Measure Theoretic probability 1

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

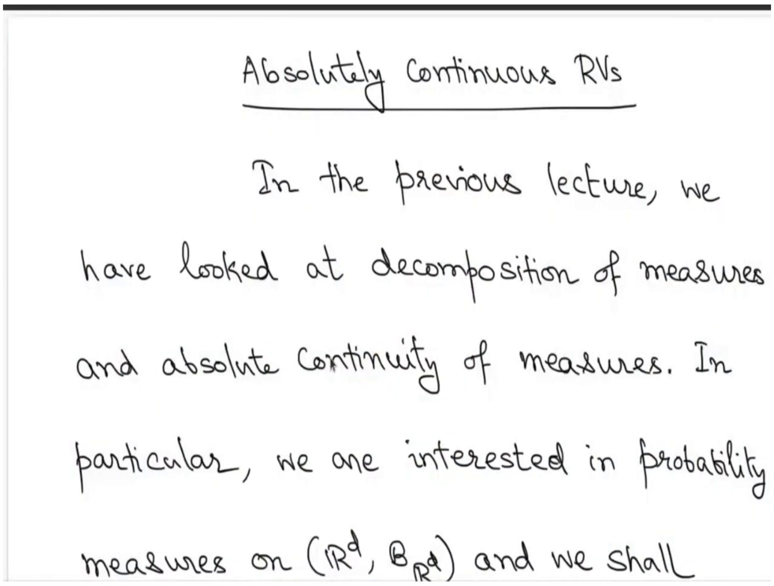
Lecture 37

Absolutely Continous RVs

Welcome to this lecture, before proceeding forward let us quickly recall what we have done in the previous lecture. In the previous lecture we had discussed certain structural properties of measures; in particular we have talked about the Lebesgue decomposition theorem and the Radon-Nikodyne theorem. These gave us some interesting structural properties involving mutual singularity of measures and absolute continuity of measures.

As we shall see using such structural properties we can finally talk about many different types of random variables and corresponding laws in particular we shall talk about absolutely continuous random variables. We are consolidating what we have learned in this course and using all those results putting together we are going to get some very interesting results. Let us move forward and discuss the slides of this lecture.

(Refer Slide Time: 01:25)



Lebesgue measure $\lambda^{(d)}$. In this lecture, we start by considering this problem in dimension one.

Let μ be a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and let F_μ denote the corresponding distribution function. The discussion

Let μ be a probability measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and let F_μ denote the corresponding distribution function. The discussion
below may be restated in terms of an RV with law μ (See the week 4 notes on constructing RVs with a

$(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and let F_μ denote the corresponding distribution function. The discussion below may be restated in terms of an RV with law μ (See the week 4 notes on constructing RVs with a
specified law).

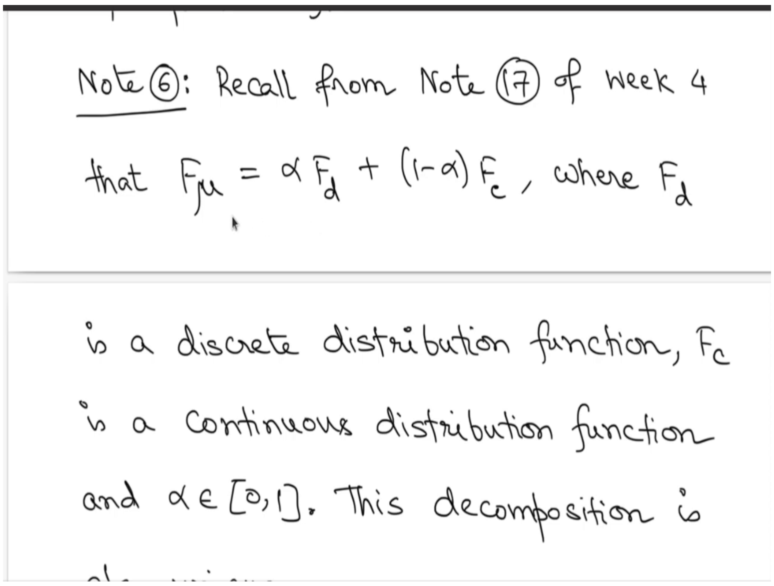
<u>Note 6</u>: Recall from Note 17 of week 4

In the previous lecture we have looked at this decomposition of measures and in particular we have also talked about this absolute continuity. Here we are interested in probability measures on such Euclidean spaces with the Borel σ-field and what we are going to do is to compare these probability measures against the well-known measure, which is the Lebesgue measure λ^d.

In this lecture what do you do we start by considering this specific problem in dimension 1. Towards the end we will talk about the situation in higher dimensions, so for the purpose of the discussion we are going to fix this notation that μ will be a probability measure on r and f_μ will denote the corresponding distribution function. We may also consider the corresponding random variable with this law μ.

