

Measure Theoretic Probability 1
Professor. Suprio Bhar
Department of Mathematics & Statistics
Indian Institute of Technology, Kanpur
Lecture No. 07
Probability Measures

Welcome to this lecture. Let us first recall what we have done so far. So, in the previous lecture, we had considered how to look at sizes of sets in a σ -field. And the motivation was that this will help us understand which events are more likely to occur or which sets in a σ -field are more important and this we are checking by the associated numerical value and if the value is larger, then we will say something is more important or more likely to happen than the other.

So, with this way or with this motivation, what we did we define something called set functions and then we understood that for disjoint sets, pairwise disjoint sets, we should be able to add up the sizes. And with that motivation, we looked at the spatial plus of set functions which we called as measures.

So, what were measures? Measures were non-negative countably additive set functions on σ -fields. But then we clarified that these measures may take the value plus infinity. And with that, we made this very specific assumption that there is your set with finite mass or finite size and that allows us to prove that measure of empty set is 0. And using that we had to prove the finite additivity of the measures and obtain certain other basic estimates involving the measures of sets or the sizes of sets. So, with that in mind, we now start over and move on to the slides for this lecture.

(Refer Slide Time: 2:09)

Probability Measures

In the previous lecture, we have discussed the notions of measures and measure spaces. Recall from Proposition 3 of the previous lecture that if μ is a measure on a measurable space

measure spaces. Recall from Proposition 3 of the previous lecture that if μ is a measure on a measurable space (Ω, \mathcal{F}) , then $0 \leq \mu(A) \leq \mu(\Omega) \forall A \in \mathcal{F}$. with this in mind, we now look at a special subclass of finite measures.

So, in the previous lectures, these notions of measures and measure spaces were discussed. Now, let us recall this important property that we proved almost at the end of the last lecture. So, this was this fact that if μ is a measure on a measurable space, then you pick up any set from your σ -field, look at the size of that. So, $\mu(A)$ is the size of that set according to that set function way of measuring. So, that is dominated from above by the size or $\mu(\Omega)$. So, Ω is the whole set.

So, what it says is that the maximum value achieved by this function μ is achieved at the whole set. And this is the maximum possible. And with this in mind, we are now going to look at the special subclass of finite measures. So, just to recall, we had looked at these two specific

subclasses of measures. One was the case when $\mu(\Omega)$, the size of Ω was infinite, we will call it an infinite measure. And the other one where $\mu(\Omega)$, the size of Ω was finite and we called it a finite measure.

(Refer Slide Time: 3:29)

Definition 6 (Probability measure and Probability space)
If μ is a measure on (Ω, \mathcal{F}) with $\mu(\Omega) = 1$, we say that μ is a Probability measure and refer to the triple $(\Omega, \mathcal{F}, \mu)$ as a Probability space.

So, we are now going to look at this special surplus of finite measures and this allows us to make this next definition. So, what is this? So, we define that μ , a measure μ is a probability measure if $\mu(\Omega)$ is 1. So, we are looking at again the size of the whole set according to the set function μ , according to the measure μ . And we are saying that if $\mu(\Omega)$ is equal to 1, then we are going to say that μ is a probability measure on (Ω, \mathcal{F}) that measurable space.

Now, as we have seen earlier, when you have a measure on a measurable space (Ω, \mathcal{F}) , then you can consider the triple. So, take all these three things together, you get a triple, and we called it a measure space earlier. But now, if it so happens that μ is a probability measure, then we shall refer to this triple, this measure of space as a probability space. And we are going to work on these kind of triples, these kind of probability spaces throughout the course.

So, again, if μ is a measure with the fact that $\mu(\Omega) = 1$, then we are going to call it a probability measure and we are going to look at the corresponding triple, corresponding measure space and also referred to it as a probability space. Now, as soon as we make the definition, it is a good idea to look at some basic examples.

(Refer Slide Time: 4:59)

Example (Dirac measure on \mathbb{R})

Fix $x \in \mathbb{R}$. Consider the non-negative set function δ_x on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ defined by

$$\delta_x(A) := \begin{cases} 0, & \text{if } x \notin A \\ 1, & \text{if } x \in A \end{cases} \quad \forall A \in \mathcal{B}_{\mathbb{R}}$$

$$\delta_x(A) := \begin{cases} 0, & \text{if } x \notin A \\ 1, & \text{if } x \in A \end{cases} \quad \forall A \in \mathcal{B}_{\mathbb{R}}$$

check that δ_x is a measure. (Exercise)

Since $\delta_x(\mathbb{R}) = 1$, δ_x is a Probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. We refer to δ_x as the Dirac measure supported at $x \in \mathbb{R}$.

