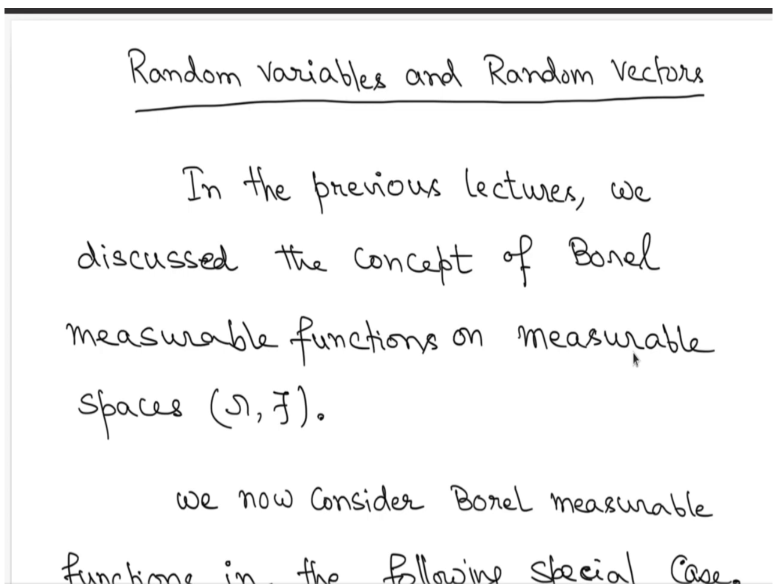
Measure Theoretic Probability 1
Professor Suprio Bhar
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur
Lecture 15
Random Variables and Random Vectors

Welcome to this lecture. This is the final lecture of this week. In this lecture we are going to talk about Random Variables and Random Vectors. Which is the core interest area of this course. So, in this week we have been talking about measurable functions more specifically about Borel measurable functions and their examples. We have also seen algebraic properties and limiting behaviors of such functions.

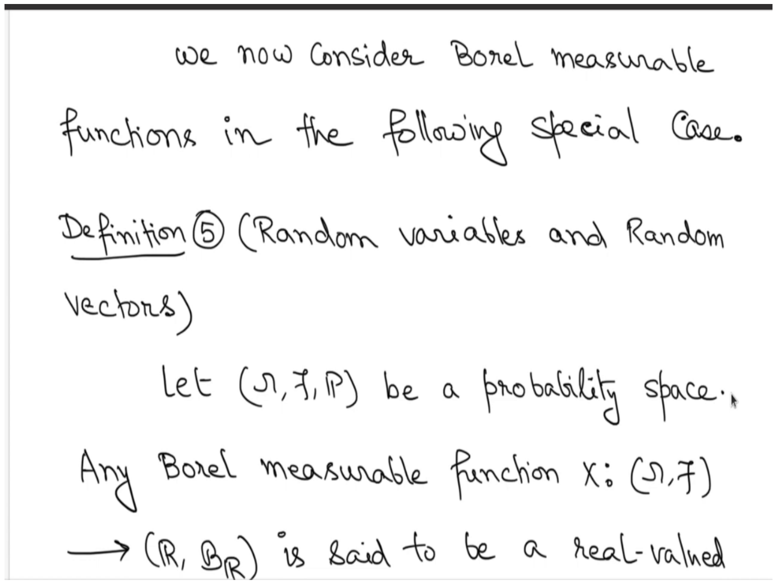
And we have extensively discussed such functions in the previous lectures. In this lecture we are going to talk about these random variables which appear as special classes of measureable functions. As usual let us move on to the slides.

(Refer Slide Time: 1:02)



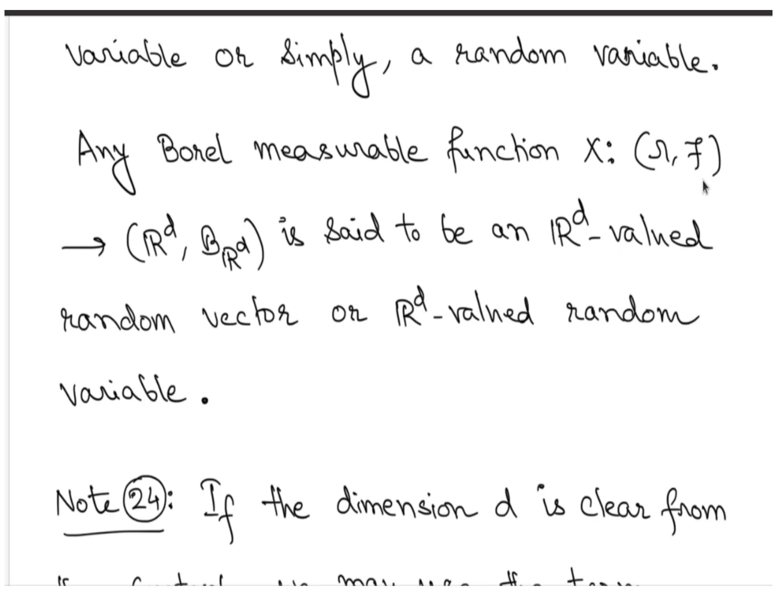
So, in the previous lectures we have discussed this concept of Borel measurable functions on a measurable space. So, we will fix the measurable space beforehand and then talk about the Borel measurable functions be it real valued, extended real valued or $\mathbb{R}^d$ valued.

(Refer Slide Time: 1:20)



Now, we are going to consider these Borel measurable functions in a very specific format. So, what do we do?

(Refer Slide Time: 1:26)



We say that we are going to look at Borel measurable functions *X* from this measurable space to real line. So, suppose you were talking about real valued Borel measurable functions. But then what you want in addition to the structure of measurability of the function. There should be a probability measure available on the domain side. So, the domain side has a probability measure and therefore the domain side is actually a probability space.

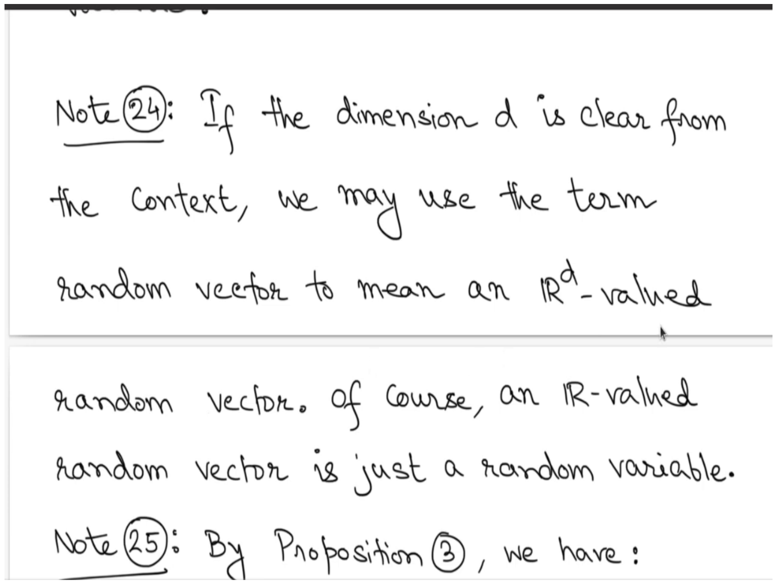
So, the measurable space actually becomes a probability space. But just for the measurability part of the function you do not require that probability measure. But as long as it is there consider any such Borel measurable function. And we are going to say that all such functions are real valued random variables. So, again you have a probability measure in addition to the usual measurable structure.