Remember in week 4 we had discussed about constructions of random variables or random vectors with a specified law. You construct a random variable and then you can talk about the corresponding distribution function. For the purpose of this lecture, we are mostly going to concentrate on the probability measure, but you can restrict all of this in terms of a random variable.

(Refer Slide Time: 02:56)



that $F_\mu = \alpha F_d + (1-\alpha) F_c$, where F_d

is a discrete distribution function, F_c is a continuous distribution function and $\alpha \in [0,1]$. This decomposition is also unique.

Note (7): Continue with the notations of Note (6). By the correspondence between distribution functions on $\mathbb{R}$ and probability measures on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ (See Note (10) of week 5), we have probability measures μ_d and μ_c corresponding to F_d and F_c respectively. But

Note (10) of week 5), we have probability measures μ_d and μ_c corresponding to F_d and F_c respectively. But

$$\begin{aligned}\mu((-\infty, x]) &= F_\mu(x) \\ &= \alpha F_d(x) + (1-\alpha) F_c(x) \\ &= \alpha \mu_d((-\infty, x]) + (1-\alpha) \mu_c((-\infty, x])\end{aligned}$$

Let us move forward, so recall from note 17 of week 4 that we had discussed this decomposition of a distribution function into a discrete distribution function and a continuous distribution function. Here this is a convex linear combination given by some scalar α between 0 and 1. We have also remarked that this decomposition is unique.

With this recall continuing with these notations of note 6, we can now appeal to the correspondence between distribution functions on the real line and probability measures on this measurable space. This was discussed earlier in week five see note 10 of week five. What do we get is that we can get probability measure μ_d and μ_c which correspond to the distribution functions f_d and f_c.

So, f subscript d is the discrete one, F_c is the continuous one. Correspondingly you get these probability measures μ_d and μ_c, but what is happening here? Let us try to identify the distribution functions.

(Refer Slide Time: 04:17)

to F_d and F_c respectively. But

$$\mu((-\infty, x]) = F_\mu(x)$$
$$= \alpha F_d(x) + (1-\alpha) F_c(x)$$
$$= \alpha \mu_d((-\infty, x]) + (1-\alpha) \mu_c((-\infty, x])$$

for all $x \in \mathbb{R}$. Since, $\alpha \mu_d + (1-\alpha)\mu_c$ is

also a probability measure by the

$$= \alpha \mu_d((-\infty,x]) + (1-\alpha)\mu_c((-\infty,x])$$

for all $x \in \mathbb{R}$. Since, $\alpha\mu_d + (1-\alpha)\mu_c$ is also a probability measure, by the uniqueness of the distribution functions (see Exercise 1 of Week 5), we have

$$\mu(A) = \alpha\mu_d(A) + (1-\alpha)\mu_c(A), \forall A \in \mathcal{B}_{\mathbb{R}}.$$

Now, we started off with this probability measure μ if you lo at the size of the set $(-\infty, x]$ for any x in the real line you get the value of the distribution function at the point x but the identification which we just recalled is exactly this, this convex linear combination in terms of a discrete distribution function and a continuous distribution function.

But then we just said that we will consider the corresponding probability measures and those are denoted by μ_d and μ_c. So, all you have to do is to look at μ_d the size that it associates to this set $(-\infty, x]$ that will give you this distribution function and if you look at μ_c and consider this set $(-\infty, x]$ this size will give you the continuous distribution function f_c.

So, that is as per the structural properties that we have discussed, but now observe that μ_d and μ_c both are probability measures and you are considering a convex linear combination of probability measures. Therefore, you get this probability measure which is this convex linear combination of probability measures.

But you have just identified that this convex linear combination of probability measures has this interesting property that the corresponding distribution function is exactly f_μ. Again, let me repeat $\alpha\mu_d + (1-\alpha)\mu_c$ this is a probability measure with the distribution function f_μ that is as per these equalities that we have written down.

Therefore, by the uniqueness of the distribution functions that was left as an exercise earlier in week five we have that these measures are actually the same, therefore, $\mu = \alpha\mu_d + (1-\alpha)\mu_c$. That means that for any Borel set you have this equality.

This is something very simple that we have just restated and appealed to the correspondences between distribution functions and probability measures. Let us continue with these notations now. We have so far got these notations probability measure μ corresponding distribution function is f_{μ} then f_{μ} got splitted into discrete part and continuous part and correspondingly we got μ_d and μ_c.