And to do that the first example and this is a very important example, in fact, as we shall find out is called the Dirac measure on the real line. So, again, just to recall the terminology that was introduced in the previous lecture, so we will call that a measure is on a measurable space, as long as the σ -fields are understood, we can as well say that the measure is on the domain Ω or equivalently we can also refer to the σ -field.

So, measure is actually defined on the σ -field. So, we can also refer to it as that measure on the corresponding similar field \mathcal{F} . So, on the real line, most often we are going to work with the Borel σ -field. So, if it is not stated, always assume that the Borel σ -field is taken on the real line.

So, when you are considering the real line, if the σ -field is not clearly stated, please consider the Borel sigma.

Now, what we do, we are going to define a set function which we write it as this δ_x . So, we are going to fix this real number x first and going to look at the corresponding set function which we are going to define now. So, this is δ_x . So, this is going to be a set function that is defined on this measurable space. So, in fact, it is a set function defined on the Borel σ -field. So, how do you define this? It is pretty simple.

So, the point x is fixed for you a priori and then you vary your Borel sets A over the Borel σ -field. So, now, you are going to check this whether your point x is in the set or not. If the point x is there in your set, then I assign the value 1. We say, yes, the point is there in the set, otherwise, we assign the value 0. So, this is simply 0, 1 valued set function. So, therefore, it is also a non-negative set function by definition correct. So, just takes values 0 or 1.

Now, this is pretty simple looking set function, but it turns out this actually happens to be countably additive. And therefore, it becomes a measure on the Borel σ -field on the real line. So, therefore, this is a very important example of a measure which simply checks whether the parametrized point, that x is in the arbitrary set A or not. If it is there, then you put 1, otherwise it is 0. So, please check that this set function is countably additive and that will imply that it is a measure. So, this is left as an exercise for you. Please check this.

But as soon as it becomes a measure, you can now try to look at the size of the whole set, which is the real line here. So, once you look at the size of the whole set, all you have to do is follow the definition. Is the point x in the real line? Yes, the answer is yes, because x was chosen like that. So, therefore, the answer to this value is 1. So, therefore, this becomes a example of a probability measure. So, as soon as a measure is with the property that the measure of the whole set is 1, then it automatically becomes a property measure as per our definition earlier.

So, therefore, you have a very nice example of a probability measure on the real name. We are going to refer to this measure as the Dirac measure supported at the points x or δ_x . So, that is the Dirac measure supported at x .

(Refer Slide Time: 8:34)

as the Dirac measure supported at $x \in \mathbb{R}$.

Exercise 1: In an analogous fashion, we may also define δ_x on a general measurable space (Ω, \mathcal{F}) for any fixed $x \in \Omega$ as

$$\delta_x(A) := \begin{cases} 0, & \text{if } x \notin A \\ 1, & \text{if } x \in A \end{cases} \quad \forall A \in \mathcal{F}.$$

check that δ_x is a Probability measure

may also define δ_x on a general measurable space (Ω, \mathcal{F}) for any fixed $x \in \Omega$ as

$$\delta_x(A) := \begin{cases} 0, & \text{if } x \notin A \\ 1, & \text{if } x \in A \end{cases} \quad \forall A \in \mathcal{F}.$$

check that δ_x is a Probability measure on (Ω, \mathcal{F}) .

Now, in an analogous fashion, you can actually extend these notions of Dirac measures on general measurable spaces. How? So, all you have to do is that start with some non-empty set Ω , take a significant \mathcal{F} then for any fixed set x in your domain, look at the set function Dirac or the δ_x , for any arbitrary set coming from the single field, you just check whether the point $x \in A$ or not. If it is there, you assign the value 1, otherwise 0, exactly as done before. And it will happen that Dirac subscript x or δ_x , that is the Dirac measure, it will turn out to be a probability measure.

So, again, please check that this is countable additive and the size of the whole set Ω now is also 1. Please check this. So, this is left as an exercise for you. Please check this. So, the motion of

Dirac measures can be extended to arbitrary measurable spaces, but most often we are going to restrict our attention to the real line as defined earlier. So, this is on the real line. So, we are going to look at these kind of structures throughout the course.

(Refer Slide Time: 9:46)

we now discuss two propositions,
which allow us to construct examples of
Probability measures.

