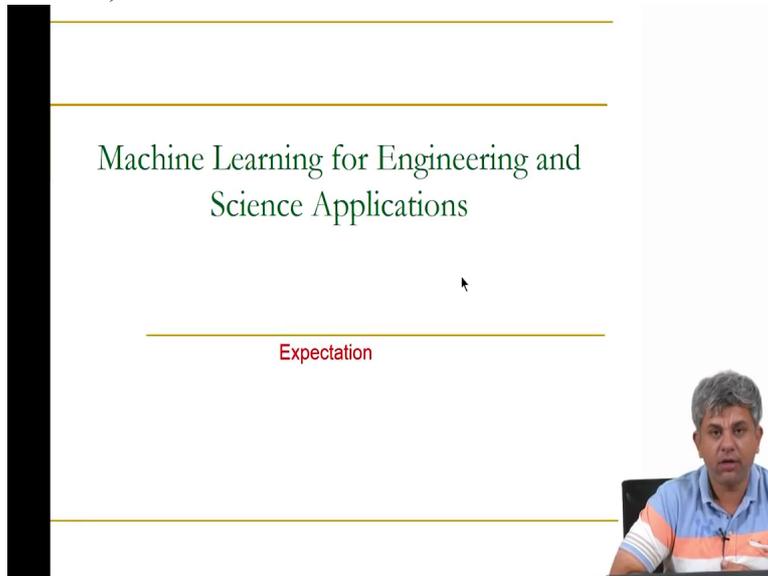


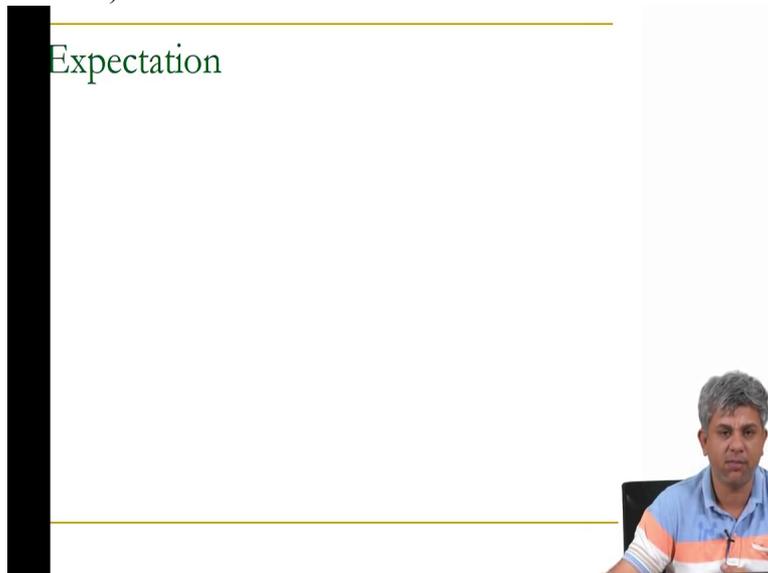
Machine Learning for Engineering and Science Applications
Professor Doctor Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology Madras
Expectation

(Refer Slide Time: 00:13)



In this video we will be looking at a very simple statistical quantity called, or the statistical function called the expectation. You are familiar with expectation as

(Refer Slide Time: 00:23)



the mean or the average, Ok.

(Refer Slide Time: 00:25)

Expectation

Random variables, by definition, result in different outcomes

The slide features a green title 'Expectation' at the top left. Below it, the text 'Random variables, by definition, result in different outcomes' is displayed. A video inset in the bottom right corner shows a man with grey hair, wearing a blue and orange striped polo shirt, sitting at a desk and speaking into a microphone.

So the context is this. If you remember, we are dealing with random variables throughout. Random variables by definition will result in different outcomes.

If I throw a dice right now, sometimes it will give a 1, sometimes it will give a 2, sometimes it will give a 6. So this is obviously why it is called a random variable in the first place,

(Refer Slide Time: 00:43)

Expectation

Random variables, by definition, result in different outcomes

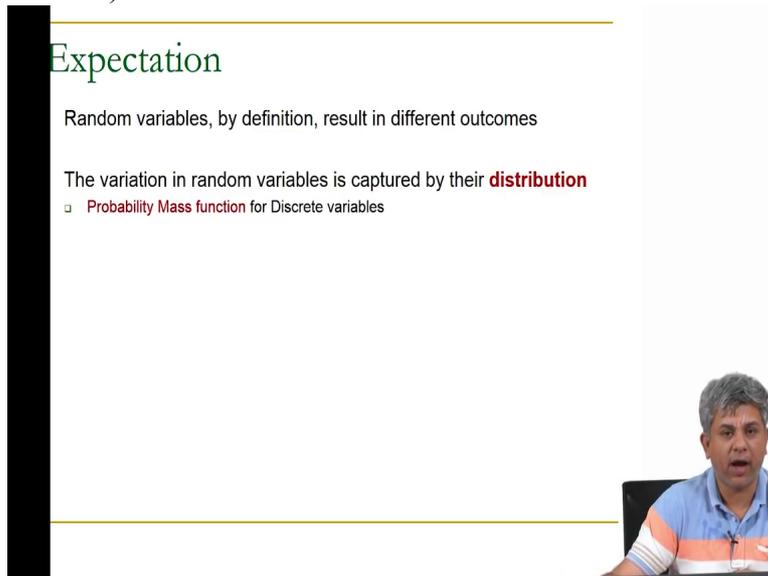
The variation in random variables is captured by their **distribution**

The slide features a green title 'Expectation' at the top left. Below it, the text 'Random variables, by definition, result in different outcomes' is displayed. The second line of text, 'The variation in random variables is captured by their **distribution**', has the word 'distribution' in bold red. A video inset in the bottom right corner shows the same man from the previous slide, speaking into a microphone.

Ok.

So which way the random variable actually varies or how it gives different outcomes is captured by what is known as the distribution if you recall, if you have

(Refer Slide Time: 00:55)



The slide is titled "Expectation" in green. It contains the following text:

Random variables, by definition, result in different outcomes

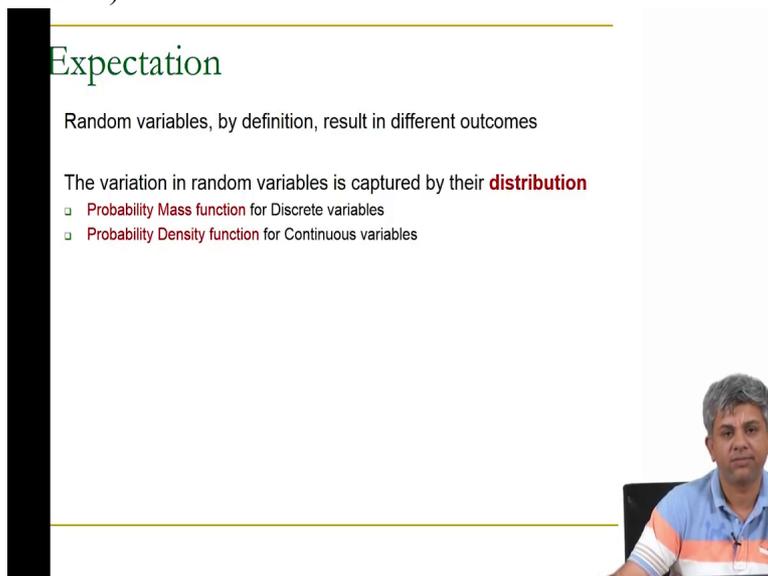
The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables

A small video inset in the bottom right corner shows a man with grey hair wearing a blue and orange striped polo shirt, speaking.

a discrete distribution such as a dice or a coin or a deck of cards, then we have something called the probability mass function that tells you how likely each one of these outcomes is.

(Refer Slide Time: 01:08)



The slide is titled "Expectation" in green. It contains the following text:

Random variables, by definition, result in different outcomes

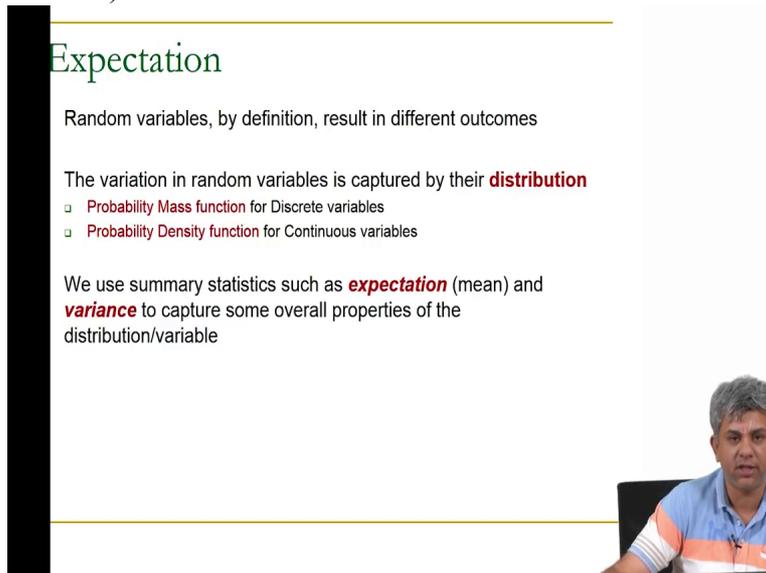
The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables
- **Probability Density function** for Continuous variables

A small video inset in the bottom right corner shows the same man from the previous slide, now silent.