(Refer Slide Time: 07:12)

$$= \alpha \mu_d((-\infty, x]) + (1-\alpha)\mu_c((-\infty, x])$$

for all $x \in \mathbb{R}$. Since, $\alpha\mu_d + (1-\alpha)\mu_c$ is also a probability measure, by the uniqueness of the distribution functions (see Exercise 1 of Week 5), we have

$$\mu(A) = \alpha\mu_d(A) + (1-\alpha)\mu_c(A), \forall A \in \mathcal{B}_{\mathbb{R}}.$$

Note 8: Continue with the notations of Note 7. The probability measure μ_d corresponds to the discrete distribution function F_d, and hence μ_d is the law of a discrete RV. Here $\mu_d \perp \lambda$, by Exercise 1 (ii).

Exercise 1 (ii).

<u>Note 9</u>: Continue with the notations of Note 7. We shall now focus on μ_c and the corresponding continuous distribution function F_c. This is in continuation of the discussion in Note 18 of Week 5. By the Lebesgue

continuation of the discussion in Note 18 of Week 5. By the Lebesgue

decomposition Theorem (Theorem 1), we have $\mu_c = \tilde{\mu}_c^1 + \tilde{\mu}_c^2$ for some measures $\tilde{\mu}_c^1$ and $\tilde{\mu}_c^2$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that $\tilde{\mu}_c^1 \perp \lambda$ and $\tilde{\mu}_c^2 \ll \lambda$. Since

Now, this probability measure μ_d that corresponds to this discrete distribution function must be discrete, meaning that it specifies its total mass only on a finite set or a counter with infinite set. Therefore, μ_d is the law of a discrete random variable, so it corresponds to a discrete distribution function and hence this is the law of a discrete random variable and as per our discussion in the previous lecture this must be mutually singular with respect to the Lebesgue measure.

This was discussed earlier in the previous lecture. Therefore, in this decomposition of μ into $\alpha\mu_d + (1 - \alpha)\mu_c$, you have identified that the first part is mutually singular with respect to the Lebesgue measure, but we are now going to focus on the remaining part this μ_c part.

Then sometimes we will refer to the corresponding distribution function which is this f_c, this is a continuous distribution function, sometimes we will focus on μ_c. We will use information from both sides. Now we would like to understand the structure of μ_c or corresponding f_c.

This is the continuation of the discussion that was left out earlier in note 18 of week 5 so there we had discussed about discrete random variables, discrete distribution functions and so on and we had mentioned that we will come back to other types of distribution functions in week 8, so this is where we are starting off. Now our focus is purely on that continuous part.

By the Lebesgue decomposition theorem that was discussed in the previous lecture we have that μ_c that can be split as the addition of two measures let us write them as $\tilde{\mu}_c^1$ and $\tilde{\mu}_c^2$. These are two measures on the real line together with the Borel σ- field.

Moreover, as per the Lebesgue decomposition theorem you must have one of them say $\tilde{\mu}_c^1$ is mutually singular with respect to the Lebesgue measure and the other one $\tilde{\mu}_c^2$ is absolutely continuous with respect to the Lebesgue measure. This is the structural properties that is given by the Lebesgue decomposition theorem.

Now, look at the size of the set real line under this measures, so here you look at this sum of these two quantities and by the definition you have that is the size of the real line under the given probability measure μ_c, but that is 1. What do you get?

(Refer Slide Time: 10:39)

measures $\tilde{\mu}_c^1$ and $\tilde{\mu}_c^2$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that $\tilde{\mu}_c^1 \perp \lambda$ and $\tilde{\mu}_c^2 \ll \lambda$. Since

$$\tilde{\mu}_c^1(\mathbb{R}) + \tilde{\mu}_c^2(\mathbb{R}) = \mu_c(\mathbb{R}) = 1,$$

we have $\tilde{\mu}_c^1(\mathbb{R}) \in [0,1]$ and $\tilde{\mu}_c^2(\mathbb{R}) \in [0,1]$.

<u>Exercise 5</u>: Continue with the notations of Note 9. Show that there exist

<u>Exercise 5</u>: Continue with the notations of Note 9. Show that there exist

probability measures μ_c^1 and μ_c^2 and a scalar $\beta \in [0,1]$ such that

(i) $\tilde{\mu}_c^1 = \beta \mu_c^1$, $\tilde{\mu}_c^2 = (1-\beta)\mu_c^2$

of Note 9. Show that there exist

probability measures μ_c^1 and μ_c^2 and a scalar $\beta \in [0,1]$ such that

(i) $\tilde{\mu}_c^1 = \beta \mu_c^1$, $\tilde{\mu}_c^2 = (1-\beta)\mu_c^2$

(ii) $\mu_c^1 \perp \lambda$, $\mu_c^2 \ll \lambda$.

(iii) $\beta = 0 \iff \mu_c^1 \equiv 0$.