Proposition ④: let μ be a finite measure
on $(\mathcal{R}, \mathcal{F})$. Consider the set function ν :
 $\mathcal{F} \rightarrow [0, \infty)$ defined by

$\mathcal{F} \rightarrow [0, \infty)$ defined by

$$\nu(A) := \frac{\mu(A)}{\mu(\mathcal{R})}, \forall A \in \mathcal{F}$$

is a Probability measure on $(\mathcal{R}, \mathcal{F})$. (Here,
implicitly, we assume $\mu(\mathcal{R}) > 0$).

Proof: By definition, ν is a non-negative

Now, with this example at hand, one concrete example at hand, we are now going to look at certain properties of these probability measures, which will now allow us to construct more examples. So, once you have some explicit examples at hand, like the Dirac measure, you can now use certain properties of probability measures, mix them up and construct more examples. So, to do that, we state two propositions.

So, the first proposition says that you take a finite measure first. You take a finite measure, call it μ . So, this is defined on this measurable space (Ω, \mathcal{F}) . Consider the set function. This we are now, as ν . So, originally, the measure was μ , which was a finite measure. Now I am calling the new set function that I am going to define as ν . So, how it is going to be defined? So, it still defined on the same σ -field that I have started off with. But then I am going to look at the values of the sets or the sizes of the sets, according to the new measure defined like this.

So, what do we do? We look at the size of the arbitrary set A with respect to the original measure μ and divide it by that total measure of Ω . So, here, we automatically have to assume that $\mu(\Omega)$ is strictly greater than 0, otherwise, this division is not well defined. So, we are not allowed to divide by 0. So, for a finite measure, as we have started off with, we are going to assume that $\mu(\Omega)$ is strictly positive, is positive, it is not 0 and we are going to divide or look at the ratio of the size of the set with respect to the whole set. So, that is the ratio I am looking at.

And automatically this value is less or equal to 1 because $\mu(A)$, the maximum value it can take is $\mu(\Omega)$ itself that we have proved earlier. So, this ratio is falls between 0s and 1s. So, $\nu(A)$ as defined actually takes values between 0s and 1s. Now, all you have to check that this is a product dimension. So, this is the claim of the proposition. So, let us try to go through this.

(Refer Slide Time: 11:58)

implying, we assume $\mu(\Omega) > 0$.

Proof: By definition, ν is a non-negative set function. To check countable additivity, let $\{A_n\}_n$ be a sequence of pairwise disjoint sets in \mathcal{F} . Then

$$\sum_{n=1}^{\infty} \nu(A_n) = \frac{1}{\mu(\Omega)} \sum_{n=1}^{\infty} \mu(A_n) = \frac{\mu\left(\bigcup_{n=1}^{\infty} A_n\right)}{\mu(\Omega)}$$

sets in \mathcal{F} . Then

$$\begin{aligned} \sum_{n=1}^{\infty} \nu(A_n) &= \frac{1}{\mu(\Omega)} \sum_{n=1}^{\infty} \mu(A_n) = \frac{\mu\left(\bigcup_{n=1}^{\infty} A_n\right)}{\mu(\Omega)} \\ &= \nu\left(\bigcup_{n=1}^{\infty} A_n\right). \end{aligned}$$

Hence ν is a measure on (Ω, \mathcal{F}) . Moreover

$$\nu(\Omega) = \frac{\mu(\Omega)}{\mu(\Omega)} = 1. \text{ This completes the proof.}$$

So, by definition, this set function ν is non-negative. Why, because μ , the original finite measure that was non-negative and you are dividing by this positive quantity, so the ratio is, of course, non-negative. So, that is not a problem. So, the next thing that we need to check is countable additivity. So, that will allow us to say that ν is a measure. So, how do you do this? So, you choose $\{A_n\}$ a sequence of pairwise disjoint sets coming from the σ -field \mathcal{F} . Then you look at this summation that you want to check countable additivity, you look at the summation.

So, on your left hand side, I am just adding up the individual sizes according to the new measure. So, put in the definition. So, this 1 by measure of the whole set according to the μ measure comes out. So, that is a common factor comes out of the summation. And you just have to add up the individual sizes according to the μ measure.

But since μ itself is a measure, this will have countable additivity and that allows us to write this ratio that on the numerator I have μ of the whole union, countable union of A_n 's. So, this is simply using the countable additivity of the measure μ . But then I am dividing the whole quantity by the measure of the whole set, and then immediately, I can write it according to the definition in terms of the ν measure.

So, according to ν , this is simply the ν of the complete union or the countable union A_n . So, therefore, we have proved that ν is countably additive non-negative set function, and as per the definition, it becomes a measure. But then, as per our proposition, we wanted to claim that this actually is a probability measure. So, how do you show this?

