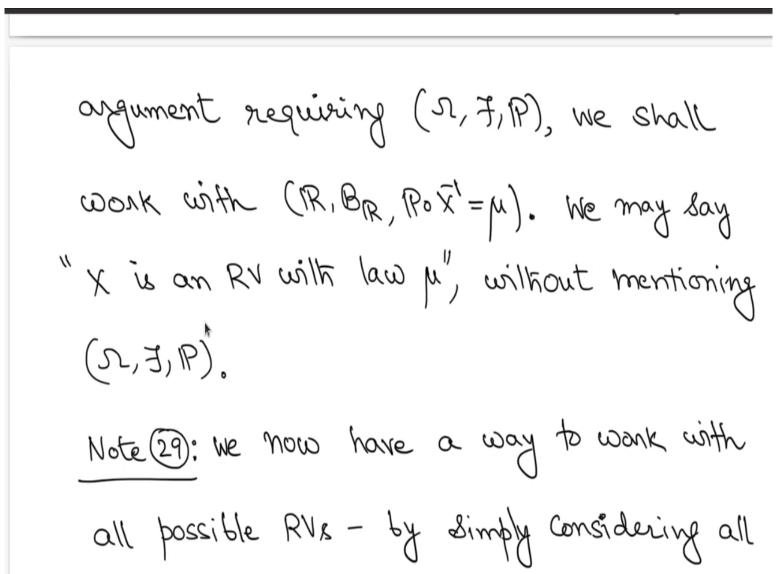
(Refer Slide Time: 10:02)



So, what we can see is that the law of a random variable captures all this important information about the random variable or the random vector. So, we have already seen some of it, when to check the properties of distribution functions. And all these properties of distribution functions that were well known and which you have studied in basic probability now follows from the probability measures properties, follows from the properties of the law $\mathbb{P} \circ X^{-1}$.

But then, we are yet to discuss things about moments and so on. But we shall also see that all of these can be written purely in terms of the law. So, the law is quite important whenever you are discussing the random variables or random vectors. So now, in this correspondence between random variables and random vectors and the corresponding law, it so happens that the law becomes quite important.

(Refer Slide Time: 11:04)



And since the law captures all the relevant information, what we can do is that we can concentrate our attention purely on the law. So, purely on the range side. So, if you are dealing with a random vector, look at $\mathbb{R}^d$, Borel σ-field of $\mathbb{R}^d$ and the probability measure, that law.

And if needed, you can construct appropriate random vectors with that specified law. So, what you really want to do is to focus on, more on the range side rather than the domain side.

So, what we shall do is that we will mostly work on the range side. So, for the (ran) case of random variables, it will be this. Real line, Borel σ-field of the real line and the probability measure. So, if you are given that, then you can extract all the relevant information about the random variable or the random vector.

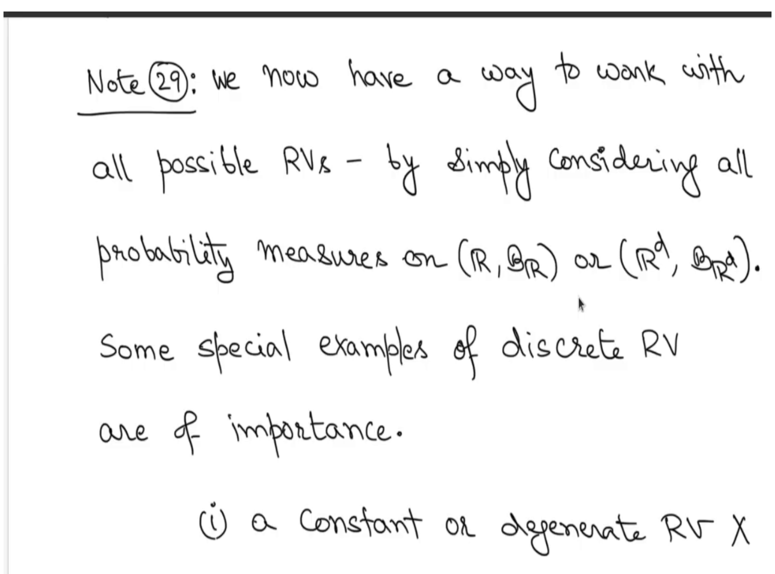
So therefore, we will usually ignore to the domain side unless strictly necessary, we will stick to the random variables only through the law. We will look at the random variables only through the law. And we are going to say such statements that X is a random variable with law μ without mentioning the domain space.