And as long as you have that presence of the probability measure or appearance of the probability measure you are going to talk about these functions and you are going to refer to them as real valued random variables. In short you can also refer to them as univariate random variables or simply random variables. So, these are all equivalent terms as long as you are interested in real valued measurable functions on top of probability spaces.

And there is of course a natural generalization to higher dimensions. So, if your function is taking values in $\mathbb{R}^d$ it is appropriate dimensional Euclidean space so you put the Borel σ-field on the range side. So, then any such $\mathbb{R}^d$ valued Borel measureable function will be said to be a random vector or $\mathbb{R}^d$ valued random vector or $\mathbb{R}^d$ valued random variable. If there is a probability measure on the domain side.

So again, in all of these structures that we are looking at the probability measure immediately does not play any role in the measurability of the functions. So, for the measurability of the functions X that we consider. We only need to consider the appropriate σ-fields on the domain side and the range side. But as long as the domain side in addition has a probability measure we shall refer to all such Borel measurable functions as random variables or random vectors as appropriate.

(Refer Slide Time: 3:47)



In particular, if the dimension d is clear from the context. So, usually you will fix that in the beginning of your discussion. Then we may use this term random vector instead of saying $\mathbb{R}^d$ valued random vector. So, as long as it is understood that the function is taking values in $\mathbb{R}^d$ you just say that it is a random vector. Of course, if you put dimension $d = 1$. So, that just becomes a real valued random vector and that is nothing but a random variable.

So specifically, there is no difference between the random variables and random vector structurally they are similar. Random vectors are taking values in d-dimensional Euclidean spaces. So, d could be 1, 2, 3 any such positive integer. So, random variables for all our practical purposes these are simply one dimensional random vectors. But then you would like to make these connections a bit more precise.

(Refer Slide Time: 4:44)

random vector is just a random variable.

Note (25): By Proposition (3), we have:

$X = (X_1, X_2, \ldots, X_d)^t$ is a random vector if and only if X_i, $i = 1, 2, \ldots, d$ are random variables.

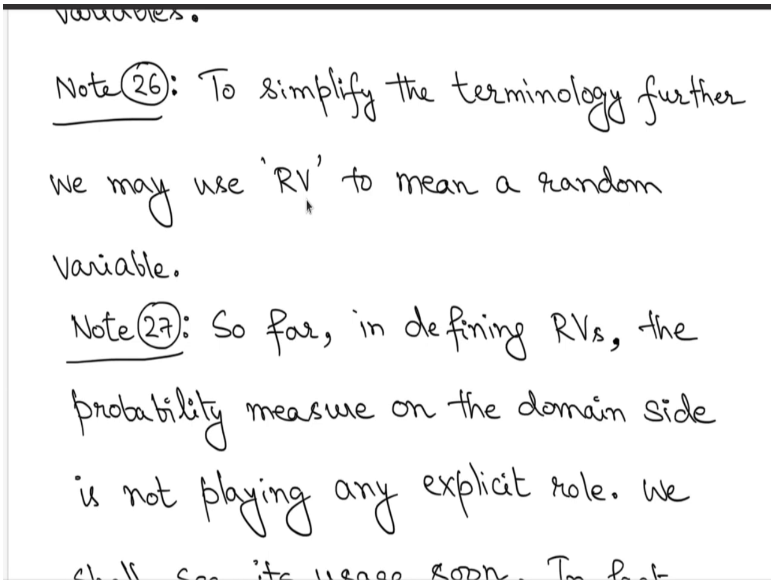
Note (26): To simplify the terminology further we may use 'RV' to mean a random variable

And we had this result in proposition 3 for Borel measurable functions. So, for Borel measurable functions if the function was taking values in $\mathbb{R}^d$. Then you said that the component functions were real valued measurable functions and conversely if the, if you have d many real valued Borel measurable functions if you put them together you get a function taking values in $\mathbb{R}^d$. And therefore, you get a Borel measurable function.

So, that identification we use here say again whatever functions you need as long as it is measurable and as long as there is a probability measure on your domain side you are going to talk about this random variables and random vectors. So, by the same argument as done for the measurable functions you immediately get this result that X which is made up of these d one dimensional component functions is a random vector. If and only if the component functions are also random variables.

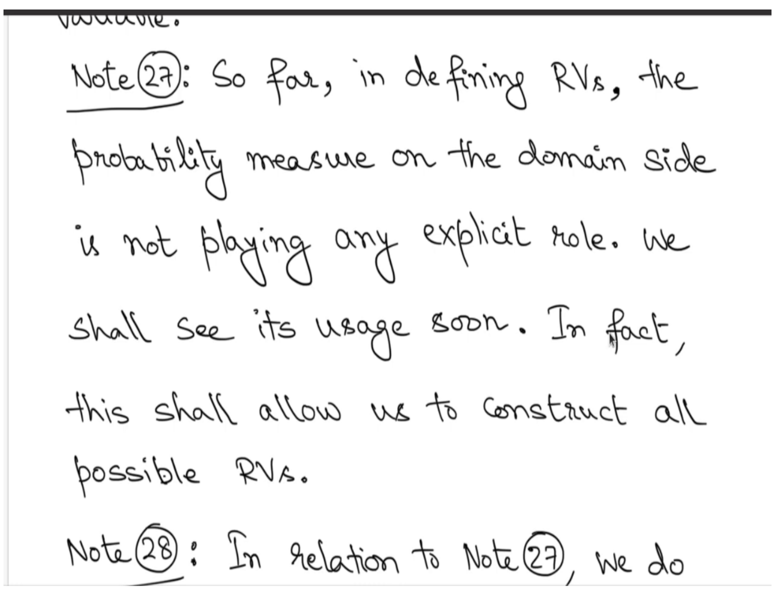
So, that is just about measurability but all these functions are defined on that same probability space. So, in the presence of probability spaces these measurable functions we are referring to them as random variables or random vectors as appropriate.

(Refer Slide Time: 5:56)



And to simplify this terminology further we will simply write our RV to mean random variables. So, instead of saying random variable so we will say RV or we will write RV.