You look at the size of the whole set according to the ν measure. So, what is this? As per the definition, this is on the numerator, this you have to simply put that set itself, so that is $\mu(\Omega)$, but then you are dividing by $\mu(\Omega)$ as per definition, and hence this ratio is exactly 1. So, therefore, this becomes a probability measure and completes the proof.

So, all we are saying, so let us go back, as soon as you have some examples of finite measures, divide the actual measure by this constant. And then immediately, what you get is this ratio, which falls between 0s and 1s, this automatically becomes a probability measure. So, therefore, as long as you have examples of finite measures, you immediately get examples of probability measures. So, that is the first result that that allows you to construct more examples.

(Refer Slide Time: 14:34)

Example (Uniform measure on a finite set)

Consider a finite set $\Omega = \{x_1, x_2, \dots, x_n\}$ with the σ -field 2^Ω and counting measure μ . We have, $\mu(\Omega) = n < \infty$ and hence μ is a finite measure. By

Proposition (4), $\nu(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{\#A}{n}$ $\forall A \in 2^\Omega$

hence ν is a finite measure. By

Proposition (4), $\nu(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{\#A}{n}$ $\forall A \in 2^\Omega$.

is a Probability measure on $(\Omega, 2^\Omega)$. Here,

ν distributes the whole mass uniformly

with $\nu(\{x_i\}) = \frac{1}{n}$, $i=1, 2, \dots, n$.

And we will immediately apply this to look at some interesting examples of probability measures. So, the example is uniform measure on our finite set. What do we do? We have to first of all, remember, we have to first of all work with finite measures. So, take our finite set Ω , call the points as x_1, x_2, \dots, x_n and look at the power sets, power set σ -field on top of it and look at the counting measure on this.

So, remember, this gives you an example of a finite measure, because the whole set has finitely many elements. So, therefore, the $\mu(\Omega)$ here is exactly equal to n , the number of elements, and

that will be finite. So, this is a finite measure. So, for this finite measure, you can now apply this previous proposition and look at this ν measure as defined there. So, what is this? This simply says look at $\nu(A)$ for any arbitrary set which is now any arbitrary subset of Ω because you are looking at the power set. So, what do you do? You look at $\mu(A)$ divided by the $\mu(\Omega)$, so the size of Ω you divided it by A .

But since we are looking at the finite set Ω , all its subsets are also finite, meaning all subsets are finite number of elements. Therefore, as per the definition of accounting measure, what you get in the numerator is simply number of elements in the set A divided by the total number of elements in the set Ω . So, you are just dividing that number and dividing it by n . And as per the proposition, this ν measure must be a probability measure. So, let us explicitly compute what happens.

So, here if you look at more explicitly on this non-empty set Ω , you are looking at all possible subsets, in particular, you can also look at singletons. So, if x_1, x_2, \dots, x_n where the individual points that you have identified, then look at the singleton sets $\{x_i\}$. So, this set is a subset of Ω , so we can look at the size of that according to the ν measure. And what is the number of elements in this set? It is exactly 1. So, therefore, according to the formula, which we have just computed, this ratio turns out to be $\frac{1}{n}$.

So, for each singleton set, you get the value $\frac{1}{n}$. And now, this gives you the uniform measure because it distributes the whole mass, the total mass is 1, it distributes the total mass 1 into n equal chunks, which is $\frac{1}{n}$. So, you distribute according to the number of elements. So, if number of elements is n , you are just getting $\frac{1}{n}$ weights for each of these points.

So, again, if you are having a doubleton set, meaning a set with two elements, you will get the size as $\frac{2}{n}$. So, that is what the meaning of the uniform measure. So, this is an explicit example of a probability measure which you have constructed out of a finite measure. Here it was the, here the finite measure was the counting measure.

(Refer Slide Time: 17:46)

Exercise ②: we may also look at the previous example in another way. Here,

$$\nu = \sum_{i=1}^n \frac{1}{n} \delta_{x_i}$$

in the sense that

$$\nu(A) = \sum_{i=1}^n \frac{1}{n} \delta_{x_i}(A), \forall A \in 2^{\Omega}.$$

check this equality.

But now, this is an interesting way to look at this previous example. So, we say that we can look at or we can write the measure ν according to a formal summation like this. What is this? I am saying look at this $\sum_{i=1}^n \frac{1}{n} \delta_{x_i}$, Dirac supported at x_i . So, we have looked at the Dirac measures and

I am saying that I am going to look at this summation, this formal summation. Now, you are going to ask what is this formal summation?