Similarly, for a continuous variable such as height, weight, temperature, pressure, stress, strain etc what you have is the probability density function that tells you how likely a range of probability is, a range of values is. So the probability that height varies between 5 point 6 and 5 point 7 is something, Ok. So that would be a probability

(Refer Slide Time: 01:32)



Expectation

Random variables, by definition, result in different outcomes

The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables
- **Probability Density function** for Continuous variables

We use summary statistics such as **expectation** (mean) and **variance** to capture some overall properties of the distribution/variable

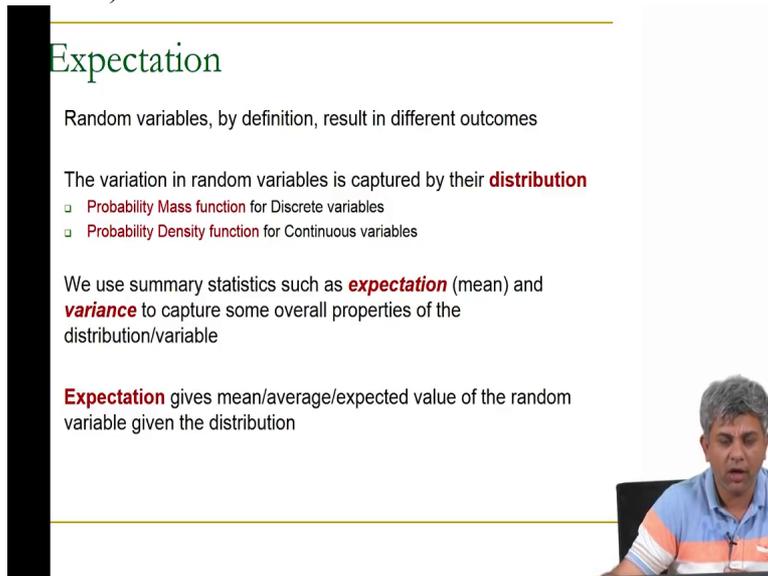


density function.

Now once you are given the distribution we start using some overall ideas, you know. It is a random variable. It has any number of, large number of values but you want to give some summary statistics, Ok. So you want to give some qualitative and quantitative picture of what the random variable is doing.

And the most common ones that we will be using at least as far as this course is concerned are two quantities called the expectation and the variance, Ok. The expectation is something that you are already familiar with. We usually call it the mean or the average. But in the context of random variables typically we tend to call it expectation, Ok.

(Refer Slide Time: 02:10)



Expectation

Random variables, by definition, result in different outcomes

The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables
- **Probability Density function** for Continuous variables

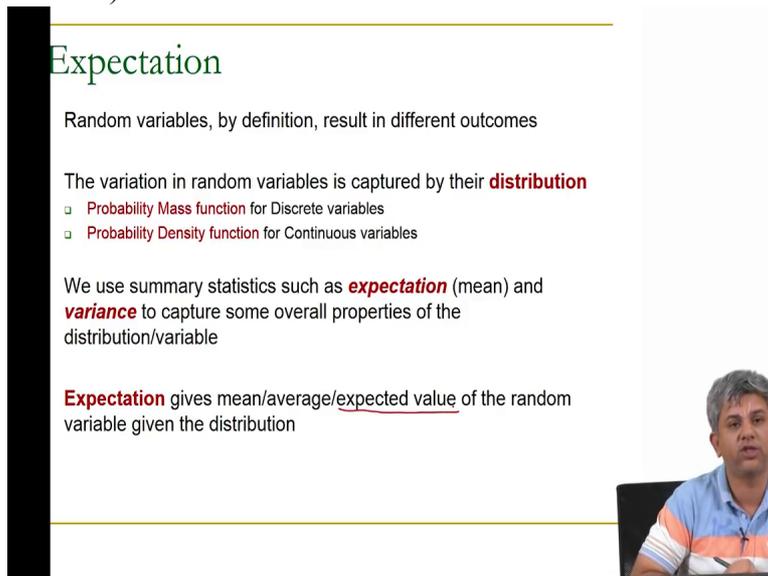
We use summary statistics such as **expectation** (mean) and **variance** to capture some overall properties of the distribution/variable

Expectation gives mean/average/expected value of the random variable given the distribution



Now the expectation gives the mean average or expected value of the random variable

(Refer Slide Time: 02:17)



Expectation

Random variables, by definition, result in different outcomes

The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables
- **Probability Density function** for Continuous variables

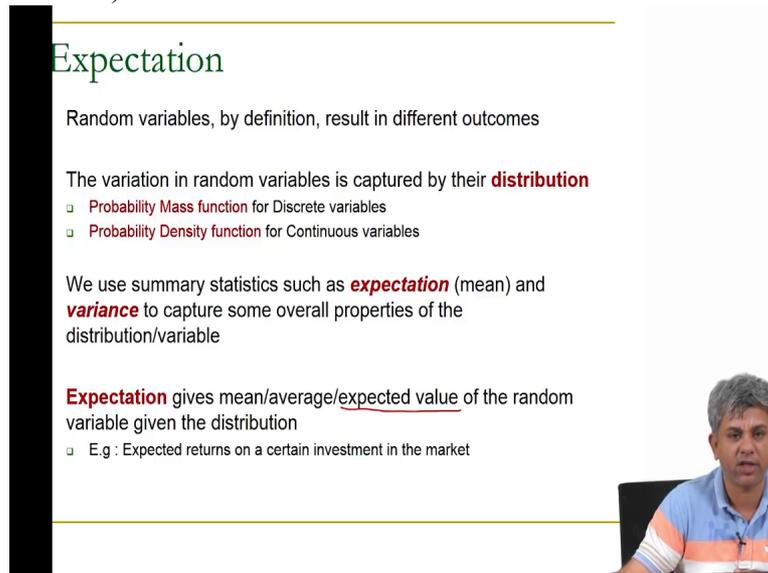
We use summary statistics such as **expectation** (mean) and **variance** to capture some overall properties of the distribution/variable

Expectation gives mean/average/expected value of the random variable given the distribution



once you know the distribution, Ok, so that is important. You need to know what the distribution of the random variable is. Then you can find out the

(Refer Slide Time: 02:24)



Expectation

Random variables, by definition, result in different outcomes

The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables
- **Probability Density function** for Continuous variables

We use summary statistics such as **expectation** (mean) and **variance** to capture some overall properties of the distribution/variable

Expectation gives mean/average/expected value of the random variable given the distribution

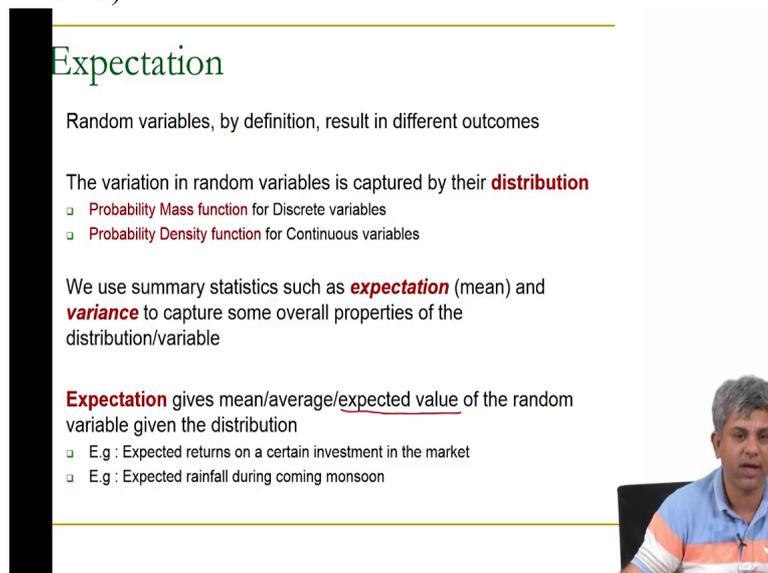
- E.g : Expected returns on a certain investment in the market

expectation.

So here are the couple of examples. You can say you know I have invested in the stock market. What are my expected returns? Obviously you know the returns are not fixed. It is the random variable. But nonetheless, given the certain investment what is my expected return from the stock market?