(Refer Slide Time: 6:11)

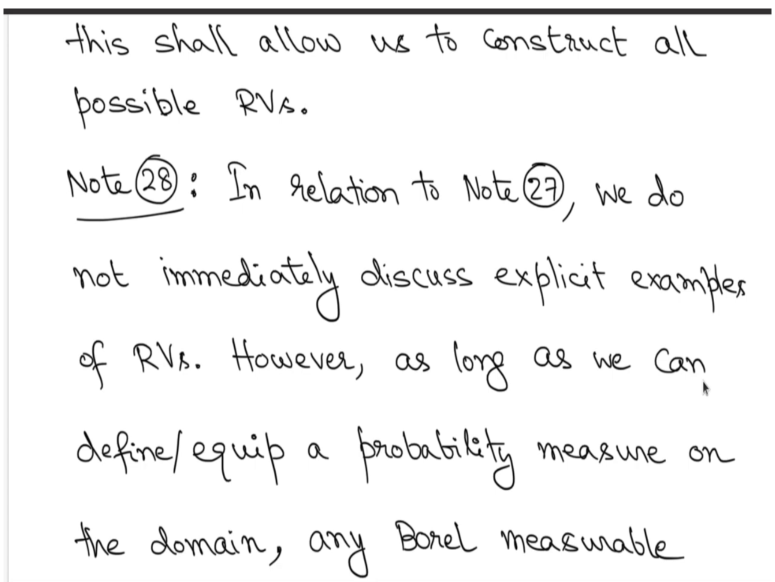


Now, it is this important point about this probability measure. So, in defining this random variable the probability measure a priori there is no explicit rule. So, all you need is a measurable function and to check the measurability you just have to specify the appropriate σ-fields on the domain side and the range side. So, the probability measure is just there. So, but in the presence of this we are talking about this random variables and random vectors.

So, we are going to see the connection of these probability measures with these measurable structures or with these functions that we have been talking about. We are going to see that in later lectures. But for the moment as long as the probability measure is there refer to these Borel measurable functions as random variables or random vectors.

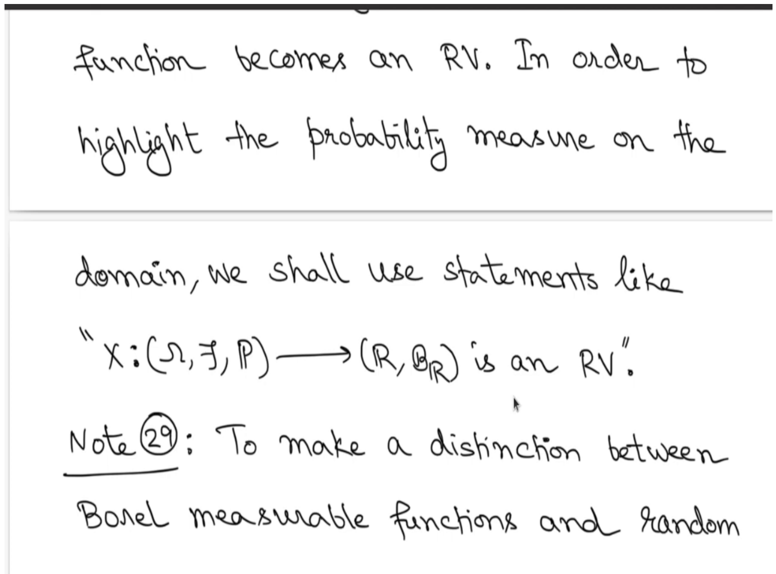
But once we discussed that part, we are actually going to see that with this connection between the appropriate measures and the appropriate measurable functions what you can do is that you can actually construct all possible random variables in one go. So, there is a general construction that will allow us to give you example of all possible random variables. So, for this reason this connection between the probability measure and the measurable structure is very important which will be discussed later on.

(Refer Slide Time: 7:27)



But keeping this in mind that we can construct all possible random variables this way by exploring this connection between the probability measure and the measurable structure of the function what we are actually going to focus on are the properties of the random variables and random vectors. And not immediately focused on explicit examples of random variables or random vectors.

(Refer Slide Time: 7:54)



So, we do not discuss this but again just for the peace of our mind just to understand that what we can do is that we can look at Borel measurable functions. And as long as you can define or equip or probability measure on the domain side you get that this Borel measurable functions becomes random variables or random vectors. So, you first look at some appropriate measurable spaces where you can define this your probability measure. And then any real valued or $\mathbb{R}^d$ valued measurable function becomes a random variable or a random vector respectively.

So, in order to highlight this probability measure in this connection what do we do, we will write the statement that $X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ is a random variable. So, it is important that for the measurable structure we wrote this σ-fields on both sides. But now we are putting this additional information coming from the probability measure on the domain side. So, this is the additional information that we are having.

And so far we have not explored this connection between the measurable structure and the probability measure. But we are going to do that and to do that what we do is that we are going to write this and we will explore this connection later on. But we will keep track of that probability measure on the domain side.

(Refer Slide Time: 9:09)

Note (29): To make a distinction between Borel measurable functions and random variables, we use small letters f, g etc. to denote the Borel measurable functions, and Capital letters X, Y etc. to denote random variables/vectors.

Using Theorem (1), we have the

But now it is an important point that there is a distinction between Borel measurable functions and random variables. So, if you have random variables, then they are Borel measurable functions defined on probability spaces, but a priori to talk about measurable functions, you do not require probability spaces.

So, in general, Borel measurable functions are defined on some appropriate measurable spaces. There need not be a measure on your domain set. So, therefore, all random variables, by definition are Borel measurable functions, but not all Borel measurable functions are random variables. So, therefore, your measurable space that you are talking about, if you do not equip a measure, or if you equip a different measure than a probability measure.

If your measure that we equip, that is not a probability measure, then you cannot say that, that Borel measurable function is a random variable, correct? So therefore, all random variables are Borel measurable functions, but not all Borel measurable functions are random variables and to distinguish this fact what do we do, we use small letters, f, g to denote these functions, Borel measurable functions and for random variables and random vectors, we shall use this capital letters, X, Y, etc.

(Refer Slide Time: 10:27)

Using Theorem 1, we have the following result.

Theorem 5: Let X and Y be RVs defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $X+Y$, $X-Y$, XY, $|X|$, $X \wedge Y$, $X \vee Y$, X^+, X^- are also RVs defined on the same probability space.