So, this means that for all arbitrary sets A from your σ -field, which is the power set here, we are going to have that the $\nu(A)$ is simply this summation. What is this? This is $\sum_{i=1}^n \frac{1}{n} \delta_{x_i}(A)$. So, we are looking at the size of the set arbitrary set A according to the individual Dirac masses and we are just adding them up with this scaling factor $\frac{1}{n}$.

So, now, this is exactly the same measure that was obtained earlier. So, that is another way to write down the same equality. This is simply counting whether the point x_i is in the set A or not, but then if you check it for all possible i , meaning if you check for all possible elements, you are just counting up the number of 1s there, Dirac simply associates the value 1 if the point is there in your set A or not.

So, therefore, if you are adding up this summation from $i = 1$ to $i = n$, you are just checking whether the points x_1, x_2, \dots, x_n is in your set or not. And the number of 1s is exactly the number of elements in the set, and therefore, you exactly get back the ratio, number of elements in the set A divided by n . So, that is the motivation for looking at this kind of a summation and formally we are going to write it as a ν as a linear combination like this with the weightage $\frac{1}{n}$ to each of the Dirac masses.

So, this is a interesting way to represent that ν measure. But then this observation has been put in the form of an exercise, please try to write it down, please check that this equality holds as discussed now.

(Refer Slide Time: 20:16)

Motivated by Exercise ②, we now look at the next result.

Proposition ⑤: Let μ_1 and μ_2 be two probability measures on (Ω, \mathcal{F}) . Fix $p \in [0, 1]$. Then the non-negative set function ν on \mathcal{F} defined by $\nu(A) := p\mu_1(A) + (1-p)\mu_2(A)$, $A \in \mathcal{F}$ is also a probability measure.

Proposition 5: Let μ_1 and μ_2 be two probability measures on (Ω, \mathcal{F}) . Fix $p \in [0, 1]$. Then the non-negative set function ν on \mathcal{F} defined by $\nu(A) := p\mu_1(A) + (1-p)\mu_2(A)$, $A \in \mathcal{F}$ is also a probability measure.

Proof: To check countable additivity, let

But motivated by this example, by this exercise, we are now going to look at some kind of combinations of such Dirac masses or more general probability measures. What do we do? We look at two probability measures call them μ_1 and μ_2 both of them are defined on the same measurable space. Now, fix a constant p between 0 and 1. So, it could be 0, it could also be 1 or it could be a proper fraction between 0 and 1.

Then we are now going to look at a convex combination like this. So, for any arbitrary set A , we are going to look at $\nu(A)$ defined as the convex combination of $\mu_1(A)$ and $\mu_2(A)$, but we are going to associate the weights p and $(1 - p)$. So, this is a convex combination of weights.

So, we are choosing this weight p , we are fixing this parameter p and looking at this convex combination p and $1 - p$. So, with that weights, we are going to look at the convex combination of the size of A according to μ_1 and the size of A according to μ_2 . Look at this convex combination. Now, that I am saying, we will give you a probability measure. So, let us quickly check this out.

So, again, by definition, $\nu(A)$ is non-negative. Why, because p and $1 - p$ both are non-negative, μ_1 and μ_2 are probability measures. So, in particular, $\mu_1(A)$ and $\mu_2(A)$ both are non-negative. So, therefore, this combination that you are looking at is non-negative. So, ν is a non-negative set function. So, that is not a problem. But let us try to check the countable additivity.

(Refer Slide Time: 22:01)

Proof: To check countable additivity, let $\{A_n\}_n$ be a sequence of pairwise disjoint sets in \mathcal{F} . Then

$$\sum_{n=1}^{\infty} \nu(A_n) = p \sum_{n=1}^{\infty} \mu_1(A_n) + (1-p) \sum_{n=1}^{\infty} \mu_2(A_n)$$

$$\begin{aligned} \sum_{n=1}^{\infty} \nu(A_n) &= p \sum_{n=1}^{\infty} \mu_1(A_n) + (1-p) \sum_{n=1}^{\infty} \mu_2(A_n) \\ &= p \mu_1\left(\bigcup_{n=1}^{\infty} A_n\right) + (1-p) \mu_2\left(\bigcup_{n=1}^{\infty} A_n\right) \\ &= \nu\left(\bigcup_{n=1}^{\infty} A_n\right). \end{aligned}$$

Hence, ν is a measure. Moreover,

$$\nu(\Omega) = p \mu_1(\Omega) + (1-p) \mu_2(\Omega) = p + (1-p) = 1.$$

Hence, ν is a measure. Moreover,

$$\nu(\Omega) = p \mu_1(\Omega) + (1-p) \mu_2(\Omega) = p + (1-p) = 1.$$

This completes the proof.