Another

(Refer Slide Time: 02:42)



Expectation

Random variables, by definition, result in different outcomes

The variation in random variables is captured by their **distribution**

- **Probability Mass function** for Discrete variables
- **Probability Density function** for Continuous variables

We use summary statistics such as **expectation** (mean) and **variance** to capture some overall properties of the distribution/variable

Expectation gives mean/average/expected value of the random variable given the distribution

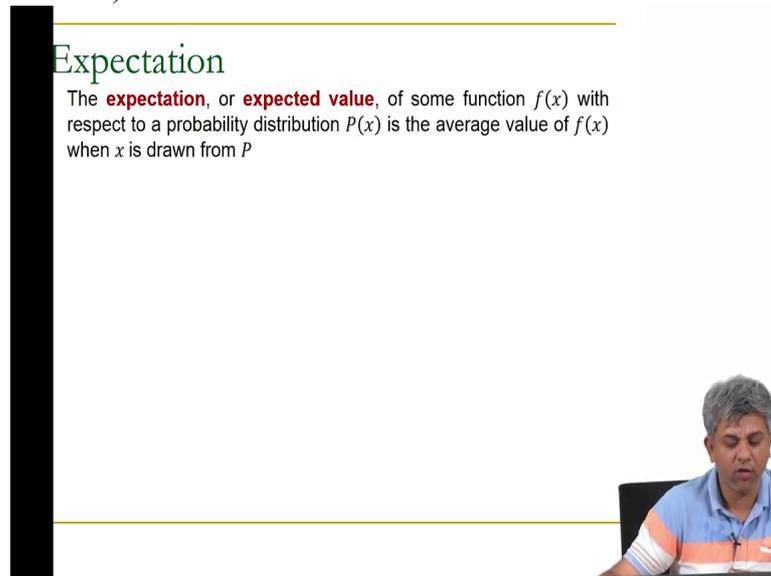
- E.g : Expected returns on a certain investment in the market
- E.g : Expected rainfall during coming monsoon

example is you know I know that monsoon is going to hit. What is going to be the expected rainfall during the coming monsoon? Ok. So you could ask questions of that sort. Again this

is a random variable but overall you would still like to know what is going to be my expected crop yield, a farmer might be interested in knowing.

So here

(Refer Slide Time: 03:03)



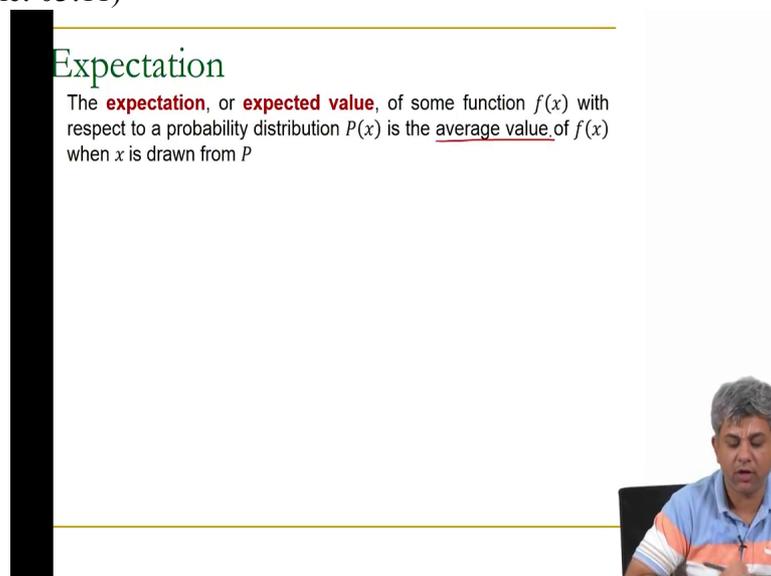
Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P

The slide features a title 'Expectation' in green, followed by a definition in black text. A video inset in the bottom right corner shows a man with grey hair wearing a blue and orange striped shirt, sitting at a desk and looking down at a document.

is how we define expectation, Ok. This is the average value.

(Refer Slide Time: 03:11)



Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P

This slide is identical to the one above, but the word 'average value' in the definition is underlined. The video inset shows the same man as in the previous slide.

So right now I am not talking about only expectation of a random variable x , but that of a function of x . So the expectation or the expected value, both these terminologies are used for some function x of a random variable where x is a random variable,

(Refer Slide Time: 03:28)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P .

Handwritten notes: "Random variable" with an arrow pointing to $f(x)$.

Video inset: A man in a blue and orange striped shirt sitting at a desk.

is the average value of f of x when x is drawn from a probability distribution P .

Recall the notation x is drawn from P is written as $x \sim P$.

(Refer Slide Time: 03:41)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$].

Handwritten notes: "Random variable" with an arrow pointing to $f(x)$.

Video inset: A man in a blue and orange striped shirt sitting at a desk.

(Refer Slide Time: 03:44)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P $[x \sim P]$

enoted by $\mathbb{E}_{x \sim P}[f(x)]$

random variable



So the denotation is, how the notation that we use is x , expectation of f of x , notice the square brackets, expectation of f of x , because it is not quite a function so expectation of f of x given that x is

(Refer Slide Time: 04:01)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P $[x \sim P]$

enoted by $\mathbb{E}_{x \sim P}[f(x)]$

random variable



drawn from P , Ok so this is the notation that we will be using. This is the most detailed or rigorous notation,

(Refer Slide Time: 04:18)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$].

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed/rigorous notation.

Handwritten notes: "random variable" with an arrow pointing to $f(x)$.

Video inset: A man in a blue and orange striped shirt speaking.

Ok.

But more often than not,

(Refer Slide Time: 04:20)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$].

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed/rigorous notation.

- If P is clear from the context $\mathbb{E}_x[f(x)]$

Handwritten notes: "random variable" with an arrow pointing to $f(x)$.

Video inset: A man in a blue and orange striped shirt speaking.

we use some shortcuts. If you know which probability distribution we are talking about, we simply say expectation of x given f of x .

(Refer Slide Time: 04:29)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$].

noted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed / rigorous notation

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$

Handwritten notes: "random variable" with an arrow pointing to $f(x)$.



Even more simply if you also know what x is, you know which random variable x we are considering you can simply say expectation of f of x

(Refer Slide Time: 04:38)

Expectation

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$].

noted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed / rigorous notation

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Handwritten notes: "random variable" with an arrow pointing to $f(x)$.



and sometimes you simply say expectation of f , Ok. So all these notations are used as far as expectation is concerned.

(Refer Slide Time: 04:46)

Expectation random variable

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$]

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed / rigorous notation

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,



So mathematically how do we calculate expectation?

(Refer Slide Time: 04:49)

Expectation random variable

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$]

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed / rigorous notation

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

Discrete



For a discrete

(Refer Slide Time: 04:51)

Expectation random variable

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$]

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed/rigorous notation

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

Discrete

$$\mathbb{E}_{x \sim P}[f(x)] = \sum_x P(x)f(x)$$


variable it is simply summation of P of f , P of x f of x , Ok and if it is a continuous

(Refer Slide Time: 05:00)

Expectation random variable

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$]

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — Most detailed/rigorous notation

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

Discrete

$$\mathbb{E}_{x \sim P}[f(x)] = \sum_x P(x)f(x)$$

Continuous



variable it

(Refer Slide Time: 05:02)

Expectation *random variable*

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$]

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — *Most detailed / rigorous notation*

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

Discrete $\mathbb{E}_{x \sim P}[f(x)] = \sum_x P(x)f(x)$

Continuous $\mathbb{E}_{x \sim P}[f(x)] = \int_x p(x)f(x)dx$



is, remember now this is a probability mass function.

(Refer Slide Time: 05:07)

Expectation *random variable*

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P [$x \sim P$]

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ — *Most detailed / rigorous notation*

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

Discrete $\mathbb{E}_{x \sim P}[f(x)] = \sum_x P(x)f(x)$ *p.m.f*

Continuous $\mathbb{E}_{x \sim P}[f(x)] = \int_x p(x)f(x)dx$



In this case it is a probability density function which is why

(Refer Slide Time: 05:10)

Expectation *random variable*

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P *[$x \sim P$]*

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ *Most detailed / rigorous notation*

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

Discrete $\mathbb{E}_{x \sim P}[f(x)] = \sum_x P(x)f(x)$ *p.m.f*

Continuous $\mathbb{E}_{x \sim P}[f(x)] = \int_x p(x)f(x)dx$ *p.d.f*



you have to multiply by dx in order to get the probability, so integral of P of x dx f of x integrated over all possible values. Now we will see a couple of

(Refer Slide Time: 05:26)

Expectation *random variable*

The **expectation**, or **expected value**, of some function $f(x)$ with respect to a probability distribution $P(x)$ is the average value of $f(x)$ when x is drawn from P *[$x \sim P$]*

denoted by $\mathbb{E}_{x \sim P}[f(x)]$ *Most detailed / rigorous notation*

- If P is clear from the context $\mathbb{E}_x[f(x)]$
- If x is also clear from the context $\mathbb{E}[f(x)]$
- Sometimes, simply denoted as $\mathbb{E}[f]$

