Discrete Random Variables

October 7, 2010

Random Variables

In many situations, we are interested in numbers associated with the outcomes of a random experiment. For example:

- Testing cars from a production line, we are interested in variables such as average emissions, fuel consumption, acceleration time etc

Random Variables

In many situations, we are interested in numbers associated with the outcomes of a random experiment. For example:

- Testing cars from a production line, we are interested in variables such as average emissions, fuel consumption, acceleration time etc
- A box of 6 eggs is rejected if it contains one or more broken eggs. If we examine 10 boxes of eggs, we may be interested in
    1. X_1 - the number of broken eggs in the 10 boxes
    2. X_2 - the number of boxes rejected

Random Variables

In many situations, we are interested in numbers associated with the outcomes of a random experiment. For example:

- Testing cars from a production line, we are interested in variables such as average emissions, fuel consumption, acceleration time etc
- A box of 6 eggs is rejected if it contains one or more broken eggs. If we examine 10 boxes of eggs, we may be interested in
    1. X_1 - the number of broken eggs in the 10 boxes
    2. X_2 - the number of boxes rejected
- An office phone system has 50 lines available and we are interested in monitoring the number of lines in use at a given time.

Definition

A random variable is a function that maps outcomes of a random experiment to real numbers.

Definition

A random variable is a function that maps outcomes of a random experiment to real numbers.

Example

A fair coin is tossed 6 times. The number of heads that come up is an example of a random variable.

$HHTTHT \to 3$, $THHTTT \to 2$.

Definition

A random variable is a function that maps outcomes of a random experiment to real numbers.

Example

A fair coin is tossed 6 times. The number of heads that come up is an example of a random variable.

$HHTTHT \to 3$, $THHTTT \to 2$.

This random variables can only take values between 0 and 6. The set of possible values of a random variables is known as its Range.

Random Variables

Example

A box of 6 eggs is rejected once it contains one or more broken eggs. If we examine 10 boxes of eggs and define the random variables X_1, X_2 as

1. X_1 - the number of broken eggs in the 10 boxes
2. X_2 - the number of boxes rejected

Random Variables

Example

A box of 6 eggs is rejected once it contains one or more broken eggs. If we examine 10 boxes of eggs and define the random variables X_1, X_2 as

1. X_1 - the number of broken eggs in the 10 boxes
2. X_2 - the number of boxes rejected

Then the range of X_1 is $\{0, 1, 2, \ldots, 59, 60\}$, while the range of X_2 is $0, 1, 2, \ldots, 10\}$.

Continuous and Discrete Random Variables

- If the range of a random variable is finite or countably infinite, it is said to be a discrete random variable. For example - Number of broken eggs in a batch or the number of bits in error in a transmitted message.
- If the range of a random variable is continuous, it is said to be a continuous random variable. For example - the current in a copper wire or the length of a manufactured part.

Continuous and Discrete Random Variables

- If the range of a random variable is finite or countably infinite, it is said to be a discrete random variable. For example - Number of broken eggs in a batch or the number of bits in error in a transmitted message.
- If the range of a random variable is continuous, it is said to be a continuous random variable. For example - the current in a copper wire or the length of a manufactured part.

Random variables are usually denoted by capital letters X. The values of the variables are usually denoted by lower case letters x.

- If the range of a random variable is finite or countably infinite, it is said to be a discrete random variable. For example - Number of broken eggs in a batch or the number of bits in error in a transmitted message.
- If the range of a random variable is continuous, it is said to be a continuous random variable. For example - the current in a copper wire or the length of a manufactured part.

Random variables are usually denoted by capital letters X. The values of the variables are usually denoted by lower case letters x. The notation

$$P(X = x)$$

stands for the probability that the random variable X takes the value x.

Probability Mass Functions

Definition

For a discrete random variable X, its probability mass function $f(\cdot)$ is specified by giving the values $f(x) = P(X = x)$ for all x in the range of X.

Probability Mass Functions

Definition

For a discrete random variable X, its probability mass function $f(\cdot)$ is specified by giving the values $f(x) = P(X = x)$ for all x in the range of X.

Example

What is the probability mass function of the random variable that counts the number of heads on 3 tosses of a fair coin?