Note ①: To denote such linear combinations

we shall write $\nu = p \mu_1 + (1-p) \mu_2$.

Exercise ③: let $\mu_1, \mu_2, \dots, \mu_n$ be probability

measures on (Ω, \mathcal{F}) choose scalars $\alpha_1, \dots, \alpha_n$

So, how do you check this? Again, start with a sequence of pairwise disjoint sets in your σ -field. Then let us start with that right hand side of the countable additivity property, add up the individual sizes. So, for the individual sizes, put in the definition. So, this is simply will give you these two individual series, one for the μ_1 and another for μ_2 with the individual weights coming from the convex linear combination p and $1 - p$. So, just put in the definition of the $\nu(A_n)$ and you will immediately get this series on your right hand side.

But then you observe that μ_1 and μ_2 are genuine probability measures, in particular, they are countably additive. And therefore, you end up having that this is a convex linear combination of the measures of A , the union of A_n s according to μ_1 and μ_2 . So, here you are simply using the countable additivity of μ_1 and μ_2 . Therefore, all you get is the convex linear combination. And as per the definition, this is nothing but the ν of the countable union of A_n 's. So, again, you have managed to prove that ν is countably additive. And hence this becomes a measure.

But then, our original claim was that this measure is in fact a probability measure. So, how do you check this? All you have to check is the measure of the whole set. So, we are going to look at the size of the whole set Ω . So, $\nu(\Omega)$, what is this? Put in the definition. So, this is nothing but the convex linear combination, size of the set Ω according to μ_1 and μ_2 , but with these weights p and $1 - p$. So, that is the convex linear combination. $\mu_2(\Omega)$ is also 1, because this is a property

measure. And hence you will end up with this formula, this is simply $p + 1 - p$ which is nothing but 1.

So, therefore, with this convex combination that you have looked at, this gives you total measure 1 for the set function ν . So, for a convex combination of measures, you end up having a probability measure. So, let us go back. So, we are saying that take two probability measures μ_1 and μ_2 and look at these convex combinations of individual sizes. And therefore, we are ending up with a probability measure.

So, again, as mentioned a few minutes back, we are now going to remove this set A from this equality as defined here. And we are going to look at this formal summation as some kind of functions, functional equation. So, therefore, we are going to write this as a convex combinations of p and $1 - p$. So, those are the weights that appear in the convex combination and it will be a convex combinations of the measures μ_1 and μ_2 with the weights p and $1 - p$. So, that is the motivation that we have looked at.

Now, this was for convex combinations of two things, two measures μ_1 and μ_2 . Now, we can do the extension of this result and this is stated in exercise 3.

(Refer Slide Time: 25:05)

we shall write $\nu = p\mu_1 + (1-p)\mu_2$.

Exercise ③: let $\mu_1, \mu_2, \dots, \mu_n$ be probability measures on (Ω, \mathcal{F}) . Choose scalars $\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$ with $\sum_{i=1}^n \alpha_i = 1$. Then show that the set function ν on \mathcal{F} defined by $\nu = \sum_{i=1}^n \alpha_i \mu_i$ is also a probability measure.

Note ⑩: We can connect Exercises ② & ③

So, take n measures, n probability measures μ_1 up to μ_n , all of these are defined on the same probability space (Ω, \mathcal{F}) . Now, you are going to choose convex combinations. So, you have to choose scalars α_1, α_2 up to α_n non-negative with the sum equal to 1. So, this will be the weights for the convex combinations. Then what you are going to look at is this ν measure given by these convex combinations α is, μ is.

Now, you should be clear what this equality means. It simply means that if you want to compute ν of some arbitrary set A , all you have to do is to look at μ_i is of A , the size of the set A according to μ_i , first compute that then multiply by the weights α_i and add them up. So, that will be the convex combinations of the sizes of the sets A in terms of the individual measures μ_1 up to μ_n .

So, then as per the exercise here, it will also be a probability measure. So, this proof goes through exactly the same arguments as considered for the convex combinations of two probability measures. So, the same argument goes through and you will get this. And this was in fact, the special case that was mentioned earlier in exercise 3 as the motivation. So, what we said was that the uniform measure on a finite set came out to be these convex combinations of the Dirac masses. So, that is exactly what was our motivation for looking at such combinations, such convex combinations.