Mathematically,

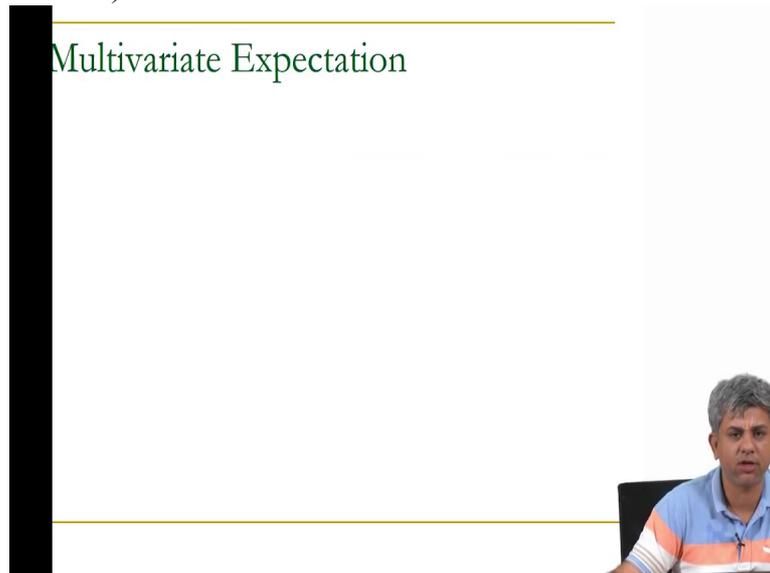
Discrete $\mathbb{E}_{x \sim P}[f(x)] = \sum_x P(x)f(x)$ *p.m.f*

Continuous $\mathbb{E}_{x \sim P}[f(x)] = \int_x p(x)f(x)dx$ *All possible values*



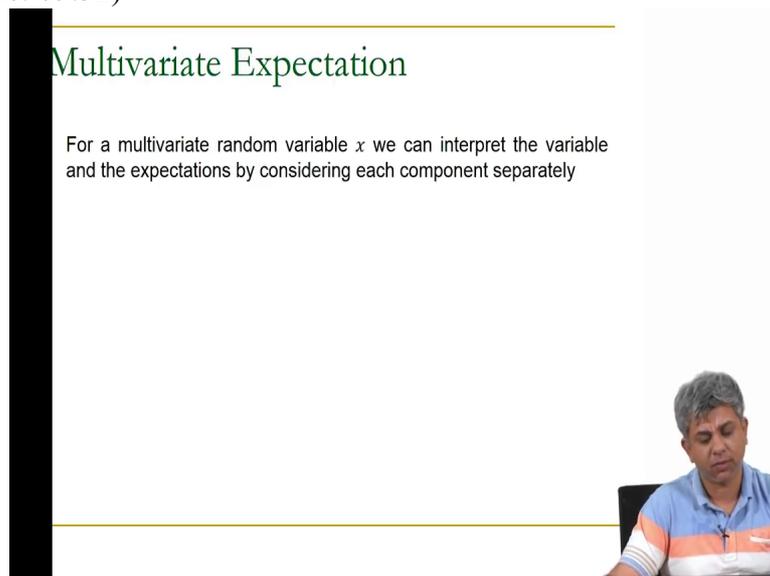
very, very trivial and simple examples on

(Refer Slide Time: 05:29)



the next slide.

(Refer Slide Time: 05:32)



So a generalization of expectation which we just saw expectation of a single variable. But usually especially within machine learning we are dealing with vectors, Ok. So this is called multivariate expectation where basically x is now a vector consisting of x_1, x_2 so on and so forth up to x_n ,

(Refer Slide Time: 06:00)

Multivariate Expectation

For a multivariate random variable x we can interpret the variable and the expectations by considering each component separately

Handwritten: vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$



Ok.

As we saw in the previous video this could be a image, this could be temperature, pressure, velocity etc. It could be any number of variables, Ok. So if you have that then you can

(Refer Slide Time: 06:15)

Multivariate Expectation

For a multivariate random variable x we can interpret the variable and the expectations by considering each component separately

that is, if $x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_D \end{bmatrix} \in \mathbb{R}^D$ then

Handwritten: vector $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$



consider each component separately.

(Refer Slide Time: 06:18)

Multivariate Expectation

For a multivariate random variable x we can interpret the variable and the expectations by considering each component separately

that is, if $x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_D \end{bmatrix} \in \mathbb{R}^D$ then

$$\mathbb{E}_x[f(x)] = \begin{bmatrix} \mathbb{E}_{x_1}[f(x_1)] \\ \mathbb{E}_{x_2}[f(x_2)] \\ \dots \\ \mathbb{E}_{x_D}[f(x_D)] \end{bmatrix}$$


And all you will do is expectation of variable 1, expectation of variable 2, so on so forth, the reason why I wrote it here is notice

(Refer Slide Time: 06:28)

Multivariate Expectation

For a multivariate random variable x we can interpret the variable and the expectations by considering each component separately

that is, if $x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_D \end{bmatrix} \in \mathbb{R}^D$ then

$$\mathbb{E}_x[f(x)] = \begin{bmatrix} \mathbb{E}_{x_1}[f(x_1)] \\ \mathbb{E}_{x_2}[f(x_2)] \\ \dots \\ \mathbb{E}_{x_D}[f(x_D)] \end{bmatrix}$$


that the variable itself, the first one is expectation over variable 1, Ok.

So if I want, let us say my x vector is temperature is temperature, pressure, humidity

(Refer Slide Time: 06:47)

Multivariate Expectation

For a multivariate random variable x we can interpret the variable and the expectations by considering each component separately

that is, if $x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_D \end{bmatrix} \in \mathbb{R}^D$ then

$$\mathbb{E}_x[f(x)] = \begin{bmatrix} \mathbb{E}_{x_1}[f(x_1)] \\ \mathbb{E}_{x_2}[f(x_2)] \\ \dots \\ \mathbb{E}_{x_D}[f(x_D)] \end{bmatrix}$$

Handwritten notes: vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$
 $\vec{x} = \begin{bmatrix} \text{Temp} \\ \text{Press} \\ \text{Humidity} \end{bmatrix}$



then I will take expectation over all possible values of temperature of, if you are interested in some function of temperature, so on and so forth, expectation over all possible values of pressure, function of pressure, Ok.

(Refer Slide Time: 07:06)

Multivariate Expectation

For a multivariate random variable x we can interpret the variable and the expectations by considering each component separately

that is, if $x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_D \end{bmatrix} \in \mathbb{R}^D$ then

$$\mathbb{E}_x[f(x)] = \begin{bmatrix} \mathbb{E}_{x_1}[f(x_1)] \\ \mathbb{E}_{x_2}[f(x_2)] \\ \dots \\ \mathbb{E}_{x_D}[f(x_D)] \end{bmatrix}$$

Handwritten notes: vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$
 $\vec{x} = \begin{bmatrix} \text{Temp} \\ \text{Press} \\ \text{Humidity} \end{bmatrix}$
 $\begin{bmatrix} \mathbb{E}_{\text{Temp}}[f(\text{Temp})] \\ \mathbb{E}_{\text{Press}}[f(\text{Press})] \\ \vdots \end{bmatrix}$



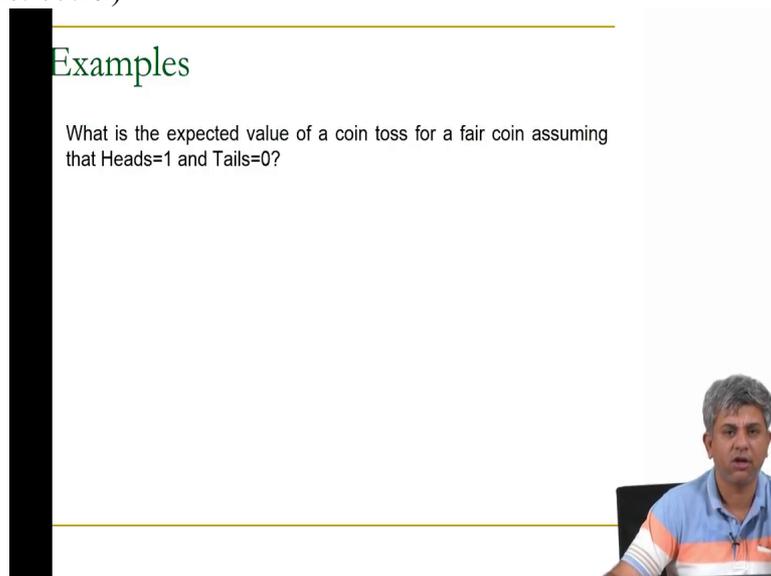
So that is multivariate expectation. Multivariate simply means multiple variables, Ok. It is not a single scalar, it is a vector.

(Refer Slide Time: 07:15)



So here are some trivial examples. I am going to do univariate examples here

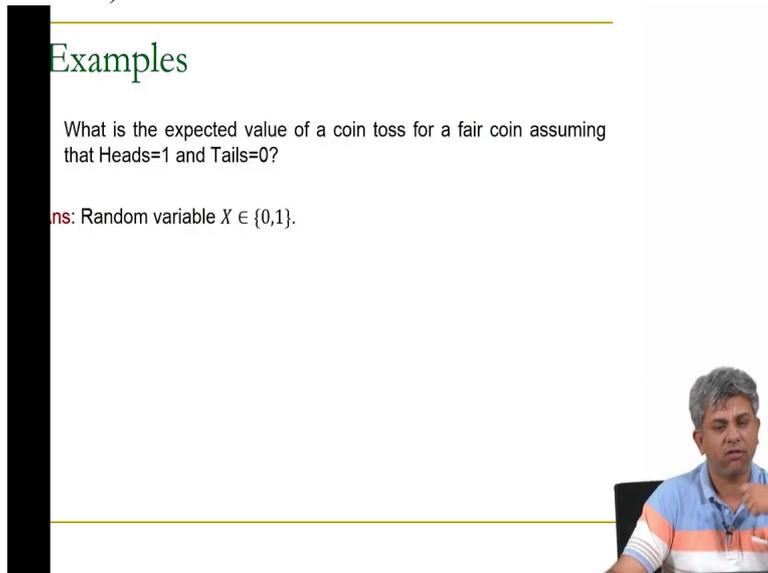
(Refer Slide Time: 07:19)



Ok.