These two quantities this addition of these two quantities is 1 but these are some non-negative real numbers and therefore, they take values between 0 and 1. That is the conclusion from this relation above, great! If you continue with these notations, what you can now try to show is that there exists two probability measures which we denote by μ_c^1 and μ_c^2, which are defined on this real line and such that there is a scalar β which takes values between 0 and 1 with these structural properties that the earlier measures, the $\tilde{\mu}$ s can be written as a scalar multiple of these probability measures which are μ_c^1 and μ_c^2.

Earlier these measures $\tilde{\mu}_c^1$ and $\tilde{\mu}_c^2$; these were not necessarily probability measures. There what we observed is that the total mass associated to the real line takes values between 0 and 1 they could be probability measures but if one of them is a probability measure then the other one is identically 0.

But now what we are saying is that you can figure out a scalar β and you can find these two probability measures such that the original measures, $\tilde{\mu}$'s can be represented as a scalar multiple of these probability measures. Of course, here β could be 0 or 1 and in those cases, we can make some comments, but before that since this $\tilde{\mu}$'s, these measures have some relation with respect to the Lebesgue measure you can make the similar comments with respect to these new probability measures that you construct.

Here μ_c^1 must be mutually singular with respect to λ and μ_c^2 must be absolutely continuous with respect to λ. Let us come to these extreme cases when β is 0 or 1. If β is 0 then the first part should contribute 0, so that should be a 0 measure. This is a if and only if condition, and if β is 1 then the second part should not contribute that should be identically 0, again this is an if and only if condition.

So, try to work this out this identification is left as an exercise, but then what we have claimed in this exercise is that you have some identification of these measures $\tilde{\mu}$ s in terms of certain probability measures. So let us put it together with the structure that we already have.

(Refer Slide Time: 13:41)

$\beta = 1 \iff \mu_c^2 \equiv 0.$

Note (10): Continue with the notations of Exercise (5). We have $\mu_c = \beta \mu_c^1 + (1-\beta)\mu_c^2$.

let F_c^1 and F_c^2 denote the distribution functions of μ_c^1 and μ_c^2 respectively.

Then, for all $x \in \mathbb{R}$,

$$F_c(x) = \mu_c((-\infty, x])$$

functions of μ_c^1 and μ_c^2 respectively.

Then, for all $x \in \mathbb{R}$,

$$F_c(x) = \mu_c((-\infty, x])$$

$$= \beta \mu_c^1((-\infty, x]) + (1-\beta)\mu_c^2((-\infty, x])$$

$$= \beta F_c^1(x) + (1-\beta) F_c^2(x).$$

Note (11): Continue with the notations of

Note (10). For any $x \in \mathbb{R}$,

Note (11): Continue with the notations of

Note (10). For any $x \in \mathbb{R}$,

$$F_c^2(x) - F_c^2(x-) = \mu_c^2((-\infty, x]) - \mu_c^2((-\infty, x))$$

$$= \mu_c^2(\{x\})$$

$$= 0$$

If you have these results from exercise 5 then put it together then μ_c that continuous part that we obtained from the given probability measure μ that part can be written as a convex linear combination of two measures, two probability measures. We would like to explore the properties of these two probability measures. Now what do you do?

Directly go to the corresponding distribution function so you know that μ_c^1 and μ_c^2 these are probability measures and you are going to look at the corresponding distribution functions. Again, if you look at any real number x then the distribution function of μ_c that is f_c that is computed as the size of the set $(-\infty, x]$ under the measure μ_c.

But as just obtained, μ_c can be written as this convex linear combination, and that is exactly telling you that the distribution function f_c that was the continuous part of the distribution function f_μ that can be written as a convex linear combination of another two distribution functions. We are going to explore the structural properties of these new distribution functions and the corresponding measures μ_c^1 and μ_c^2.

So, that is what we are going to concentrate on. Here we are focusing specifically on the continuous part of the original distribution function f_μ; we already understand the discrete part. So, continue with these notations but focus on this F_c^2 that corresponds to μ_c^2, remember μ_c^2 as per the construction is absolutely continuous with respect to Lebesgue measure.

So, let us use this information, here what is happening is that you can now try to compute the jumps of these F_c^2 at any point x. So, if you fix a point x look at this jump that is nothing but the difference of the function values of x and the left limit at that function value.

That is the difference and that is giving you the jump size, but let us write them in terms of the corresponding measure which is μ_c^2 and as per the standard computation you have to look at the size of the singleton set $\{x\}$ under this measure μ_c^2 and we claim that this value is 0, Why?

(Refer Slide Time: 16:47)

Since $\mu_c^2 << \lambda$ and $\lambda(\{x\}) = 0$. Hence, F_c^2 is a Continuous distribution function.

