Lecture

Random variables

M.W.

Mathematics Teaching and Distance Learning Centre
Gdańsk University of Technology

2012-2016

Definition (Random variable)

A random variable is a function that assigns a real number to each outcome in the sample space of a random experiment.
A random variable is denoted by an upper-case letter such as X. After an experiment is conducted, the measured value of the random variable is denoted by a lower-case letter such as x.

- A discrete random variable is a random variable with a finite (or countably infinite) range.
- A continuous random variable is a random variable with an interval (either finite or infinite) of real numbers for its range.

Example

- Examples of continuous random variables:
  electrical current, length, pressure, temperature, time, voltage, weight.
- Examples of discrete random variables:
  number of scratches on a surface, proportion of defective parts among 1000 tested, number of transmitted bits received in error.

The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X. For a discrete random variable, the distribution is often specified by just a list of the possible values along with the probability of each. In some cases, it is convenient to express the probability in terms of a formula.

Example

I select at random a student from the class and measure his or her height in centimetres.

Example

I throw a six-sided die twice; I am interested in the sum of the two numbers. Here the sample space is

$$\Omega = \{(i,j);\ 1 \leq i,j \leq 6\}$$

and the random variable F is given by $F(i,j) = i + j$. The target set is the set $\{2, 3, \ldots, 12\}$.

Definition (Mass function)

For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$, a probability mass function is a function such that

1. $f(x_i) \geq 0$
2. $\sum_{i=1}^{n} f(x_i) = 1$
3. $f(x_i) = P(X = x_i)$

Example

I toss a fair coin three times. The random variable X gives the number of heads recorded. The possible values of X are $0, 1, 2, 3$, and its p.m.f. is

a	0	1	2	3
$f(a) = P(X = a)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Exercises

1. Consider the experiment of throwing a fair die. Let X be the r.v. which assigns 1 if the number that appears is even and 0 if the number that appears is odd.
    1. What is the range of X?
    2. Find $P(X = 1)$ and $P(X = 0)$.
2. Consider the experiment of tossing a fair coin three times. Let X be the r.v. that counts the number of heads in each sample point. Find the following probabilities: (a) $P(X \leq 1)$; (b) $P(X > 1)$; and (c) $P(0 < X < 3)$.
3. Consider the experiment of tossing a coin three times. Let X be the r.v. giving the number of heads obtained. We assume that the tosses are independent and the probability of a head is p.
    1. What is the range of X?
    2. Find the probabilities $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, and $P(X = 3)$.

1. Consider the experiment of throwing two fair dice. Let X be the r.v. indicating the sum of the numbers that appear.
    1. What is the range of X?
    2. Find (i) $P(X = 3)$; (ii) $P(X \leq 4)$; and (iii) $P(3 < X \leq 7)$.

Definition (Cumulative distribution function)

The cumulative distribution function of a discrete random variable X, denoted as $F(x)$, is

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

For a discrete random variable X, $F(x)$ satisfies the following properties.

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

Example

Suppose a discrete r.v. X has the following pmfs:

$$p_X(1) = \frac{1}{2}, \quad p_X(2) = \frac{1}{4}, \quad p_X(3) = \frac{1}{8}, \quad p_X(4) = \frac{1}{8}$$

1. Find and sketch the cdf $F_X(x)$ of the r.v. X.
2. Find (i) $P(X \leq 1)$, (ii) $P(1 < X \leq 3)$, (iii) $P(1 \leq X \leq 3)$.

Exercises

1. 1. Verify that the function $p(x)$ defined by

   $$p(x) = \begin{cases} \frac{3}{4}\left(\frac{1}{4}\right)^x & x = 0, 1, 2, \dots \\ 0 & otherwise \end{cases}$$

   is a pmf of a discrete r.v. X.

   2. Find (i) $P(X = 2)$, (ii) $P(X \leq 2)$, (iii) $P(X \geq 1)$.

2. Consider the experiment of tossing an honest coin repeatedly. Let the r.v. X denote the number of tosses required until the first head appears.

   1. Find and sketch the pmf $p_X(x)$ and the cdf $F_X(x)$ of X.
   2. Find (i) $P(1 < X \leq 4)$, (ii) $P(X > 4)$.

Mean and Variance of a Discrete Random Variable

Definition

The mean or expected value of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_x x f(x)$$

The variance of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

The standard deviation of X is $\sigma = \sqrt{\sigma^2}$.

Example

I toss a fair coin three times; X is the number of heads. What are the expected value and variance of X? $E(X) = 0 \cdot (1/8) + 1 \cdot (3/8) + 2 \cdot (3/8) + 3 \cdot (1/8) = 3/2$,
$Var(X) = 0^2 \cdot (1/8) + 1^2 \cdot (3/8) + 2^2 \cdot (3/8) + 3^2 \cdot (1/8) - (3/2)^2 = 3/4$.

$Var(X) = \left(-\frac{3}{2}\right)^2 \cdot \frac{1}{8} + \left(-\frac{1}{2}\right)^2 \cdot \frac{3}{8} + \left(\frac{1}{2}\right)^2 \cdot \frac{1}{8} + \left(\frac{3}{2}\right)^2 \cdot \frac{1}{8} = \frac{3}{4}$

Two properties of expected value and variance can be used as a check on your calculations.

- The expected value of X always lies between the smallest and largest values of X.
- The variance of X is never negative. (For the formula is a sum of terms, each of the form $(a_i - \mu)^2$ (a square, hence non-negative) times $P(X = a_i)$ (a probability, hence non-negative).

Example

Let a r.v. X denote the outcome of throwing a fair die. Find the mean and variance of X.