So all of us know, let us say you want to know the expected value of a toss of a coin, for a fair coin assuming that heads has a value 1 and tails has a value 0,

(Refer Slide Time: 07:30)



Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

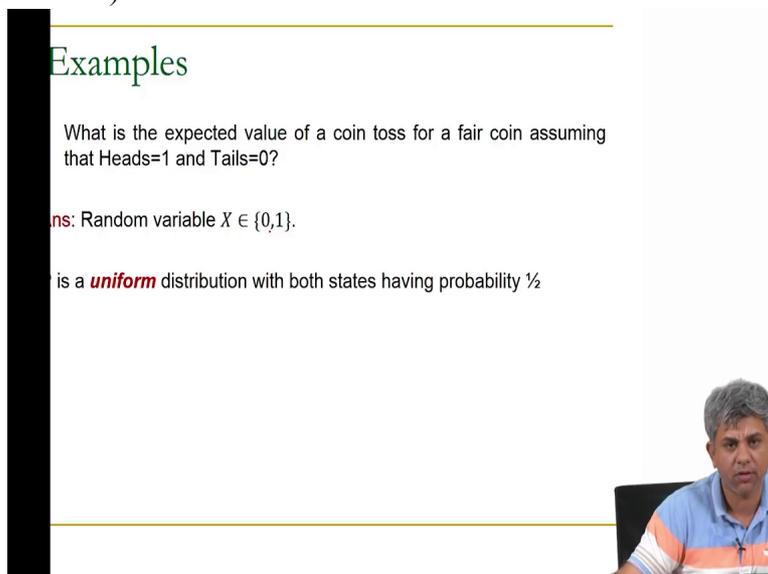
Ans: Random variable $X \in \{0,1\}$.



then I will just do it in detail so that you get used to this kind of calculation if you are not already used to it.

So you first identify the random variable you are considering. Here the random variable is the result of the toss of the coin and it now belongs to 0 1,

(Refer Slide Time: 07:45)



Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

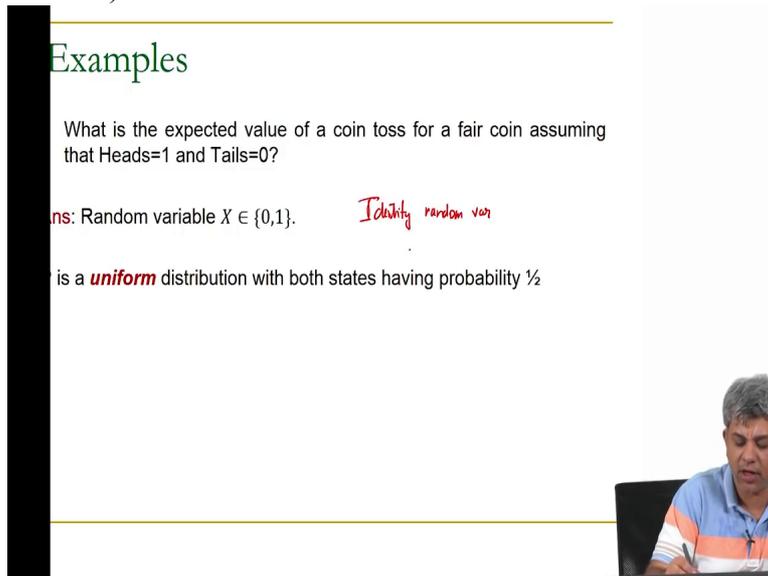
Ans: Random variable $X \in \{0,1\}$.

is a **uniform** distribution with both states having probability $\frac{1}{2}$



Ok. What is P, you now need to know what the distribution P is. So first identify random variable.

(Refer Slide Time: 07:59)



Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

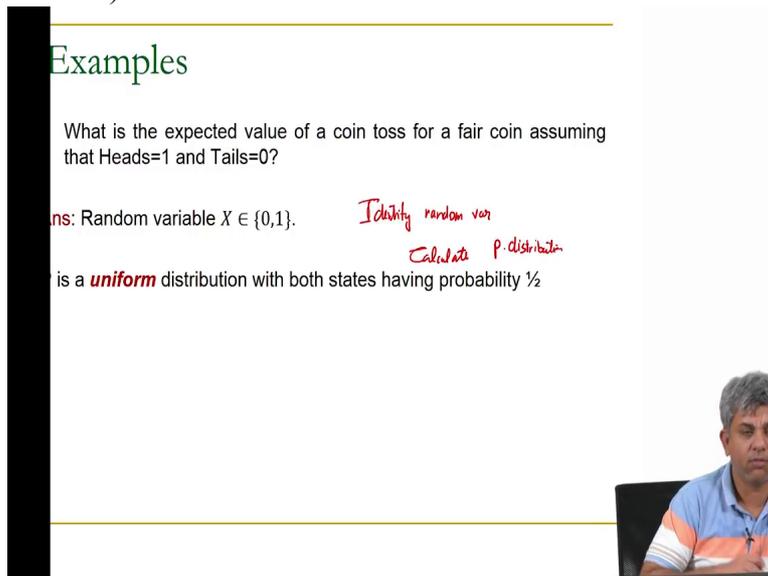
ns: Random variable $X \in \{0,1\}$. *Identity random var*

is a **uniform** distribution with both states having probability $\frac{1}{2}$



Next calculate the probability distribution. In this case it is a mass distribution

(Refer Slide Time: 08:10)



Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

ns: Random variable $X \in \{0,1\}$. *Identity random var*
Calculate P-distribution

is a **uniform** distribution with both states having probability $\frac{1}{2}$



because it is a discrete random variable.

Now probability distribution is very simple. If I have x and P of x then x takes the value 0 with the probability half, x takes the value 1

(Refer Slide Time: 08:25)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

x	0	1
P(x)	$\frac{1}{2}$	$\frac{1}{2}$

Ans: Random variable $X \in \{0,1\}$. Identify random var

is a **uniform** distribution with both states having probability $\frac{1}{2}$. Calculate P-distribution



also with the probability half.

(Refer Slide Time: 08:30)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

x	0	1
P(x)	$\frac{1}{2}$	$\frac{1}{2}$

Ans: Random variable $X \in \{0,1\}$. Identify random var

is a **uniform** distribution with both states having probability $\frac{1}{2}$. Calculate P-distribution

So, $E_{x \sim P}[x] = \sum_x xP(x) = \left[0 \times \frac{1}{2} + 1 \times \frac{1}{2}\right] = \frac{1}{2}$



So if you find out the expectation, it is simply 0 times half plus 1 times half which is equal to half, Ok.

(Refer Slide Time: 08:38)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

Ans: Random variable $X \in \{0,1\}$. *Identify random var*

is a **uniform** distribution with both states having probability $\frac{1}{2}$. *Calculate P-distribution*

So, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[0 \times \frac{1}{2} + 1 \times \frac{1}{2}\right] = \frac{1}{2}$ ✓

x	0	1
P(x)	$\frac{1}{2}$	$\frac{1}{2}$



All of us know this. Another way to look at it is the average value of the toss that you will obtain is basically going to be half. So notice that the expectation, even though we have given it as the average value, or the expected value, obviously you cannot say that the expected value of a toss of a coin is half.

Because neither heads nor tails is the actual half. It just represents an average or a weighted average of values that come out, Ok.

(Refer Slide Time: 09:05)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

Ans: Random variable $X \in \{0,1\}$. *Identify random var*

is a **uniform** distribution with both states having probability $\frac{1}{2}$. *Calculate P-distribution*

So, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[0 \times \frac{1}{2} + 1 \times \frac{1}{2}\right] = \frac{1}{2}$ ✓

Similarly, the expected value of a fair dice throw is?

x	0	1
P(x)	$\frac{1}{2}$	$\frac{1}{2}$



Similarly, if you want to find out the expected value of a fair dice

(Refer Slide Time: 09:09)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

Ans: Random variable $X \in \{0,1\}$. *Identify random var*

is a **uniform** distribution with both states having probability $\frac{1}{2}$. *Calculate P-distribution*

So, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[0 \times \frac{1}{2} + 1 \times \frac{1}{2}\right] = \frac{1}{2}$ ✓

Similarly, the expected value of a fair dice throw is?

Random variable $X \in \{1,2,3,4,5,6\}$. P is uniform with probability $\frac{1}{6}$



throw this is going to be the average of 1, 2, 3, 4, 5, 6.

(Refer Slide Time: 09:14)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

Ans: Random variable $X \in \{0,1\}$. *Identify random var*

is a **uniform** distribution with both states having probability $\frac{1}{2}$. *Calculate P-distribution*

So, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[0 \times \frac{1}{2} + 1 \times \frac{1}{2}\right] = \frac{1}{2}$ ✓

Similarly, the expected value of a fair dice throw is?