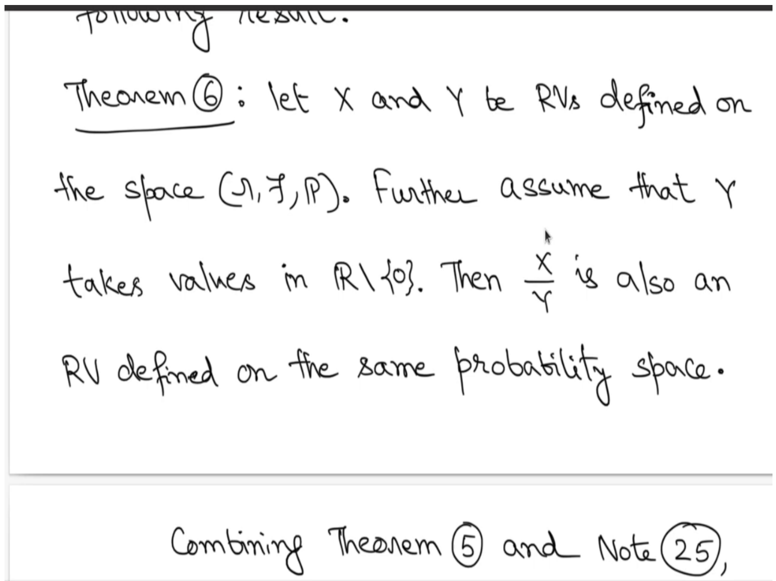
Theorem 5: Let X and Y be RVs defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $X+Y$, $X-Y$, XY, $|X|$, $X \wedge Y$, $X \vee Y$, X^+, X^- are also RVs defined on the same probability space.

Using Theorem 2, we have the following result.

So now, earlier we talked about these algebraic operations between measurable functions and since random variable are also Borel measurable functions. So, we go back to this theorem on where we discuss these algebraic properties and rewrite these properties. So, take $X + Y$ and Y to be random variables defined on this probability space, then you can talk about the addition, subtraction and so on.

These are just Borel measurable functions. So, that is exactly what we are looking at. But then these are also Borel measurable as per theorem 1 considered earlier, just edition of Borel measurable function is Borel measurable and so on. But these functions $X + Y$, for example, they are defined on the same probability space, and therefore, we are going to refer to them again as random variables. So, these $X + Y$, $X - Y$ and so on, all of these are Borel measurable functions defined on this given probability space and therefore, you should refer to them as random variables.

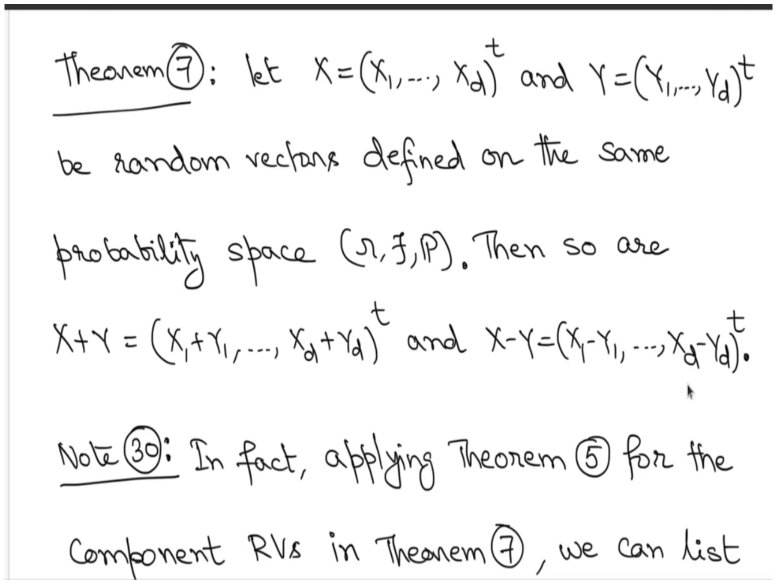
(Refer Slide Time: 11:25)



But then you can also talk about the division operation. So, in the algebraic operation considered earlier we looked at addition, subtraction, multiplication. Then certain maximum minimum operations and so on. But then here what we are doing, we are looking at the division operation. So, again keeping track of the result in measurable functions in mind. So, there what we do the function that we divide by so here we are dividing by Y that should not take the value 0.

So, as long as it avoids the value 0 it takes the values in $\mathbb{R}\setminus\{0\}$ then you can talk about the function $\frac{X}{Y}$ as long as X takes the value 0 the division can take the value 0. So, therefore $\frac{X}{Y}$ is a real valued Borel measurable function. But then $\frac{X}{Y}$ is still defined on the same probability space and therefore you get a random variable. So, as long as your Y the random variable that you are dividing by that appears in the denominator does not take the value 0 you can talk about $\frac{X}{Y}$. So, this is a genuine random variable once more. So, this is the opposition division appearing here.

(Refer Slide Time: 12:31)

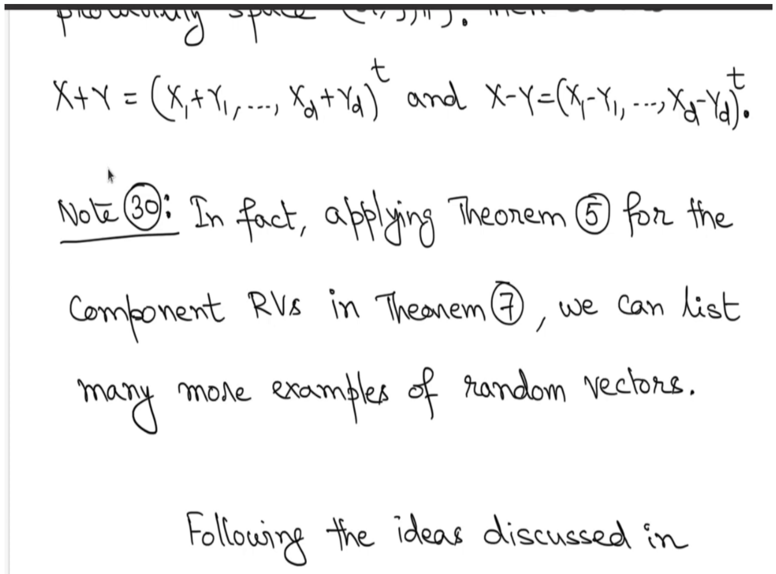


So, with these things in mind let us talk about things in higher dimensions. For random vectors the identification that we made earlier was that the component functions are random variables. So, if you take a random vector X taking values in $\mathbb{R}^d$ then you have d many $\mathbb{R}$ valued real valued components each of which are random variables. So, then take such two random vectors X and Y taking values in $\mathbb{R}^d$ then you can construct other examples of random vectors.

How? So, you can think of this component wise operations. So, if you do these componentize additions these are again Borel measurable since $X_1 + Y_1$ is again a random variable. Therefore, if you put them together so $X_1 + Y_1$, $X_2 + Y_2$, $X_d + Y_d$ each of these are random variables. So, if you put them together you still get $X + Y$ that is a d dimensional random vector.