Moreover, by Note ⑤, if f is the p.d.f of μ_c^2, then for all $x \in \mathbb{R}$,

$$F_c^2(x) = \int_{(-\infty, x]} f(y)\,\lambda(dy) = \int_{-\infty}^{x} f(y)\,dy.$$

Moreover, by Note ⑤, if f is the p.d.f of μ_c^2, then for all $x \in \mathbb{R}$,

$$F_c^2(x) = \int_{(-\infty, x]} f(y)\,\lambda(dy) = \int_{-\infty}^{x} f(y)\,dy.$$

Then F_c^2 is an absolutely Continuous distribution function as per the following definition.

<u>Definition</u> ⑤ (Absolutely Continuous Distribution functions)

A distribution function $F: \mathbb{R}^d \to [0,1]$ is said to be absolutely Continuous if $\mu << \lambda^{(d)}$, where μ is the probability measure on $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$ Corresponding to the distribution function F.

Because μ_c^2 is absolutely continuous with respect to the Lebesgue measure and this singleton set is an set of measure 0 under the Lebesgue measure, therefore, this must be a set of measure 0 under the μ_c^2 that measure. Therefore, what is the conclusion from here that there is no jump, so therefore the distribution function corresponding to μ_c^2 that has no jumps this is a continuous distribution function.

Moreover, by the discussion in the previous lecture which is in note 5 if f is the PDF of this so μ_c^2 is a probability measure which is absolutely continuous with respect to the Lebesgue measure and in this case, we have discussed about the existence of the probability density function, which is nothing but the Radon-Nikodyne derivative of the probability measure μ_c^2 with respect to the Lebesgue measure, so that is the PDF here.

Take that PDF we call it f then for all x in the real line look at the distribution functions value again what is this, this is nothing but the size of the set $(-\infty, x]$ under the measure μ_c^2, but as per the identification in the Radon-Nikodyne theorem that is this integration of the function over the set $(-\infty, x]$.

This is the identification of the measure μ_c^2 with respect to the Lebesgue measure, so that is the identification and we are using the probability density function in this expression, but then use the standard notation of writing the Lebesgue integrations in terms of this standard integration notation. That is again our standard notation that we are following, this is just for the simplification of the notation. Therefore, what do we get, we get that F_c^2 this is an absolutely continuous distribution function.

So, now we will ask what is an absolutely continuous distribution function, this is what we are going to discuss in the next definition. You already have a structural property of F_c^2 that it is given as an integration of the corresponding probability density function over the range $-\infty$ to x. That is giving back the distribution function value at x.

Let us look at this definition, we state it in higher dimensions in general that you say that this is absolutely continuous if the corresponding probability measure μ is absolutely continuous with respect to the Lebesgue measure in this d dimensional setup. If that happens then you say that the corresponding distribution function is absolutely continuous. Absolute continuity of measures is being transferred to the absolute continuity of the distribution function f. Now in this terminology F_c^2 turns out to be absolutely continuous, great.

(Refer Slide Time: 20:13)

Moreover, by Note ⑤, if f is the p.d.f of μ_c^2, then for all $x \in \mathbb{R}$,

$$F_c^2(x) = \int_{(-\infty, x]} f(y)\,\lambda(dy) = \int_{-\infty}^{x} f(y)\,dy.$$

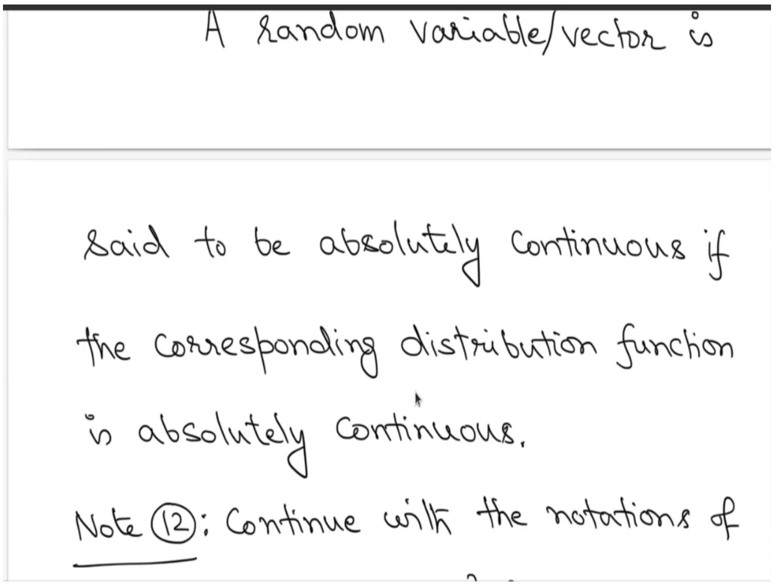
Then F_c^2 is an absolutely continuous distribution function as per the following definition.

the distribution function F.

<u>Definition ⑥</u> (Absolutely Continuous random variables/vector)

A random variable/vector is

said to be absolutely continuous if



Corresponding to these absolutely continuous measures, absolutely continuous distribution functions you can talk about corresponding class of random variables or random vectors. A random variable or a vector is said to be absolutely continuous if the corresponding distribution function is absolutely continuous.