So, from now on, whatever discussion we have done for probability spaces which appeared on the domain side, we will now ignore. We will look at purely on the range side, which is of course another probability space, the induced probability space.

The induced probability space appeared due to the domain space and the measurability structure but all that information is now concentrated on the range side and we will restrict our attention there and ignore the domain side.

Unless strictly necessary, we are not even going to mention the domain side. We are just going to say things like X is a random variable or a random vector with some law μ. And the law means it is a probability measure on the range side. So, that is all we are going to concentrate on.

(Refer Slide Time: 13:14)

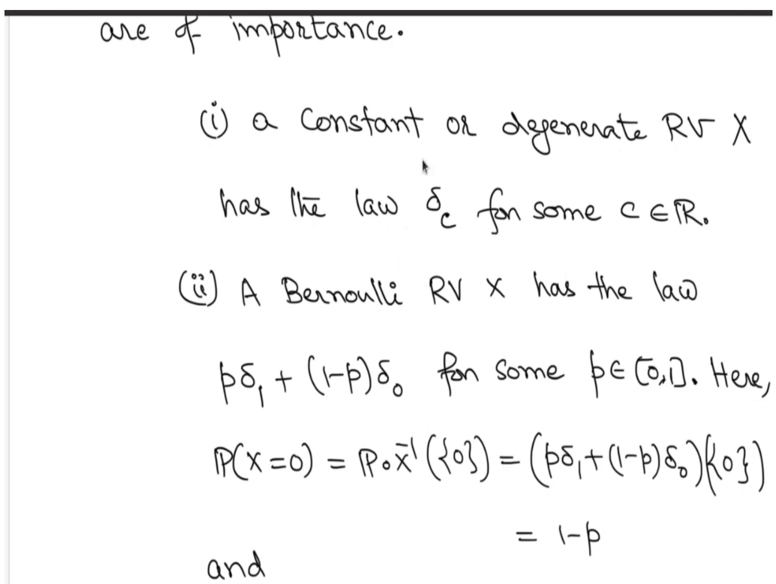


Now, observe this that we have a way to work with all possible random variables now, or all possible random vectors. Why? Simply consider all possible probability measures on the real line or all $\mathbb{R}^d$. So, if you consider all such laws, then of course, all such random variables that you really needed, that, information about those random variables or random vectors are actually concentrated in the law.

So, as long as you study all possible probability measures on the range side, then you are done. You just look at the random vectors, look at its law, then as long as you know everything about probability measures, you are done. So, therefore, what you need to do is to look at purely on the range side.

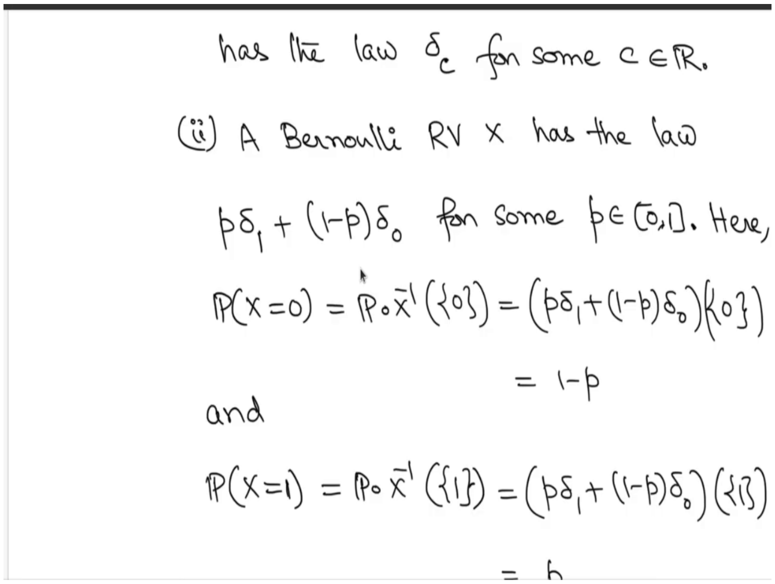
So therefore, we will now again, just to repeat, we will ignore all possible domain side information, we will restrict our attention to the range side, but this way, it is not restricting our informations. We can still get back all the possible information about all possible random variables or all possible random vectors, by simply looking the corresponding law.