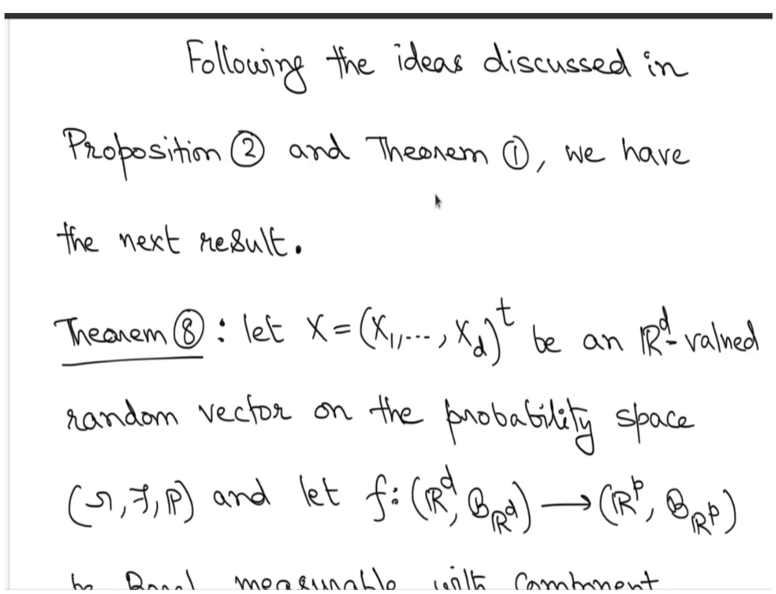
So, with this kind of operations you can also consider subtractions. So, again all of these are going back to the individual real valued components. The coordinate components, the coordinate wise functions that you get are giving you all these operations all these properties. So, therefore you can continue on and construct more examples of random vectors from given examples.

(Refer Slide Time: 13:51)



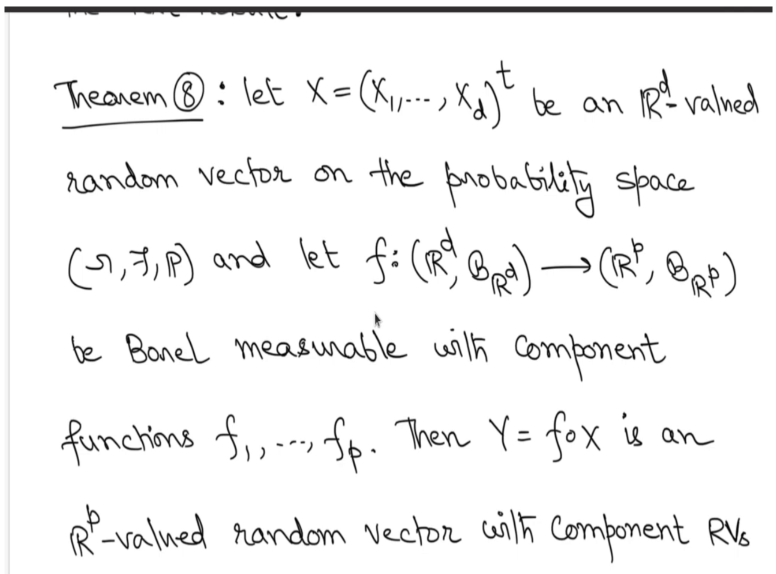
So, that is mentioned in note 30 that in fact applying these operations, algebraic operations you can get much more examples. For example, one of the operations that we skipped in this previous result was multiplication. So, if you choose to do coordinate-wise multiplication so you will get in the first coordinate you will get X_1Y_1, second coordinate you will get X_2Y_2 and so on. So, that will also be a random vector because individual components are random variables. So, all of these are defined on the same probability space. So, this is quite important.

(Refer Slide Time: 14:25)



Now, what we have discussed earlier in the algebraic operations was that these algebraic operations will keep measurability operations fixed. So, we have already discussed these algebraic operations involving this addition, subtraction, multiplication, division and so on. One of the important operations that we have also discussed there earlier in proposition 2 and theorem 1 earlier was the composition operation.

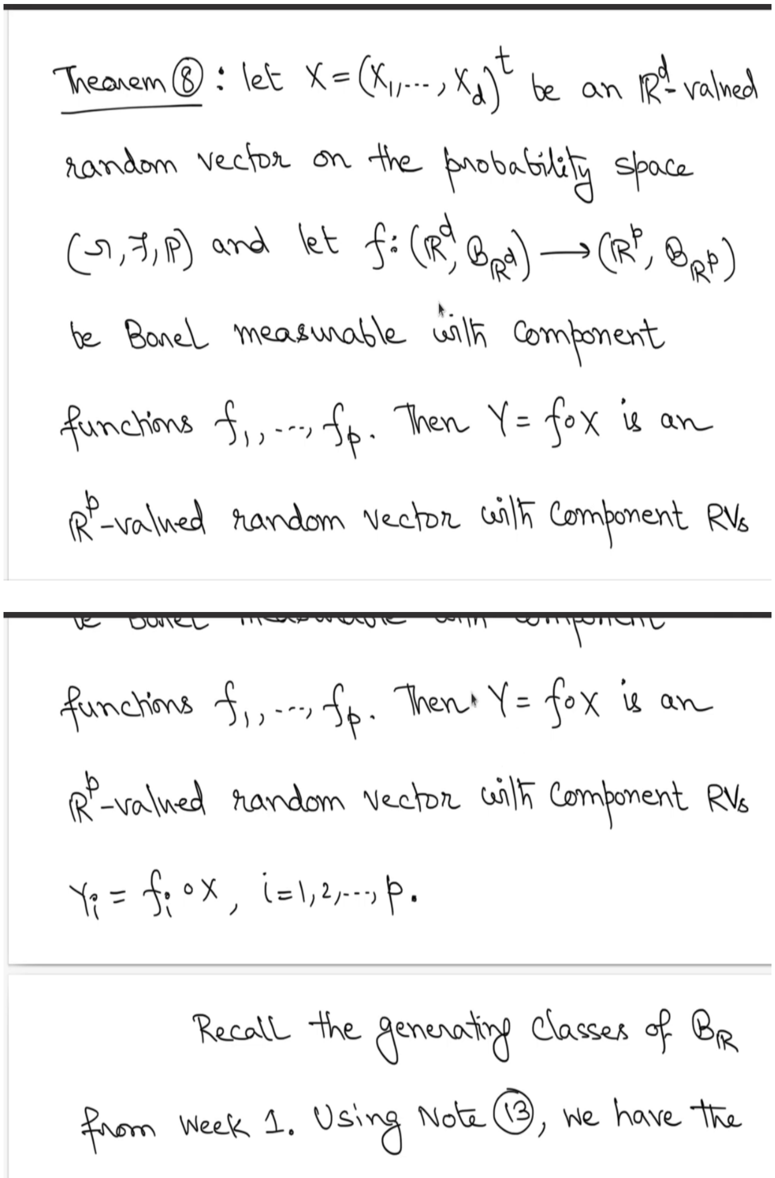
(Refer Slide Time: 14:51)



So, this is what we revisit for random variables and random vectors. So, again here if d equals to 1 this statement can be read for random variables. So, how so take X to be a random vector taking values in $\mathbb{R}^d$. Then if you look at a function, a Borel measurable function from $\mathbb{R}^d$ to $\mathbb{R}^p$ let us say. So, you go to a possibility dimension p could be d. So, it will remain in the same dimension if p is different you go to another dimension.