Random variable $X \in \{1,2,3,4,5,6\}$. P is uniform with probability $\frac{1}{6}$

So, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \dots + 6 \times \frac{1}{6}\right] = 3.5$



Given that all of these are equally probable, assuming this is a fair dice and it is not like sort of a loaded dice or something, again from the same idea you get 3 point 5.

(Refer Slide Time: 09:23)

Examples

What is the expected value of a coin toss for a fair coin assuming that Heads=1 and Tails=0?

Ans: Random variable $X \in \{0,1\}$. *Identify random var*

is a **uniform** distribution with both states having probability $\frac{1}{2}$. *Calculate P distribution*

o, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[0 \times \frac{1}{2} + 1 \times \frac{1}{2}\right] = \frac{1}{2}$ ✓

Similarly, the expected value of a fair dice throw is?

random variable $X \in \{1,2,3,4,5,6\}$. P is uniform with probability $\frac{1}{6}$

o, $\mathbb{E}_{x \sim P}[x] = \sum_x xP(x) = \left[1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \dots + 6 \times \frac{1}{6}\right] = 3.5$ ✓



Let us look at a slightly more complex

(Refer Slide Time: 09:26)

Examples

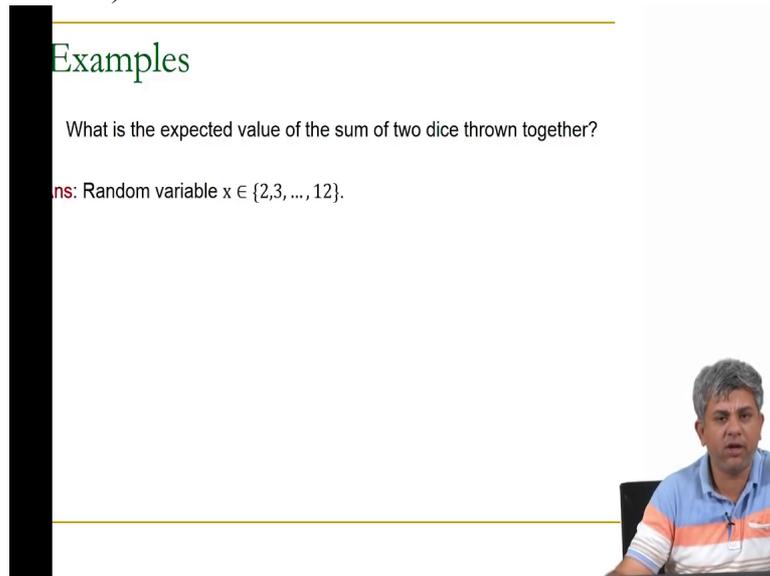
What is the expected value of the sum of two dice thrown together?



example. This is just to see, you know may be a slight increase over simple averages. Ok so what is the expected value of the sum of two dice thrown together?

Now the random variable here

(Refer Slide Time: 09:40)



Examples

What is the expected value of the sum of two dice thrown together?

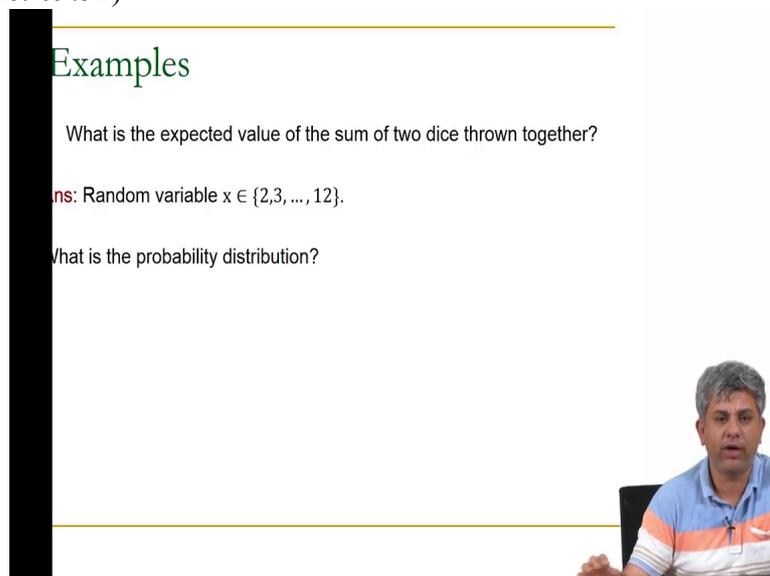
Ans: Random variable $x \in \{2, 3, \dots, 12\}$.

A video inset shows a man with grey hair wearing a blue and orange striped polo shirt, sitting at a desk.

is any number between 2 and 12, obviously 1 cannot occur if you are throwing 2 dice and taking the sum.

So you have the variable that goes between

(Refer Slide Time: 09:51)



Examples

What is the expected value of the sum of two dice thrown together?

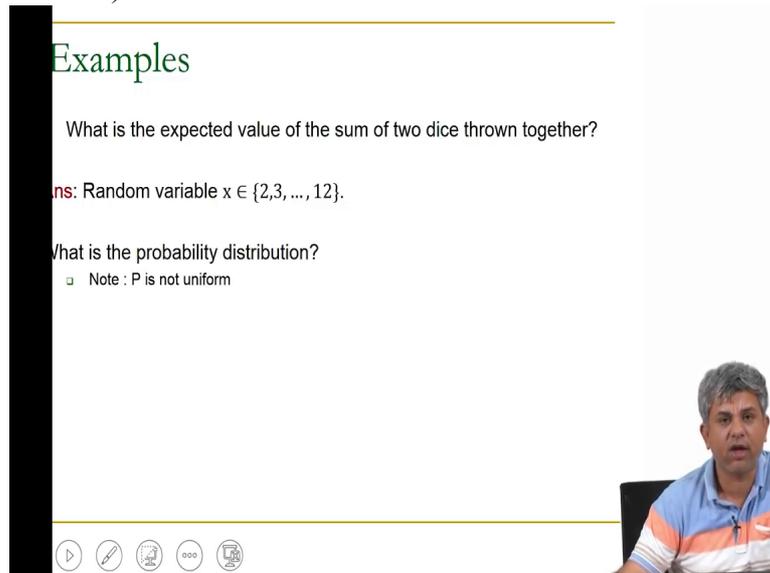
Ans: Random variable $x \in \{2, 3, \dots, 12\}$.

What is the probability distribution?

A video inset shows the same man from the previous slide, sitting at a desk.

2 and 12. The probability distribution you have to be a little careful now. Ok. Now notice that the probability of 2 is not the same as the probability of 3,

(Refer Slide Time: 10:03)



Examples

What is the expected value of the sum of two dice thrown together?

Ans: Random variable $x \in \{2,3, \dots, 12\}$.

What is the probability distribution?

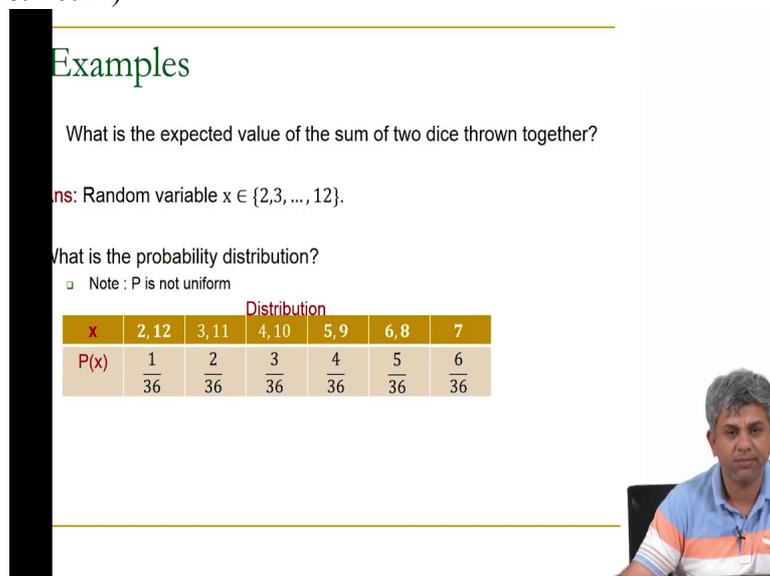
- Note : P is not uniform

The slide features a presenter in the bottom right corner and a navigation bar at the bottom with icons for back, forward, search, and refresh.

Ok, unlike the previous case where we had uniform probability distributions; in this case each of these probabilities is different, Ok.

So here is the distribution.

(Refer Slide Time: 10:14)



Examples

What is the expected value of the sum of two dice thrown together?

Ans: Random variable $x \in \{2,3, \dots, 12\}$.

What is the probability distribution?

- Note : P is not uniform

x	2, 12	3, 11	4, 10	5, 9	6, 8	7
P(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$

The slide features a presenter in the bottom right corner and a navigation bar at the bottom with icons for back, forward, search, and refresh.

Both 2 and 12, if x is 2 as well as x is 12, both of these can occur in only one way, and so for 2, you need 1, 1 and for 12 you need 6, 6 and that can occur in 1 by 6 multiplied by 1 by 6 which is 1 by 36.

Similarly, 3 can occur

(Refer Slide Time: 10:34)

Examples

What is the expected value of the sum of two dice thrown together?