(Refer Slide Time: 26:46)

1=1

Note ⑩: We can connect Exercises ② & ③
 by taking $\mu_i = \delta_{x_i}$, $i = 1, 2, \dots, n$
 and $\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{1}{n}$.

Exercise ④: Extend Exercise ③ to the case
 involving a sequence $\{\mu_n\}_n$ of probability
 measures and non-negative scalars
 $\{\alpha_n\}$ with $\sum_{n=1}^{\infty} \alpha_n = 1$. To the set

But with that in hand, we can now identify that thing. So, we can now connect these exercises 2 and 3 which was our motivation that you just say that μ is where this Dirac of x_i 's and α_1 up to α_n is $\frac{1}{n}$. So, that is the convex combination that is given earlier.

(Refer Slide Time: 27:08)

and $\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{1}{n}$.

Exercise ④: Extend Exercise ③ to the case involving a sequence $\{\mu_n\}_n$ of probability measures and non-negative scalars $\{\alpha_n\}_n$ with $\sum_{n=1}^{\infty} \alpha_n = 1$. Is the set

function $\nu := \sum_{n=1}^{\infty} \alpha_n \mu_n$ a probability measure

involving a sequence $\{\mu_n\}_n$ of probability measures and non-negative scalars $\{\alpha_n\}_n$ with $\sum_{n=1}^{\infty} \alpha_n = 1$. Is the set

function $\nu := \sum_{n=1}^{\infty} \alpha_n \mu_n$ a probability measure

Further examples of probability measures

As an application of the above

Now, we can extend this exercise 3, which was for these finite combinations, two infinite combinations involving series. What you have to do? You now look at a sequence μ_n of probability measures and a corresponding sequence of weights. So, this should be non-negative and they should add up to 1. So, you are now looking at series. Now, the question is, is the set function ν given by this countable summation $\sum_{n=1}^{\infty} \alpha_n \mu_n$ a probability measure and the answer will be, yes. Please try to work this out. Again, the meaning is same.

If you take arbitrary set A , then all you have to look at is first μ_n s of A , the weights of the set A according to μ_n and then multiply by α_n and add them up. So, that will give you a probability measure, please check this. Now, these exercises allow us to construct more examples.

(Refer Slide Time: 28:04)

(i) Applying Proposition ② above, $p\delta_1 + (1-p)\delta_0$ will be an example of a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

(ii) Applying Exercise ③ above, for any positive integer n and $p \in [0, 1]$ $\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \delta_k$ will be a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

any positive integer n and $p \in [0, 1]$ $\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \delta_k$ will be a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

(iii) Fix $\lambda > 0$. Applying Exercise ④ above, $\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$ will be a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

So, how, we are now going to look at convex combinations of probability measures. Remember, the explicit example of property measures that we have are the Dirac measures. Now, on the real line, consider Dirac measures δ_1 and δ_0 , meaning the Dirac measures supported at 1 and Dirac measures supported at 0.

If you choose an appropriate p between 0 and 1 and look at this convex linear combination, as per the proposition above, you can immediately get this probability measure. So, let us go back. So, this is a proposition 5, the correction will be made. So, this will be provision 5, sorry. So, the proposition 5 will imply that probability of $p\delta_1 + (1 - p)\delta_0$ that will be a probability measure.

Now, you apply exercise 3, which was for combinations of more than two measures. So, again, look at Dirac measures, but support it at integers from 0 to n . And look at these combinations given from the binomial coefficients. So, choose some $p \in [0, 1]$, fix a positive integer n , look at these coefficients, these coefficients as before, these are familiar coefficients, they add up to 1 as given by the binomial theorem. So, these coefficients add up to 1. And this will give you a convex linear combination for you of the Dirac measures. And as per the exercise earlier, this will be the probability measure.

So, let us go back. So, this is the exercise that I am referring to that you heard you are looking at convex linear combinations of n measures, n probability measures. But then we also looked at the series version of this, where you are looking at a sequence of probability measures. And as I already mentioned, that exercise 4, I left it as a question, but I mentioned that this answer to that easiest. You can get this as a probability measure. So, then you look at these Dirac measures supported at the points k , at the integers k from 0 to infinity. So, look at all possible non-negative integers and look at these weights. So, which is this, what is this?

So, this is $e^{-\lambda} \left(\frac{\lambda^k}{k!}\right)$. So, these weights if you just look at them all alone adding them up you get 1.

So, this also denotes some convex linear combinations of these Dirac measures. So, this will again be a probability measure. Now, let us just go back and look at all these examples. So, these examples that we have so far said from these exercises are following certain parameters.