Whatever it is the range of X is $\mathbb{R}^d$ and the domain of f that function the Borel measurable function is $\mathbb{R}^d$. So, domain of f contains the range of X so you can consider the $f \circ X$.

(Refer Slide Time: 15:35)



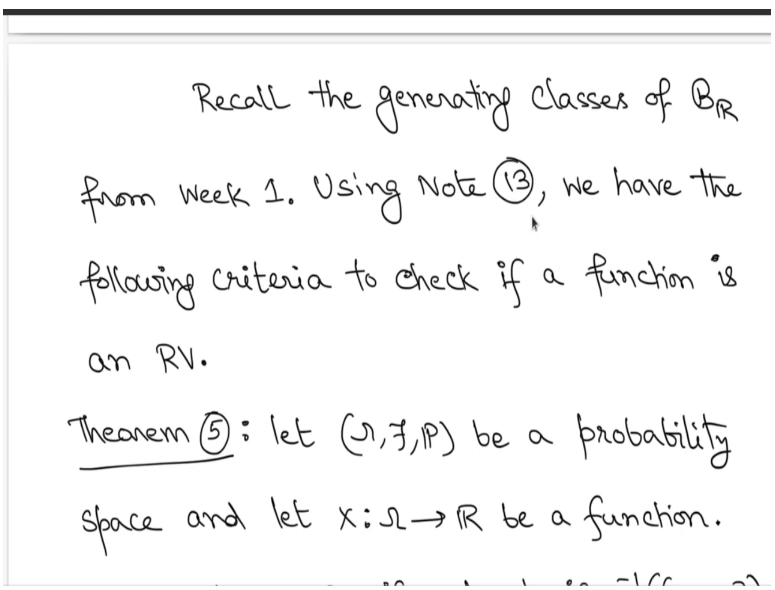
So, now what you can do is that you can look at this composition function. So, that is again measurable because f is measurable and X is measurable. So, the composition preserves measurability but you have to use the appropriate σ-fields on the domain side and the range side. So, here the domain is the probability space $(\Omega,\ \mathcal{F},\ \mathbb{P})$ and the range is $\mathbb{R}^p$ with the Borel σ- on $\mathbb{R}^p$.

So therefore, Y will become a $\mathbb{R}^p$ valued random vector but then you can also identify the individual component random variables. How? You look at the component functions of the given function f. So, f was a $\mathbb{R}^p$ valued function. So, it will have this p many $\mathbb{R}$ valued components. So, which we write it as $f_1, \ldots, f_p$. So, each of these with f_i's, $f_1, \ldots, f_p$'s are defined on $\mathbb{R}^d$ and takes values in $\mathbb{R}$. So, these are all Borel measurable components.

So, if you want to identify the components of the random vector Y which is $\mathbb{R}^p$ valued you will get p many components which are actually given as follows. So, Y_i's will be given as $f_i \circ X$. So, $X: \Omega \to \mathbb{R}^d$ and $f_i: \mathbb{R}^d \to \mathbb{R}$. So, therefore this composition goes from $\Omega \to \mathbb{R}$. So, Y_i's are real valued random variables you have p many components of the random vector Y.

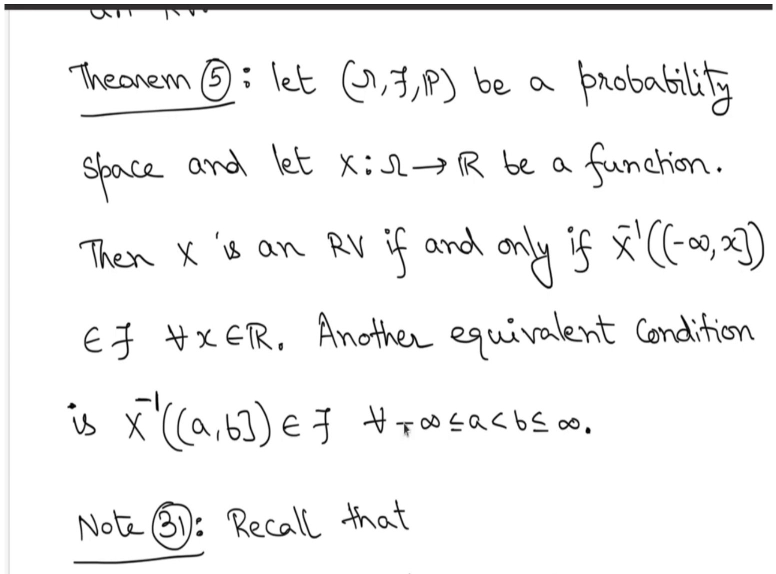
So again, in all of this what we are doing we are ensuring that we are working on the same probability space. So, as long as you apply some nice functions on random variables or random vectors you end up having other examples of random variables or random vectors.

(Refer Slide Time: 17:12)



Here an important comment about the measurable structure is required. So, you look at the generating classes of the Borel σ-field. So, if you are considering random variables, real valued random variables. Then the generating classes will help you identify about the measurability of that function. So, this was discussed earlier in note 13. But then if you are talking about these random variables.

(Refer Slide Time: 17:42)