Probability Mass Functions

Definition

For a discrete random variable X, its probability mass function $f(\cdot)$ is specified by giving the values $f(x) = P(X = x)$ for all x in the range of X.

Example

What is the probability mass function of the random variable that counts the number of heads on 3 tosses of a fair coin?

The range of the variable is $\{0, 1, 2, 3\}$.

Probability Mass Functions

Definition

For a discrete random variable X, its probability mass function $f(\cdot)$ is specified by giving the values $f(x) = P(X = x)$ for all x in the range of X.

Example

What is the probability mass function of the random variable that counts the number of heads on 3 tosses of a fair coin?

The range of the variable is $\{0, 1, 2, 3\}$.

$$\begin{aligned} P(X = 0) &= (\frac{1}{2})^3 \; P(X = 1) = 3(\frac{1}{2})^3 \\ P(X = 2) &= 3(\frac{1}{2})^3 \; P(X = 3) = (\frac{1}{2})^3 \end{aligned}$$

Example

Consider the following game. A fair 4-sided die, with the numbers $1, 2, 3, 4$ is rolled twice. If the score on the second roll is strictly greater than the score on the first the player wins the difference in euro. If the score on the second roll is strictly less than the score on the first roll, the player loses the difference in euro. If the scores are equal, the player neither wins nor loses. If we let X denote the (possibly negative) winnings of the player, what is the probability mass function of X? (X can take any of the values $-3, -2, -1, 0, 1, 2, 3$.)

Example

Example

The total number of outcomes of the experiment is $4 \times 4 = 16$.

- $P(X = 0)$: X will take the value 0 for the outcomes $(1, 1), (2, 2), (3, 3), (4, 4)$. So $f(0) = \frac{4}{16}$.
- $P(X = 1)$: X will take the value 1 for the outcomes $(1, 2), (2, 3), (3, 4)$. So $f(1) = \frac{3}{16}$.
- $P(X = 2)$: X will take the value 2 for the outcomes $(1, 3), (2, 4)$. So $f(2) = \frac{2}{16}$.
- $P(X = 3)$: Similarly $f(3) = \frac{1}{16}$.

Continuing we find the probability mass function is Continuing in the same way we see that the probability mass function is

x	-3	-2	-1	0	1	2	3
$f(x)$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

Probability Mass Functions

A function f can only be a probability mass function if it satisfies certain conditions.

Probability Mass Functions

A function f can only be a probability mass function if it satisfies certain conditions.

(i) As $f(x)$ represents the probability that the variable X takes the value x, $f(x)$ can never be negative. So $f(x) \geq 0$ for all x.

Probability Mass Functions

A function f can only be a probability mass function if it satisfies certain conditions.

(i) As $f(x)$ represents the probability that the variable X takes the value x, $f(x)$ can never be negative. So $f(x) \geq 0$ for all x.

(ii) Also if we sum over all values of x (in the range of X), the total must be equal to one.

$$\sum_x f(x) = \sum_x P(X = x) = 1.$$

Probability Mass Functions

Example

Suppose the range of a discrete random variable is $\{0, 1, 2, 3, 4\}$. If the probability mass function is $f(x) = cx$ for $x = 0, 1, 2, 3, 4$, what is the value of c?

Probability Mass Functions

Example

Suppose the range of a discrete random variable is $\{0, 1, 2, 3, 4\}$. If the probability mass function is $f(x) = cx$ for $x = 0, 1, 2, 3, 4$, what is the value of c?

- First of all, $c \geq 0$ as $f(x) \geq 0$.

Probability Mass Functions

Example

Suppose the range of a discrete random variable is $\{0, 1, 2, 3, 4\}$. If the probability mass function is $f(x) = cx$ for $x = 0, 1, 2, 3, 4$, what is the value of c?

- First of all, $c \geq 0$ as $f(x) \geq 0$.
- $$\begin{aligned} f(0) + f(1) + f(2) + f(3) + f(4) &= 1 \\ c(0+1+2+3+4) = 10c &= 1 \end{aligned}$$

  so we must have $c = \frac{1}{10}$.

Cumulative Distribution Functions

Definition

The cumulative distribution function of a random variable X is the function

$$F(x) = P(X \leq x).$$

The cumulative distribution function gives the probability that the variable takes a value less than or equal to x and is defined for all real x .