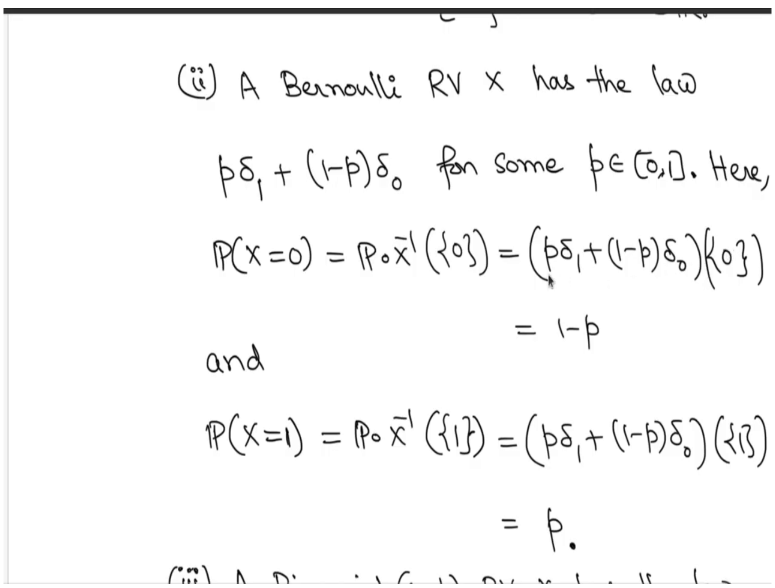
(Refer Slide Time: 14:27)



What we are going to do, is to look at certain special cases of discrete random variables. And we will understand why. So, remember a constant or degenerate random variable has the law given by some appropriate Dirac mass. So, here, c is some constant in the real line.

(Refer Slide Time: 14:48)



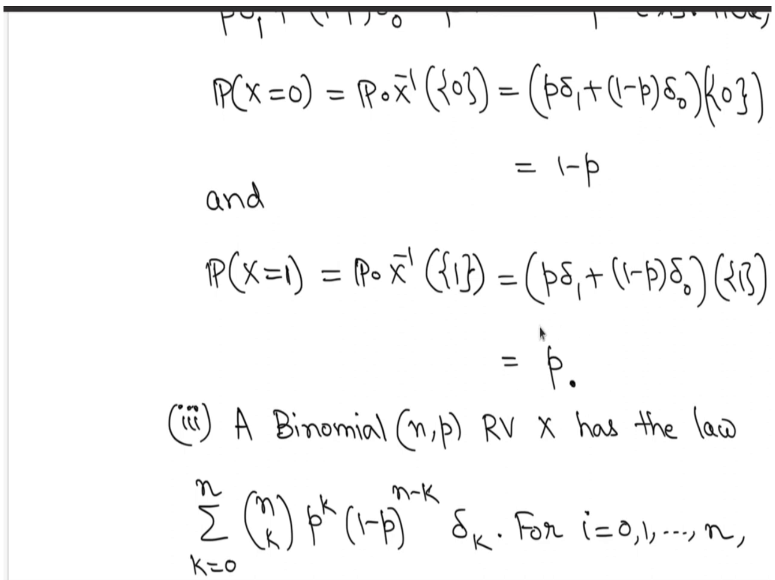


But then if you go to a non-trivial random variable now, random variable taking two possible values, then there is this nice example of a Bernoulli random variable, and it has the law which is given by, now a convex linear combination of the Dirac masses.

So, remember, in the previous lecture we have discussed that our discrete random variable, the law is exactly a convex linear combination of Dirac masses. So, this is a finite linear combination, finite linear combination of Dirac masses, which is a convex linear combination.