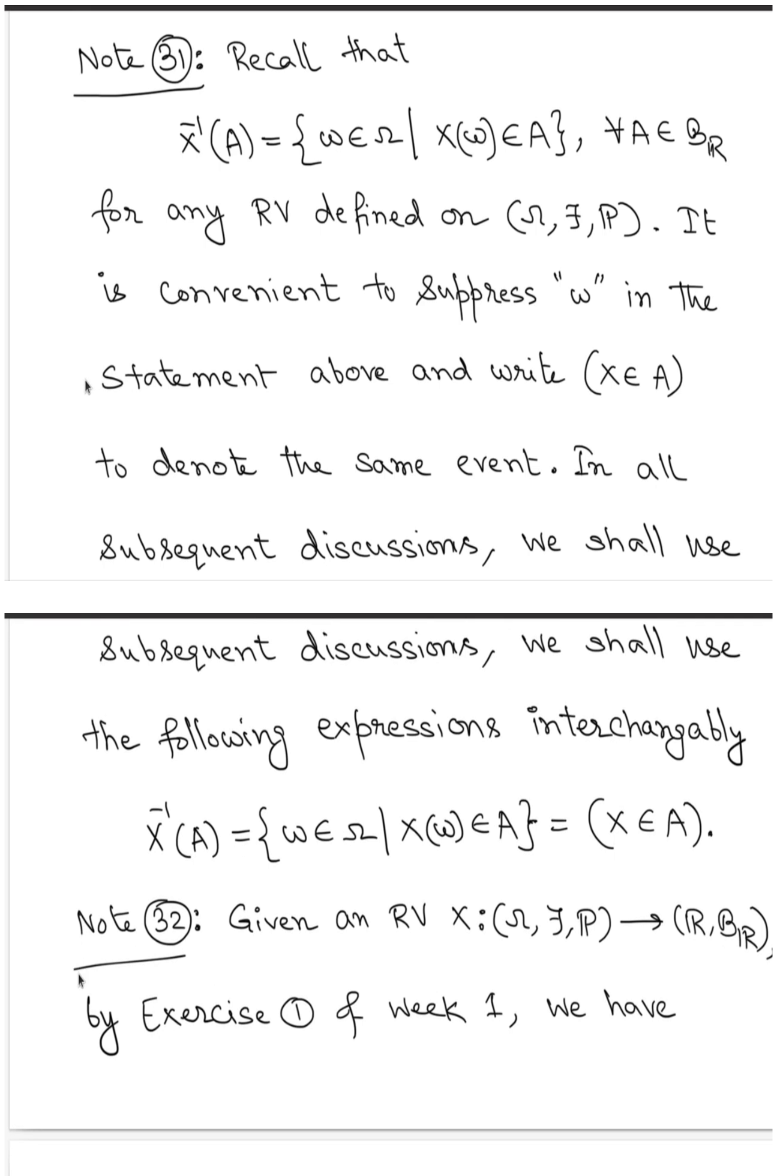
I am talking about the measurability conditions. So, you can restate that result for measurable functions for these random variables. So, what do you do? You take a function X defined on a probability space. Then X is a random variable meaning it is measurable with respect to the appropriate σ-fields if and only if the preimage of all this type of sets are in the domain side σ -field.

So, as long as you can ensure that the function X is measurable you can call it a random variable since there is a probability measure on the domain side. But then you just look at this type of intervals minus infinity to close x. So, these types of intervals for all x will generate the Borel σ -field on the real line. So, therefore if all these preimages are in the domain side σ-field you immediately claim that X is measurable.

But on the other hand, if X is measurable of course all these preimages are of course in the domain side σ-field. So, therefore this is an if and only if condition that was discussed earlier for measurable functions but then we are just restating it for random variables. Of course, you can choose to look at other types of generating sets. For example, you might look at this left open right close intervals for $a < b$, $a,\ b$ two real numbers possibly $a,\ b$ also could be ∞ or $-\ \infty$, the interpretations has been discussed in week 1.

So again, if the preimages for such sets are in the domain side σ-field you will say that X is a random variable and this is an equivalent condition. So, again these conditions help you identify the measurable structures. And these are all following from the discussions that we did for usual Borel measurable functions. Now we are going to look at some special structures of these preimages.

(Refer Slide Time: 19:24)



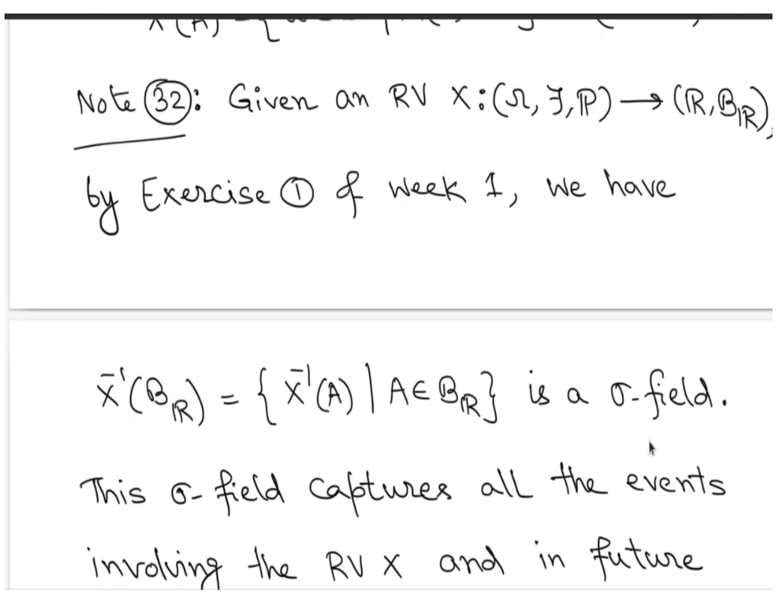
So again, if X is a random variable, it is a Borel measurable function and take a look at all possible Borel sets on the range side. Then their preimages are on the range side σ-field. So, what are these preimages once more. So, this we chose to write using this notation $X^{-1}(A)$ but that actually means these subsets of Ω defined this way that you collect all those sample points, ω such that $X(\omega) \in A$.

So, that is the description of that preimage of the set A under the mapping X. Now, what you can choose to do is to suppress this symbol or this notation ω from this expression. So, if you remove this ω so then this whole part gets removed then you remove the ω from here also. Then what you get is this notation at the end you put some appropriate brackets on the outside just to identify this condition separately.

So, then what you do is that you remove all this occurrence of ω in this notation and simplify it to this notation. So, what happens is this that you get a commonly used notation in probability theory. So, these preimages are as per definitions are these sets but for convenient notations what we do we choose to suppress this ω and just write $X \in A$ this notation. So, this will denote the exactly the same events, this will denote the exactly the same preimages.