This is again the same type of motivation that we followed in the discrete case, we look at the corresponding distribution function and if the corresponding distribution function is absolutely continuous you call the original random variable or a vector to be absolutely continuous. In the discrete case we looked at the support and then identified that the corresponding distribution function must be discret, but this is just a alternative description in terms of the corresponding distribution functions properties.

Now, we have identified one part of the, continuous part of the original distribution function, so let me repeat f_{μ} is the original distribution function, F_c is the continuous part of that and we have identified F_c^2 as a absolutely continuous distribution function. This is the part that we have identified; let us focus on the other part.

(Refer Slide Time: 21:36)

Then, for all $x \in \mathbb{R}$,

$$
\begin{aligned}
F_c(x) &= \mu_c((-\infty, x]) \\
&= \beta \mu_c^1((-\infty, x]) + (1-\beta)\mu_c^2((-\infty, x]) \\
&= \beta F_c^1(x) + (1-\beta) F_c^2(x).
\end{aligned}
$$

Note (11): Continue with the notations of

Note (10) For any $x \in \mathbb{R}$,

Note (12): Continue with the notations of Note (10). By Note (11), F_c^2 is continuous. Recall from Note (6) that F_c is also continuous. Therefore, F_c^1 is a continuous distribution function. Now, F_c^1 corresponds to the probability

continuous distribution function. Now,

F_c^1 corresponds to the probability

measure μ_c^1 with $\mu_c^1 \perp \lambda$. Then F_c^1 is a singular continuous distribution function in the sense of the following definition.

Definition (7) (Singular Continuous

But then what we have so far understood is that F_c^2 this is absolutely continuous and in particular what we have already justified is that this is a continuous distribution function, but then the decomposition, the original decomposition that we had started off with in that f_c that continuous part was definitely taken to be continuous.

So, now F_c has this decomposition in terms of F_c^1 and F_c^2, so let us go back to the decomposition, here we say that F_c that is continuous and F_c^2 is also continuous and that will immediately tell you that F_c^1 that must be continuous.

There are now two cases that you would have to verify to make this statement if β is 0 so this term will not contribute, you do not have to look at it, if β is 0 you ignore this term and F_c is purely of the second type here absolutely continuous, but if β is non-zero if there is a non-trivial contribution coming up here then F_c^1 must be continuous, so that is the upshot of this discussion.

Let us go back to that, what we have is that F_c^1 that is a continuous distribution function. However, this corresponds to this probability measure which we denoted as μ_c^1 and by the structural property this must be mutually singular with respect to the Lebesgue measure. Then we are going to call F_c^1 to be a singular continuous distribution function. So this notion we define in the next definition.

(Refer Slide Time: 23:33)

Note (12): Continue with the notations of Note (10). By Note (11), F_c^2 is Continuous. Recall from Note (6) that F_c is also Continuous. Therefore, F_c^1 is a Continuous distribution function. Now, F_c^1 Corresponds to the probability

Definition (7) (Singular Continuous Distribution functions)

A continuous distribution function $F: \mathbb{R}^d \to [0,1]$ is said to be singular if $\mu \perp \lambda^{(d)}$, where μ is the probability measure on $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$

function F.

Note (13): We have briefly discussed an example of Singular Continuous distribution function in Note (3). In this course we do not discuss such examples in detail.

What is a singular continuous distribution function this is the continuous distribution function again defined on the d dimensional setup itself and this will be called to be singular if μ the corresponding probability measure is mutually singular with respect to the Lebesgue measure.

So, in this setup what we have observed is that the continuous part of the original distribution function that was F_{μ} the continuous part of that is F_c this F_c has a convex linear combination or a decomposition into one part which is absolutely continuous and another part which is singular continuous.

We have briefly discussed an example of a singular continuous distribution function in note 3 earlier; this was regarding the null sets which are possibly uncountable. In particular we mentioned the example of the cantor set. Here the Lebesgue measure associates 0 mass and you can still consider certain probability measures which live on that set, meaning on the complement of the canter set the probability measure should associate 0 mass only on the cantor set it assigns the full mass.

With such probability measures it is possible to get examples of corresponding distribution functions which are continuous. This is the type of example that we are looking for, but in this course we are not going into the details and we are not going to discuss this much further but it is for our information that such examples are known and we should keep track of them whenever we are doing any kind of a general discussion involving random variables or measures, probability measures or corresponding distribution functions.