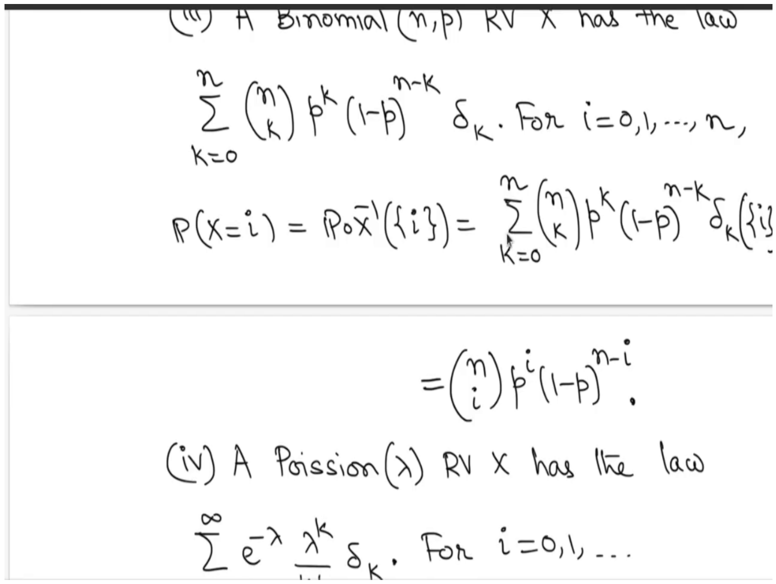
And these p and $1 - p$ exactly refer to the probability of success and failure. So, if you compute $\mathbb{P}(X = 0)$, that is nothing but $\mathbb{P} \circ X^{-1}(\{0\})$, so that is what it means. But $\mathbb{P} \circ X^{-1}$ is this linear combination and you can simplify this, and it will exactly give you this scalar because 1 does not belong to the set 0. So that is all, try to simplify, you will get this scalar.

(Refer Slide Time: 15:48)



Similarly, $\mathbb{P}(X = 1)$ can be computed exactly in the similar way and you can get the value p. So, you get the usual Bernoulli random variable that you have studied in your basic probability courses with property of success p. It is a 0, 1 valued random variable.

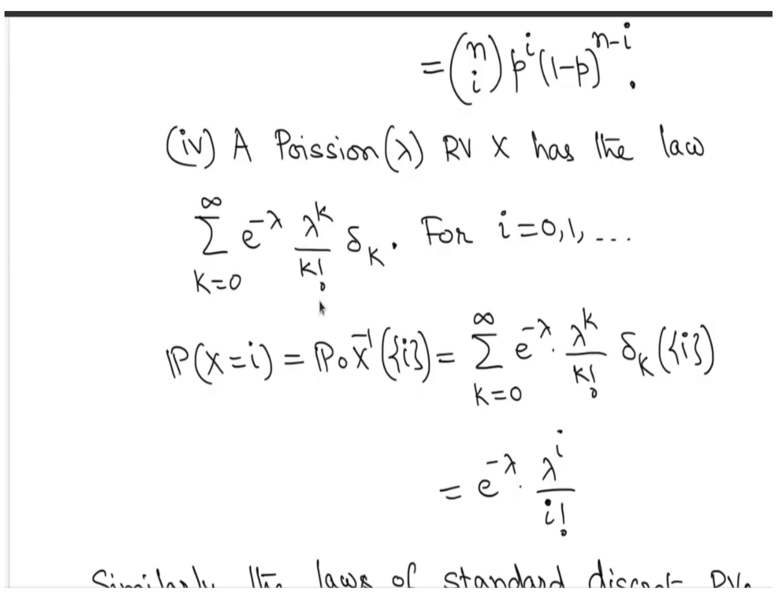
(Refer Slide Time: 16:03)



Similarly, if you go to a Binomial random variable with the n being positive integer and p being a number between 0 and 1, then you get a law which is now convex linear combination of the Dirac masses with these scalars coming from this appropriate Binomial coefficients.

So, here again, if you compute $\mathbb{P}(X = i)$, so that you can simplify and you will exactly get this exact quantity that is the probability mass function for the binomial random variable that you have seen in your basic probability course.