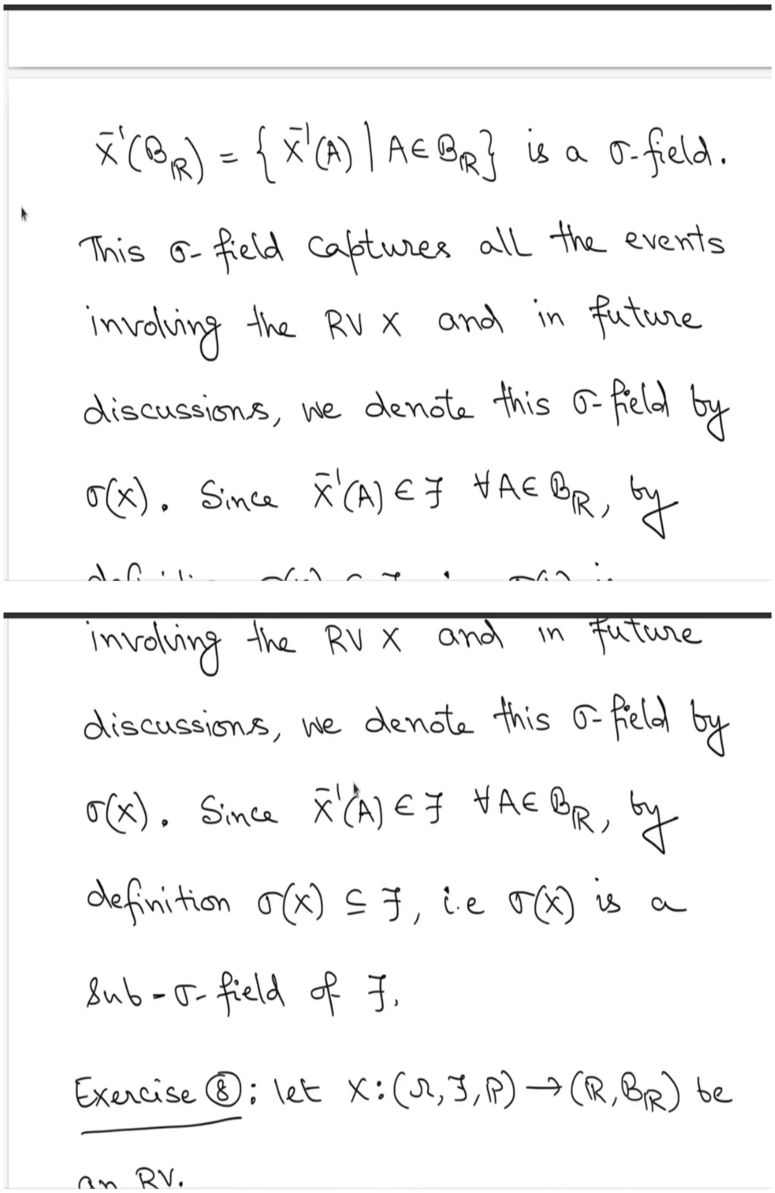
So, in all subsequent discussions just to be clear all these things mean the same event. Here $X^{-1}(A)$ means you look at the subset of Ω. Such that $X(\omega) \in A$ but this also is denoted by the same notation this thing $\frac{X}{Y} \in A$ this is your familiar notation involving events. But then there is this interesting structure for these preimages.

(Refer Slide Time: 21:37)



So, you start with this random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, then by exercise 1 if you collect all possible such preimages. So, this exercise 1 of week 1 then if you choose or if you collect all such preimages then you will get a σ-field on the domain side. So, if you collect all the preimages then you get a σ-field on the domain side.

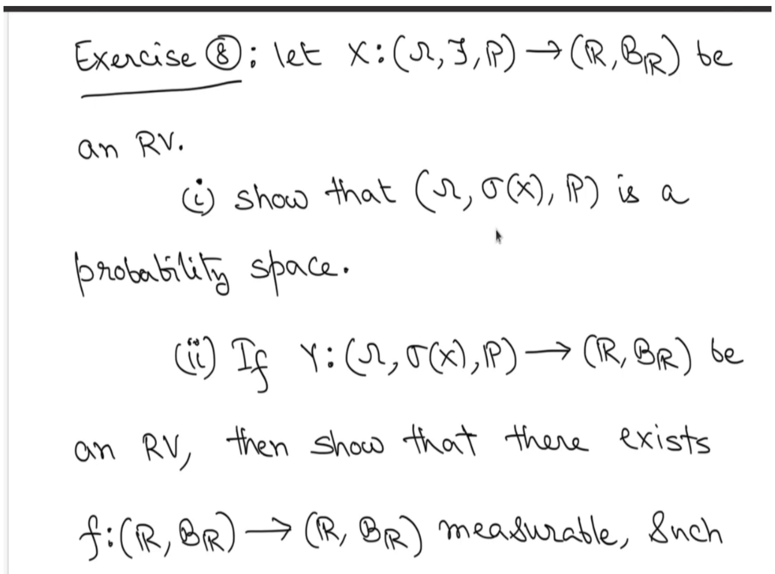
(Refer Slide Time: 21:59)



These σ-field captures all the events involving the random variable X. And then in future what we are going to call this σ-field as $\sigma(X)$. So, this is the notation that we are going to use to denote this σ-field. Now since $X^{-1}(A)$ belongs to this general σ-field that we have started off with for all possible Borel sets on the range side. Then this collection that you have now constructed this is a sub collection of this original given σ-field scripted F.

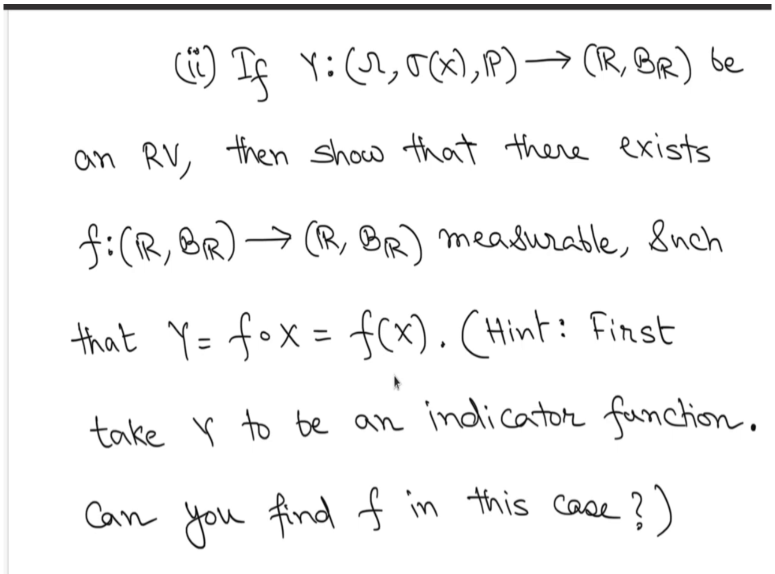
So, $X^{-1}(A)$ was already there in this σ-field that you have started off with you have now constructed a smaller σ-field. And therefore, this $\sigma(X)$ is a sub σ-field of scripted F. This sub σ-field $\sigma(X)$ contains all the events or all the information involving the random variable X.

(Refer Slide Time: 22:54)



You can now choose to check all these nice properties. So, for example you can choose to check that if you continue to use the same probability measure then this becomes a probability space. So, if you replace the σ-field by $\sigma(X)$ here just choose to look at the probability of these events that you get in this smaller class of σ-fields you still get a probability space, please check this.

(Refer Slide Time: 23:21)



Another interesting example. Another interesting property is this that you can now choose to look at functions that are defined or that are measurable with respect to this new smaller σ-field. So, take such Borel measurable functions we of course call them a random variable because this is a probability measure as per the first part. Then what you can show you can construct some functions Borel measurable functions ℝ to ℝ such that you can write Y as $f \circ X$ or a function of X.

So, you can explicitly construct such functions. So, originally if Y is measurable with respect to this σ-field coming from X itself you can write Y as a function of X. So, that is the statement of this exercise. So, the hint here is that you should start with Y to be an indicator function, try to construct that function f explicitly it should be a function from ℝ to ℝ. And then try to move on to simple functions non negative measurable functions and then to general measurable functions.

So, try to figure out the function for the case of indicator functions and it will be easy job afterwards. So, with this we stop the discussion here. So, we will continue the discussions in the next week's lectures.