Ans: Random variable $x \in \{2,3, \dots, 12\}$.

What is the probability distribution?

- Note : P is not uniform

x	2, 12	3, 11	4, 10	5, 9	6, 8	7
P(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$



in two ways which is 1 comma 2, 2 comma 1, Ok so therefore you get 2 by 36. Similarly, for 11.

For 4 you have three ways, for 5 you have four ways, for 6 you have five ways, you know 1, 5; 2,4; 3,3 and 5,1 and 4,2, that put together. 7 can actually occur in six different ways and so these are the probabilities.

So notice that unlike the previous examples that we took, in this case, P of x is an actual non-uniform distribution.

(Refer Slide Time: 11:06)

Examples

What is the expected value of the sum of two dice thrown together?

Ans: Random variable $x \in \{2,3, \dots, 12\}$.

What is the probability distribution?

- Note : P is not uniform

x	2, 12	3, 11	4, 10	5, 9	6, 8	7
P(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$

So, $E_{x \sim P}[x] = \sum_x xP(x) = \left[2 \times \frac{1}{36} + 3 \times \frac{2}{36} + \dots + 12 \times \frac{1}{36} \right] = 7$



So if you find out expectation, the same thing, so 2 multiplied by 1 by 36, 3 multiplied by 2 by 36, so on and so forth up till 12 multiplied by 1 by 36, if we calculate it, it comes out to be 7.

But the calculation is a little bit lengthy, Ok. So the question is,

(Refer Slide Time: 11:26)

Examples

What is the expected value of the sum of two dice thrown together?

Ans: Random variable $x \in \{2, 3, \dots, 12\}$.

What is the probability distribution?

Note: P is not uniform

x	2, 12	3, 11	4, 10	5, 9	6, 8	7
P(x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$

So, $E_{x \sim P}[x] = \sum_x xP(x) = \left[2 \times \frac{1}{36} + 3 \times \frac{2}{36} + \dots + 12 \times \frac{1}{36} \right] = 7$

Question: Is there an easier way of calculating this case?

is there an easier way of calculating this case?

So for this we use a

(Refer Slide Time: 11:31)

Linearity of Expectation

simple idea called the linearity of expectation,

(Refer Slide Time: 11:35)

Linearity of Expectation

Important Property of expectation

A video inset in the bottom right corner shows a man with grey hair wearing a blue and orange striped polo shirt, sitting at a desk.

Ok. This is an extremely important property of expectations.

(Refer Slide Time: 11:39)

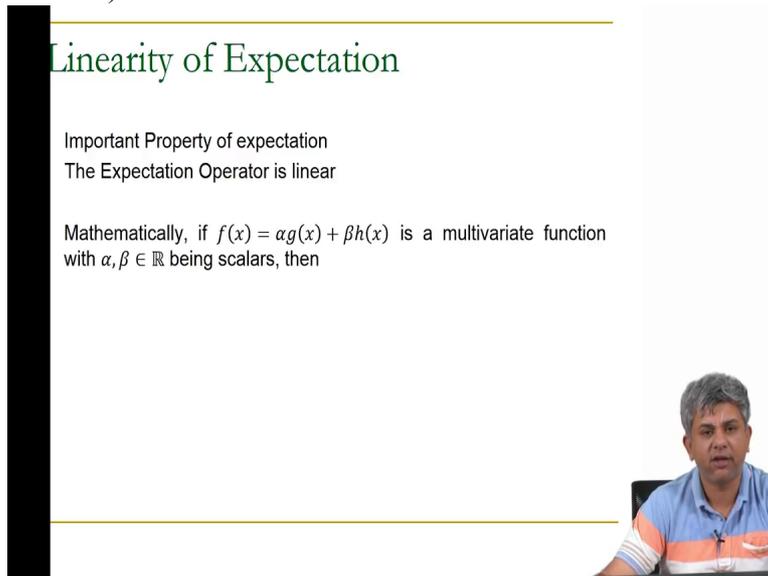
Linearity of Expectation

Important Property of expectation
The Expectation Operator is linear

A video inset in the bottom right corner shows the same man from the previous slide, looking down at his desk.

The idea is that

(Refer Slide Time: 11:42)



Linearity of Expectation

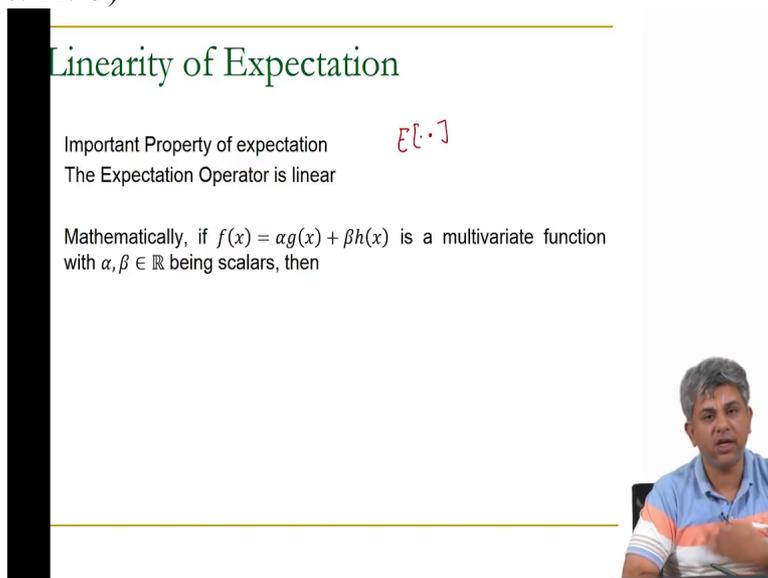
Important Property of expectation
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

[A video inset in the bottom right corner shows a man with grey hair wearing a blue and orange striped polo shirt, speaking.]

the expectation operator, so this thing, this is a linear operator.

(Refer Slide Time: 11:49)



Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

[A video inset in the bottom right corner shows the same man from the previous slide, speaking.]

What do I mean by linear?

That is if you have f which is a linear combination, please remember linear combination from our discussion, if f is a linear combination

(Refer Slide Time: 12:04)

The slide is titled "Linearity of Expectation" in green. It contains the following text: "Important Property of expectation", "The Expectation Operator is linear", and "Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then". There are handwritten red notes: " $E[\cdot]$ " next to the first line, and " $\alpha g(x)$ " above the function definition. A video feed of a presenter is visible in the bottom right corner.

of two other functions g and h , α and β let us say are scalars, so α times g of x plus β times h of x is f of x

(Refer Slide Time: 12:13)

The slide is titled "Linearity of Expectation" in green. It contains the same text as the previous slide, but with the formula $E[f] = \alpha E[g] + \beta E[h]$ added below the text. There are handwritten red notes: " $E[\cdot]$ " next to the first line, and " $\alpha g(x)$ " above the function definition. A video feed of a presenter is visible in the bottom right corner.

then expectation of f can be written as α times expectation of g plus β times expectation of h , Ok.

So this is an important property. We will just prove it in the next slide,

(Refer Slide Time: 12:29)

Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$E[f] = \alpha E[g] + \beta E[h]$ Note the use of compact notation

When Comp



Ok. Also notice that I have used a compact notation here instead of writing expectation of f etc, I simply used expectation of f is alpha times $E[g]$ plus beta times $E[h]$.

(Refer Slide Time: 12:44)

Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$E[f] = \alpha E[g] + \beta E[h]$ Note the use of compact notation

Applying this to our example, we note that $X = D_1 + D_2$ where D_1 and D_2 are the number obtained on the first and second dice respectively.



So you can apply this. I will prove this shortly but before that let us simply apply this to our two dice case. Remember that in order to find out the expectation of the two dice we actually had to find out first the probability distribution of each of those occurrences, different outcomes and then we had to do the expectation calculation, Ok.

But suppose we notice that the two dice are essentially two different random variables coming together, one is D_1 and one is D_2 , where D_1 is the value that you got out of the first dice, and D_2 is the value that you got out of the second dice, Ok.

(Refer Slide Time: 13:21)

Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then
 $E[f] = \alpha E[g] + \beta E[h]$ Note the use of compact notation

Applying this to our example, we note that $X = D_1 + D_2$ where D_1 and D_2 are the number obtained on the first and second dice respectively.

Then, $E[X] = E[D_1] + E[D_2] = 3.5 + 3.5 = 7$

Handwritten notes: $E[\cdot]$ (top right), $E[X] = E[D_1 + D_2] = E[D_1] + E[D_2]$ (bottom right)

Video inset: A lecturer in a blue and orange striped shirt is speaking.

So through our linearity we can write expectation of x is expectation of D_1 plus D_2 which by linearity is equal to expectation of D_1 plus expectation of

(Refer Slide Time: 13:37)

Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then
 $E[f] = \alpha E[g] + \beta E[h]$ Note the use of compact notation

Applying this to our example, we note that $X = D_1 + D_2$ where D_1 and D_2 are the number obtained on the first and second dice respectively.

Then, $E[X] = E[D_1] + E[D_2] = 3.5 + 3.5 = 7$

Handwritten notes: $E[X] = E[D_1 + D_2] = E[D_1] + E[D_2]$ (bottom right)

Video inset: A lecturer in a blue and orange striped shirt is speaking.