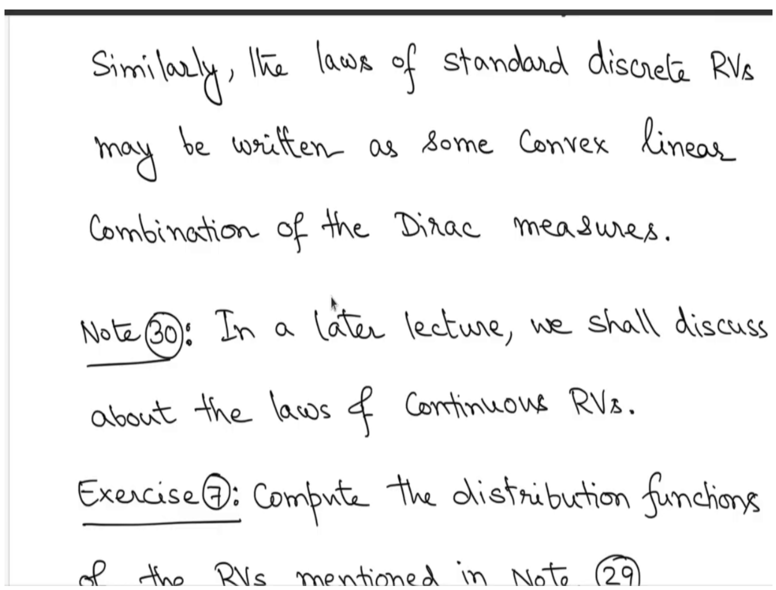
(Refer Slide Time: 16:39)



A more general random variable, for example the Poission random variable with parameter λ has the exact similar description. Here the law is a infinite combination but still convex linear combination of Dirac masses and what happens here again, if you choose these values i from 0, 1, 2, 3, 4, all these non-negative integers, then $\mathbb{P}(X = i)$ exactly appears to be this quantity, which is again the familiar probability mass function for a Poission random variable.

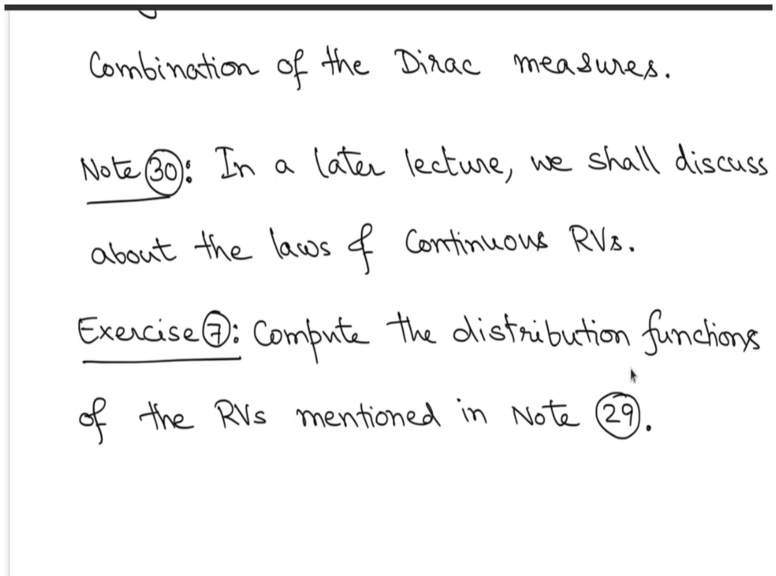
So, now we are going to identify all these random variables through their law and these discrete random variables has their law given by some appropriate convex linear combination of the Dirac masses. So, all these familiar random variables are now very clear. They are appearing though these specific convex linear combinations of the Dirac mass.

(Refer Slide Time: 17:38)



So, other standard discrete random variables, you can still consider, and you just have to identify the appropriate jump values, or the jump sizes and you can write down the corresponding linear combination in terms of the Dirac masses.

(Refer Slide Time: 17:53)



But then, just to repeat this comment from the previous lecture that we are yet to discuss about continuous random variables, and we will come back to this in Week 8. But then, just as a exercise you can now take up the computation of these distribution functions, these well-known distribution functions.

But now, since these are now defined in terms of the law, for these discrete random variables, the Bernoulli case, Binomial case and the Poission case, try to compute and check if you get back the usual distribution functions that you have computed, or that you have seen in our basic probability course. So, try to find this out. So, we will continue this discussion in the next lecture.