(Refer Slide Time: 25:41)

Note (14): Continue with the notations of Notes (10) and (11). Combining the above results, we have the decomposition into a convex linear combination:

$$F_\mu = \alpha F_d + (1-\alpha)\left[\beta F_c^1 + (1-\beta) F_c^2\right]$$

Combination:

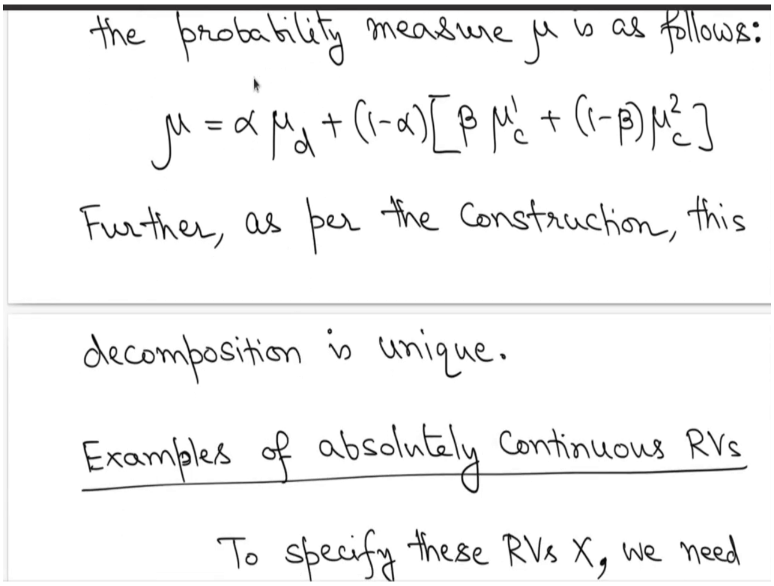
$$F_\mu = \alpha F_d + (1-\alpha)\left[\beta F_c^1 + (1-\beta) F_c^2\right]$$

where (i) F_d is discrete

(ii) F_c^1 is singular continuous

(iii) F_c^2 is absolutely continuous.

A corresponding decomposition for



Continue with these notations of this discussion so far what we have identified is that we have this convex linear decomposition or convex linear combination of the original distribution function f_{μ} that is in terms of these three things. So first we decomposed into F_d and F_c and F_d turned out to be discrete, but then F_c further was decomposed into a convex linear combination of a singular continuous distribution function and an absolutely continuous distribution function.

So, that is the complete decomposition that we have at this moment, but correspondingly we also have this decomposition for the probability measure that we had started off with. Now here the same terms will correspond to the similar terms for the measures. What here you have is that you have this μ_d which is the law of a discrete random variable and then you have these two other measures one which is singular with respect to the Lebesgue measure, but its distribution function is continuous and the second part is absolutely continuous with respect to the Lebesgue measure.

Now, if you go through the construction, you can again try to show that this decomposition is unique. So, we will go forward assuming this factor. We are interested in looking at examples of absolutely continuous random variables. Let us look at them.

(Refer Slide Time: 27:29)

To specify these RVs X, we need to look at examples of probability measures μ_X on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ with $\mu_X \ll \lambda$.

For such measures μ_X, we require p.d.fs f_X, i.e., non-negative Borel measurable functions with

For such measures μ_X, we require p.d.fs f_X, i.e., non-negative Borel measurable functions with

$$\int_{\mathbb{R}} f_X \, d\lambda = \int_{-\infty}^{\infty} f_X(x)\, dx = 1.$$

(i) ($X \sim \text{Uniform}(0,1)$, i.e. Uniform distribution on $(0,1)$)

$\mathbb{R}$ $-\infty$

(i) ($X \sim \text{Uniform}(0,1)$, i.e. Uniform distribution on $(0,1)$)

Take $f = 1_{(0,1)}$.

Here, $\int_{-\infty}^{\infty} f_X(x)\, dx = \lambda((0,1)) = 1.$

(ii) ($X \sim \text{Exp}(\alpha)$, i.e. Exponential distribution with rate $\alpha > 0$)

To specify these random variables what we exactly need are examples of probability measures. Let us call them as μ_x, so to specify these random variables x we want to specify the corresponding law. As soon as you have a law you can construct such random variables. We restrict our attention to probability measures on the real line. Now what do you want in this case you want that these probability measures must be absolutely continuous with respect to the Lebesgue measure.

For such measures as per the Radon-Nikodyne theorem what do you require? You require probability density functions which we write it as f_X, which has this property that these are non negative Borel measurable functions with the total integration over the real line to be one.

Here we are following this standard notation that Lebesgue integration is essentially written as Riemann integration or that same integration notation. This is again for notational convenience. All you have to do is to specify appropriate property density functions and what do you do you look back at standard examples that you may have covered in your basic probability courses.