D_2 and we know this already because the expected value from one dice is 3 point 5,

(Refer Slide Time: 13:44)

Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then
 $E[f] = \alpha E[g] + \beta E[h]$ Note the use of compact notation

Applying this to our example, we note that $X = D_1 + D_2$ where D_1 and D_2 are the number obtained on the first and second dice respectively.
 $E[X] = E[D_1 + D_2] = E[D_1] + E[D_2]$

Then, $E[X] = E[D_1] + E[D_2] = 3.5 + 3.5 = 7$

the expected value from the other dice is also 3 point 5 so the expectations add up and you can see that this is a remarkably simple calculation,

(Refer Slide Time: 13:52)

Linearity of Expectation

Important Property of expectation $E[\cdot]$
The Expectation Operator is linear

Mathematically, if $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then
 $E[f] = \alpha E[g] + \beta E[h]$ Note the use of compact notation

Applying this to our example, we note that $X = D_1 + D_2$ where D_1 and D_2 are the number obtained on the first and second dice respectively.
 $E[X] = E[D_1 + D_2] = E[D_1] + E[D_2]$

Then, $E[X] = E[D_1] + E[D_2] = 3.5 + 3.5 = 7$

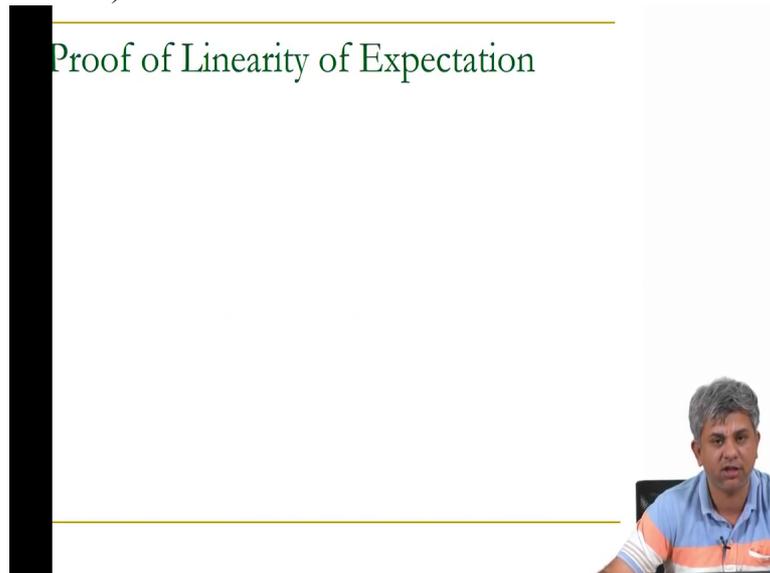
Note: Much simpler, since the distribution of X need not be found

Ok.

This is a much, much simpler calculation compared to actually finding out the overall, you know probability of, probability distribution of x .

So that is the advantage of using linearity of expectations

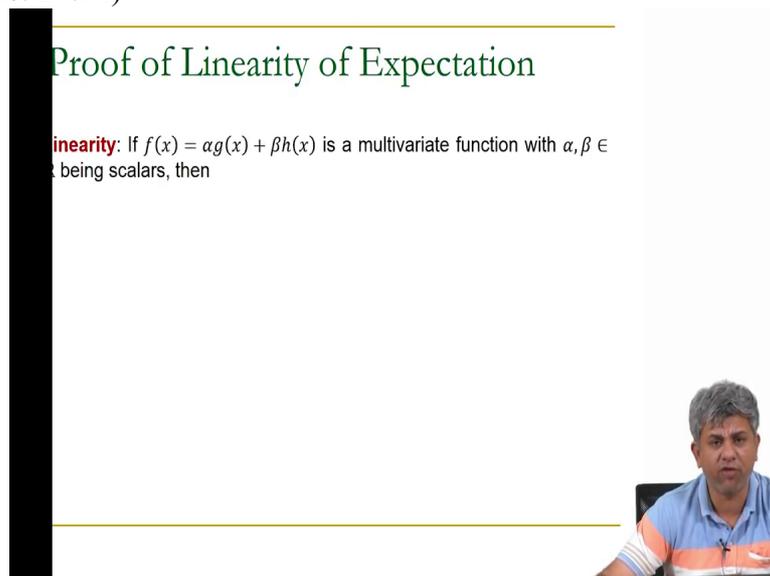
(Refer Slide Time: 14:07)



The slide features a title "Proof of Linearity of Expectation" in green text at the top. A video inset in the bottom right corner shows a man with grey hair wearing a blue and orange striped polo shirt, speaking. The slide content is otherwise blank.

and this is a very commonly used property. Let me give a very quick

(Refer Slide Time: 14:14)



The slide features a title "Proof of Linearity of Expectation" in green text at the top. Below the title, the text reads: "Linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then". A video inset in the bottom right corner shows the same man from the previous slide speaking.

proof. I will be doing it just for continuous variables. So assume that f of x is a multivariate function. This is true not only for univariate but for multivariate functions. And

(Refer Slide Time: 14:26)

Linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h]$$

it is a linear combination of g and h.

We want to prove, this is our claim, that the expectation of

(Refer Slide Time: 14:34)

Linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Claim}$$

f is alpha E g plus beta E h. So here

(Refer Slide Time: 14:38)

Linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Clear}$$

proof: For continuous distributions



is a quick proof

(Refer Slide Time: 14:40)

Linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Clear}$$

proof: For continuous distributions

$$\begin{aligned} \mathbb{E}[f] &= \int f(x)p(x)dx \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ \Rightarrow \mathbb{E}[f] &= \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \end{aligned}$$


So assuming that this is a continuous distribution, you can write E of f as integral of f of x p of x d x, this is of course the definition of expectation.

(Refer Slide Time: 14:53)

Proof of Linearity of Expectation

linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Claim}$$

proof: For continuous distributions

$$\begin{aligned} \mathbb{E}[f] &= \int f(x)p(x)dx \rightarrow \text{Definition} \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ \Rightarrow \mathbb{E}[f] &= \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \end{aligned}$$


Now f of x

(Refer Slide Time: 14:56)

Proof of Linearity of Expectation

linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Claim}$$

proof: For continuous distributions

$$\begin{aligned} \mathbb{E}[f] &= \int f(x)p(x)dx \rightarrow \text{Definition} \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ \Rightarrow \mathbb{E}[f] &= \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \end{aligned}$$


is alpha g plus beta h, Ok. Now we can split this integral into two different integrals.

Since alpha is a constant you can take it out of the integral. alpha does not depend on h, x. So alpha times integral g x p x d x plus beta times integral h x p x d x, this of course is the definition of E of g

(Refer Slide Time: 15:20)

Proof of Linearity of Expectation

linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Claim}$$

proof: For continuous distributions

$$\begin{aligned}\mathbb{E}[f] &= \int f(x)p(x)dx \rightarrow \text{Definition} \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ &\Rightarrow \mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h]\end{aligned}$$


and this is the definition of E of h.

(Refer Slide Time: 15:25)

Proof of Linearity of Expectation

linearity: If $f(x) = \alpha g(x) + \beta h(x)$ is a multivariate function with $\alpha, \beta \in \mathbb{R}$ being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Claim}$$

proof: For continuous distributions

$$\begin{aligned}\mathbb{E}[f] &= \int f(x)p(x)dx \rightarrow \text{Definition} \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ &\Rightarrow \mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h]\end{aligned}$$


And therefore you end up at the proof that E of f

(Refer Slide Time: 15:29)

Proof of Linearity of Expectation

linearity: If $f(x) = \alpha g(x) + \beta h(x)$, is a multivariate function with $\alpha, \beta \in \mathbb{R}$ & being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Clear}$$

proof: For continuous distributions

$$\begin{aligned} \mathbb{E}[f] &= \int f(x)p(x)dx \rightarrow \text{Definition} \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ &\Rightarrow \mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \end{aligned}$$

discrete can be proved similarly – Try it as an exercise!



is alpha times E g plus beta times E h.

Now you can also start from the other definition. E is equal to sigma or E of f is equal to sigma f x p x d x and since this is also

(Refer Slide Time: 15:47)

Proof of Linearity of Expectation

linearity: If $f(x) = \alpha g(x) + \beta h(x)$, is a multivariate function with $\alpha, \beta \in \mathbb{R}$ & being scalars, then

$$\mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \rightarrow \text{Clear}$$

proof: For continuous distributions

$$\begin{aligned} \mathbb{E}[f] &= \int f(x)p(x)dx \rightarrow \text{Definition} \quad \mathbb{E}[f] = \sum f(x)p(x) \\ &= \int (\alpha g(x) + \beta h(x))p(x)dx \\ &= \alpha \int g(x)p(x)dx + \beta \int h(x)p(x)dx \\ &\Rightarrow \mathbb{E}[f] = \alpha \mathbb{E}[g] + \beta \mathbb{E}[h] \end{aligned}$$

discrete can be proved similarly – Try it as an exercise!



a summation you can prove a discrete also similarly. You can, I suggest that you try this as an exercise. Thank you.