Cumulative Distribution Functions

Definition

The cumulative distribution function of a random variable X is the function

$$F(x) = P(X \leq x).$$

The cumulative distribution function gives the probability that the variable takes a value less than or equal to x and is defined for all real x .

If f is the probability mass function of a discrete random variable X with range $\{x_1, x_2, \ldots, \}$ and F is its cumulative distribution function, then

$$F(x) = \sum_{x_i \leq x} f(x_i).$$

Cumulative Distribution Functions

The cumulative distribution function of a discrete random variable must satisfy:

1. $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.

Cumulative Distribution Functions

The cumulative distribution function of a discrete random variable must satisfy:

1. $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.
2. $0 \leq F(x) \leq 1$ for all x.

Cumulative Distribution Functions

The cumulative distribution function of a discrete random variable must satisfy:

1. $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.
2. $0 \leq F(x) \leq 1$ for all x.
3. If $x \leq y$ then $F(x) \leq F(y)$.

Cumulative Distribution Functions

The cumulative distribution function of a discrete random variable must satisfy:

1. $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$.
2. $0 \leq F(x) \leq 1$ for all x.
3. If $x \leq y$ then $F(x) \leq F(y)$.

Example

Suppose the range of a discrete random variable is $\{0, 1, 2, 3, 4\}$ and its probability mass function is $f(x) = \frac{x}{10}$. What is its cumulative distribution function?

Example

First of all, for any $x < 1$, $F(x) = \sum_{x_i \le 0} f(x_i) = f(0) = 0$.

Cumulative Distribution Functions

Example

First of all, for any $x < 1$, $F(x) = \sum_{x_i \leq 0} f(x_i) = f(0) = 0$.

Next, for $1 \leq x < 2$, $F(x) = \sum_{x_i \leq 1} f(x_i) = f(0) + f(1) = \frac{1}{10}$

Example

First of all, for any $x < 1$, $F(x) = \sum_{x_i \leq 0} f(x_i) = f(0) = 0$.

Next, for $1 \leq x < 2$, $F(x) = \sum_{x_i \leq 1} f(x_i) = f(0) + f(1) = \frac{1}{10}$

For $2 \leq x < 3$, $F(x) = f(0) + f(1) + f(2) = \frac{3}{10}$
Continuing in the same way, we find that:

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{10} & 1 \leq x < 2 \\ \frac{3}{10} & 2 \leq x < 3 \\ \frac{6}{10} & 3 \leq x < 4 \\ 1 & 4 \leq x. \end{cases}$$

Cumulative Distribution Functions

Example

A discrete random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{10} & 0 \leq x < 1 \\ \frac{3}{10} & 1 \leq x < 2 \\ \frac{5}{10} & 2 \leq x < 4 \\ \frac{8}{10} & 4 \leq x < 5 \\ 1 & 5 \leq x. \end{cases}$$

Determine the probability mass function of X

- The cumulative distribution function only changes value at $0, 1, 2, 4, 5$. So the range of X is $\{0, 1, 2, 4, 5\}$.

Cumulative Distribution Functions

- The cumulative distribution function only changes value at $0, 1, 2, 4, 5$. So the range of X is $\{0, 1, 2, 4, 5\}$.
- $F(0) = \frac{1}{10}$ so $f(0) = \frac{1}{10}$.

Cumulative Distribution Functions

- The cumulative distribution function only changes value at $0, 1, 2, 4, 5$. So the range of X is $\{0, 1, 2, 4, 5\}$.
- $F(0) = \frac{1}{10}$ so $f(0) = \frac{1}{10}$.
- $\frac{3}{10} = F(1) = f(0) + f(1)$ so $f(1) = \frac{2}{10}$.

- The cumulative distribution function only changes value at $0, 1, 2, 4, 5$. So the range of X is $\{0, 1, 2, 4, 5\}$.
- $F(0) = \frac{1}{10}$ so $f(0) = \frac{1}{10}$.
- $\frac{3}{10} = F(1) = f(0) + f(1)$ so $f(1) = \frac{2}{10}$.
- Continuing in the same way we see that the probability mass function is

x	0	1	2	4	5
$f(x)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{2}{10}$