Here it was n and p and here there is a parameter λ which is taken to be positive then what will happen is that, we are going to connect these explicit examples of property measures with some known random variables known discrete random variables. This we are going to see them later on when we start discussions about random variables, but these examples are coming simply from looking at the convex linear combinations of the Dirac masses.

(Refer Slide Time: 31:15)

Exercise 5: let μ_1, μ_2, \dots be measures on a measurable space (Ω, \mathcal{F}) . Then

(i) The set function $\mu_1 + \mu_2: \mathcal{F} \rightarrow [0, \infty]$ defined by $(\mu_1 + \mu_2)(A) := \mu_1(A) + \mu_2(A), \forall A \in \mathcal{F}$ is a measure.

(ii) The set function $\sum_{n=1}^{\infty} \mu_n: \mathcal{F} \rightarrow [0, \infty]$

(ii) The set function $\sum_{n=1}^{\infty} \mu_n: \mathcal{F} \rightarrow [0, \infty]$ defined by $\left(\sum_{n=1}^{\infty} \mu_n\right)(A) := \sum_{n=1}^{\infty} \mu_n(A), \forall A \in \mathcal{F}$ is a measure.

(iii) Fix $\alpha > 0$ and $i \in \{1, 2, 3, \dots\}$. The set function $\alpha \mu_i: \mathcal{F} \rightarrow [0, \infty]$ defined by $(\alpha \mu_i)(A) := \alpha \mu_i(A), \forall A \in \mathcal{F}$ is a measure.

Now, some final exercises to finish off the lecture. So, earlier we had looked at convex linear combinations of property measures and that gave us examples of probability measures once more. But then we had also looked at another proposition which said that if you start off with a finite measure and scale it accordingly divide it by the whole mass, then you get, also get a proper dimension.

But then this is about dividing by the whole mass which is in some sense multiplying by a factor, but this motivates us to look at this exercise. And this you can do for any measure. You do not have to restrict your attention to probability measures, but you can do it for any measures. So,

this is just a repeat of the earlier exercises that was done for the probability measures. So, what is this? So, if you have μ_1, μ_2 and so, be measures on a fixed measurable space (Ω, \mathcal{F}) , then you have this set of exercises.

First look at two of them μ_1 and μ_2 , then I say that I can add them up and get a measure. How? So, $\mu_1 + \mu_2$ that is defined as follows; so for any arbitrary set A what you do, you add up the individual sizes according to μ_1 and according to μ_2 , add up the sizes and I am saying that you will get a non-negative set function which is also countably additive. So, all you have to follow are the earlier proofs or earlier methodologies that you have done for the exercises and you will immediately get the countable additivity.

By definition, this first example is of course, non negative, because μ_1 and μ_2 both are known to be measures. So, they will take non-negative values. So, that is your first part. In the second part, I am saying instead of adding have two have them, look at the series. So, now, what will happen is that this series will give you a way of defining the series or summation of measures. What is this? So, the countable sum of the measures is defined as follows. So, for any set A , you look at this left hand side and this is defined as adding up the individual sizes of the same set according to different measures μ_n . So, that gives you the second example or second part of the exercise.

And the final exercise is that you can now scale it by appropriate scalars α . So, choose any arbitrary α and choose this indices i from 1, 2, 3 and so on, then if you multiply any one of the measures that was considered earlier $\mu_1, \mu_2, \dots, \mu_n$, any measure that you have considered on here and multiplied by α , you get a measure. So, how it is defined?

So, α times μ_i so that associates the weight or the size to arbitrary set A according to this formula that you first look at $\mu_i(A)$, meaning you look at the size of the set A according to μ_i the measure given and multiplied by α , but then this will be a non-negative set function because α is non-negative and you can immediately prove the countable additivity exactly as discussed earlier. But then, if you put them together, you are getting a non-negative set function which is countably additive and this becomes a measure.

So, now, these three part exercises what does it tell you? It says that you can multiply by positive scalars α and look at additions. So, therefore, you can look at, in general, linear combinations of measures. So, that is the takeaway from this exercise that if you have a some finite number of measures or a countable number of measures then you can look at linear combinations or the corresponding countable sum of the measures, multiply by the appropriate positive scalars, you will still end up with a measure.

So, this is a very interesting construction that will also allow you to construct measures in general from known examples, $\mu_1, \mu_2, \dots, \mu_n$ and so on. So, that is the takeaway from this exercise. We are going to stop here and we are going to continue the discussion in the next lecture.