So, start with this uniform distribution on $(0, 1)$. You say that X follows uniform $(0, 1)$ on the open interval $(0, 1)$ and for that what you need to do is to take the corresponding probability density function to be the indicator of $(0, 1)$, and here all you have to verify that it is a non-negative measurable function. That is already true and you also have to verify that the total integration is 1 and that immediately follows because this is nothing but the length of the interval under the Lebesgue measure and that is equal to 1.

(Refer Slide Time: 29:20)

(ii) ($X \sim \text{Exp}(\alpha)$, i.e. Exponential distribution with rate $\alpha > 0$)

Take $f_X(x) = \alpha e^{-\alpha x} \mathbb{1}_{(0,\infty)}(x), \ \forall x \in \mathbb{R}$.

Here, $\int_{-\infty}^{\infty} f_X(x)\, dx = \int_0^{\infty} \alpha e^{-\alpha x} dx = 1$.

(iii) ($X \sim N(\mu, \sigma^2)$, i.e. Normal distribution with parameters $\mu \in \mathbb{R}$, $\sigma > 0$)

(iii) ($X \sim N(\mu, \sigma^2)$, i.e. Normal distribution with parameters $\mu \in \mathbb{R}$, $\sigma > 0$)

Take $f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \ \forall x \in \mathbb{R}$.

The verification of $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$ is

The verification of $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$ is

usually covered in basic probability courses and we do not go into the details for brevity.

Note (15): Further standard examples of absolutely continuous RVs, as

Look at a slightly more complicated example this is the exponential distribution with rate α, α is taken to be some positive scalar. You say that X follows exponential α distribution if the corresponding probability density function is given as this; this is $\alpha e^{-\alpha x}$ multiplied by this indicator function.

Then again you verify that this is a non-negative measurable function because it is the product of one continuous function against this measurable function which is an indicator function. So, this whole product makes sense and this becomes a measurable function. Moreover, these values that are being specified on the right-hand side these values are non-negative.

Therefore, f_X that you are considering now it is a non-negative measurable function, all you have to now verify is that the total integration is 1 but then that exactly comes down to the integration from 0 to ∞ because of this presence of this indicator function, but then this is a simple example of a Riemann integration and you connect it with the Lebesgue integration and that turns out to be 1. So, this is a very standard verification.

You now look at another standard example which is the normal distribution with parameters μ which is in the real line and σ which is some positive quantity, so please do not get confused with this μ which we have been using so far. This is now a real number. So far we have been using μ for the general probability measure on the real line, but now we are restricting our attention to this specific examples of absolutely continuous random variables.

And here in this example we are taking μ to be a real scalar. Now if you specify the density function in this standard format you again try to verify that this is a non-negative measurable function and the total integration is 1. Again, the total integration is one is usually covered in your basic probability courses and you must have seen this. We are not going into the details for brevity. This is assumed that this is giving you a genuine probability density function.

(Refer Slide Time: 31:39)

Note (15): Further standard examples of absolutely continuous RVs, as seen in basic probability courses, can be discussed by specifying the p.d.f.

Note (16): The decomposition of probability measures on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, as discussed

be discussed by specifying the p.d.f.

Note (16): The decomposition of probability measures on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, as discussed

in this lecture, extends in an analogous manner to higher dimensions.

We can similarly discuss standard

analogous manner to higher dimensions.

We can similarly discuss standard

examples of absolutely continuous random vectors, such as Gaussian random vectors, as discussed in basic probability courses.

But you may have seen other nice examples of absolutely continuous random variables in your basic probability courses and all you have to do is to specify that corresponding probability density function and specify the corresponding law and that will give you examples of the random variables, if you look at with the specified law that is it.

You can specify the PDF that specifies the probability measure that specifies the random variable; you can construct such random variables. Now this decomposition of probability measures as we have discussed in this lecture you can extend this further in an analogous manner to higher dimensions.

We have split a general probability measure into three parts, one part corresponds to discrete random variables, one part corresponds to singular continuous distribution functions and you can also refer to the corresponding random variables as singular continuous random variables and the final part is the absolutely continuous random variables that is the part that we have been concentrating on so far.

These three parts that constitutes a general probability measure, so that one also has this corresponding decomposition in terms of the distribution functions. These results can be extended in higher dimensions again similar discussions will go through and you can also get standard examples of absolutely continuous random vectors. So, this is now in higher dimensions so you will get random vectors and in particular you can talk about these Gaussian random vectors and other things that you have already seen in your basic probability courses.

Again, all you have to do is to specify the appropriate probability density function with respect to the Lebesgue measure on this d-dimensional setup, so that is all you have to do. If you do that you will get all these nice examples, but again for brevity we are not going into the details, but it should be assumed that these are now being covered the structural properties has been discussed you have connected it with your existing knowledge that you have gained from your basic probability theory.

That covers the discussion in this lecture and in the next lecture we are going to finally talk about the expectations of absolutely continuous random variables. That is what we are going to do in the next lecture, so we stop here.