Exercises

1. Consider a discrete r.v. X whose pmf is given by

$$p_X(x) = \begin{cases} \frac{1}{3} & x = -1, 0, 1 \\ 0 & otherwise \end{cases}$$

   1. Plot $p_X(x)$ and find the mean and variance of X
   2. Repeat (a) if the pmf is given by

$$p_X(x) = \begin{cases} \frac{1}{3} & x = -2, 0, 2 \\ 0 & otherwise \end{cases}$$

2. Let X denote the number of heads obtained in the flipping of a fair coin twice
   1. Find the pmf of X.
   2. Compute the mean and the variance of X.

Theorem

If X is a discrete random variable with probability mass function $f(x)$,

$$E[h(X)] = \sum_{x} h(x) f(x)$$

Definition

A random variable X has a discrete uniform distribution if each of the n values in its range, say, $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$f(x_i) = 1/n$$

Continuous Random Variables and Probability Distribution

Definition

For a continuous random variable X, a probability density function is a function such that

1. $f(x) \geq 0$
2. $\int\limits_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a \leq X \leq b) = \int\limits_{a}^{b} f(x)dx$

Theorem

If X is a continuous random variable, for any x_1 and x_2,

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2)$$

Definition

The cumulative distribution function of a continuous random variable X is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

for $-\infty < x < \infty$.

Theorem

The c.d.f. is an increasing function (this means that $F(x) \leq F(y)$ if $x < y$), and approaches the limits 0 as $x \to -\infty$ and 1 as $x \to \infty$.

Theorem

$$f(x) = \frac{d}{dx}F(x)$$

Example

Consider the experiment of throwing a dart onto a circular plate with unit radius. Let X be the r.v. representing the distance of the point where the dart lands from the origin of the plate. Assume that the dart always lands on the plate and that the dart is equally likely to land anywhere on the plate.

1. What is the range of X ?
2. Find (i) $P(X < a)$ and (ii) $P(a < X < b)$, where $a < b \leq 1$.

Example

Consider the experiment of throwing a dart onto a circular plate with unit radius, again. Sketch the cdf $F_X(x)$ of the r.v. X defined by it. and specify the type of X.

Exercises

1. Consider the function given by

$$F(x) = \begin{cases} 0 & x < 0 \\ x + \frac{1}{2} & 0 \leq x < \frac{1}{2} \\ 1 & x > \frac{1}{2} \end{cases}$$

   1. Sketch $F(x)$ and discuss its properties.
   2. If X is the r.v. whose cdf is given by $F(x)$, find (i) $P(X \leq \frac{1}{4})$ (ii) $P(0 < X \leq \frac{1}{4})$, (iii) $P(X = 0)$ and (iv) $P(0 \leq X \leq \frac{1}{4})$.
   3. Specify the type of X.

1. The pdf of a continuous r.v. X is given by

$$f_X(x) = \begin{cases} \frac{1}{3} & 0 < x < 1 \\ \frac{2}{3} & 1 < x < 2 \\ 0 & otherwise \end{cases}$$

   Find the corresponding cdf $F_X(x)$ and sketch $f_X(x)$ and $F_X(x)$.

2. Let X be a continuous r.v. X with pdf

$$f_X(x) = \begin{cases} kx & 0 < x < 1 \\ 0 & otherwise \end{cases}$$

   1. Determine the value of k and sketch $f_X(x)$.
   2. Find and sketch the corresponding cdf $F_X(x)$.
   3. Find $P(\frac{1}{4} < X \leq 2)$.

Mean and Variance of a Continuous Random Variable

Definition

Suppose X is a continuous random variable with probability density function $f(x)$. The mean or expected value of X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

The variance of X, denoted as $V(X)$ or σ^2, is

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x) - \mu^2 dx$$

The standard deviation of X is $\sigma = \sqrt{\sigma^2}$.

Example

Let X be a continuous r.v. X with pdf

$$f_X(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & otherwise \end{cases}$$

Find the mean and variance of the r.v. X.

Theorem

If X is a continuous random variable with probability density function $f(x)$,

$$E[h(X)] = \int_{-\infty}^{\infty} h(x)f(x)dx$$

Definition

A continuous random variable X with probability density function

$$f(x) = 1/(b-1), \quad a \leq x \leq b$$

is a continuous uniform random variable.

Proposition

$$F(x) = \begin{cases} 0 & \textit{if } x < a \\ (x-a)/(b-a) & \textit{if } a \leq x \leq b \\ 1 & \textit{if } x > b \end{cases}$$

Example

Let X be a uniform r.v. over (a, b). Find the mean and variance of the r.v. X.

Certain distributions

- Bernoulli $p_X(x = 1) = p$, $p_X(x = 0) = 1 - p$ for some $p \in [0, 1]$
- Binomial $p_X(x = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, where n is positive integer, $0 \leq k \leq n$ and $p \in [0, 1]$
- Geometric $p_X(x = k) = (1 - p)p^k$, k is nonnegative integer, $p \in (0, 1)$
- Poisson $p_X(x = k) = e^{-\lambda} \lambda^k / k!$, k is nonnegative integer, $\lambda > 0$
- Uniform $p_X(x = k) = 1/n$, n some positive integer, $1 \leq k \leq n$

- Uniform $f(x) = \mathbb{1}_{[a,b]}\frac{1}{b-a}$
- Exponential $f(x) = \lambda e^{-\lambda x}$, $\lambda > 0$, $x \in [0, \infty)$
- Standard normal $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
- $\mathcal{N}(\mu, \sigma^2)$ $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$