Random variables

Wenda Zhou

May 25, 2017

What is a random variable?

- random numerical quantity
- can characterise the randomness

Die roll

The result of a die roll is a number from 1 to 6.
This can be seen as a random variable.
Randomness: each number is equally likely.

Random variables
Examples

Some examples

- temperature tomorrow
- weight of a baby at birth
- FTSE100 Index

Random variables

Examples

Some examples

- temperature tomorrow
- weight of a baby at birth
- FTSE100 Index

Function of a random variable

If X is a random variable, and g is a function, then $g(X)$ is a random variable.

Random variables and events

Every statement we can make about a random variable is an event.

Examples

Let X be a random variable, e.g. the temperature tomorrow.

- $X < 32$ is an event
- $X > 70$ is an event
- $32 \leq X \leq 70$ is an event

Hence we can (in principle) compute probabilities.

Discrete random variables

Discrete random variables take a finite (or at most countable) number of values.

Examples

- Number of heads in 10 coin tosses
- Number of customers for a store today

They are easy to characterise

Probability mass function

The probability mass function (p.m.f.) offers a full characterisation of a discrete random variable. The p.m.f. f_X of a random variable X is defined by:

$$f_X(x) = \mathrm{P}(X = x). \tag{1}$$

Die roll

Let X be the outcome of a die roll. Then for $x = 1, \ldots, 6$, we have

$$\mathrm{P}(X = x) = 1/6. \tag{2}$$

Hence we have $f_X(x) = 1/6$ for all $x = 1, \ldots, 6$.

Probability mass function

The p.m.f. allows us to compute the probability of any event involving X. Let $X \in A$ be the event, then:

$$\mathrm{P}(X \in A) = \sum_{x \in A} \mathrm{P}(X = x) = \sum_{x \in A} f_X(x). \tag{3}$$

Die roll

Let X be the outcome of a die roll. The event $X \leq 2$ can also be written as $X \in \{1, 2\}$, and hence we may compute

$$\mathrm{P}(X \leq 2) = f_X(1) + f_X(2) = 1/3. \tag{4}$$

Probability mass function

- It is enough to give the p.m.f. of a random variable to fully characterise the random variable.
- We will often define distributions by giving their p.m.f.
- The p.m.f. is a probability, it must be between 0 and 1.
- We have that $\sum_x f_X(x) = 1$ (normalization).

Probability mass function

Let us try to compute the p.m.f. of X, where X is the number of heads in two coin flips.
Four scenarios are possible: HH, HT, TH, TT, all being equally likely. Hence the number of heads:

$$
\begin{aligned}
f_X(2) &= \mathrm{P}(X = 2) = 1/4 \\
f_X(1) &= \mathrm{P}(X = 1) = 1/2 \\
f_X(0) &= \mathrm{P}(X = 0) = 1/4
\end{aligned}
$$

Expectation

The expectation of random variable X (written $\mathbb{E}\,X$) is the "long run average" of X. For a discrete random variable, we define it as:

$$\mathbb{E}\,X = \sum_x x\,\mathrm{P}(X = x) = \sum_x x f_X(x), \tag{5}$$

where the sum goes over all possible values of x.

Expectation

Die roll

Let X be the outcome of a die roll. We have:

$$\mathbb{E}\, X = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} \tag{6}$$

Hence adding it all up, we get:

$$\mathbb{E}\, X = 3.5 \tag{7}$$

Expectation

Two coins

Let X be the number of heads in two coin tosses. We have

$$\mathbb{E}\, X = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} \tag{8}$$

Summing it all up, this gives

$$\mathbb{E}\, X = 1 \tag{9}$$

Expectation of functions

We will often be interested in computing the expectation of a function of X. Let g be a function, then we define

$$\mathbb{E}\, g(X) = \sum_x g(x)\,\mathsf{P}(X = x) = \sum_x g(x) f_X(x) \tag{10}$$

Expectation of functions

Let X be the number of heads in two coin tosses. Let $g(x) = x^2$, we compute $\mathbb{E}\,X$.

$$\mathbb{E}\,g(X) = \mathbb{E}\,X^2 = 0^2 \times \frac{1}{4} + 1^2 \times \frac{1}{2} + 2^2 \times \frac{1}{4} \tag{11}$$

Hence we deduce that

$$\mathbb{E}\,X^2 = \frac{3}{2} \tag{12}$$

Note that $\mathbb{E}\,X^2 \neq (\mathbb{E}\,X)^2$.

Variance

Expectation allows us to define population variance.

$$\sigma^2 = \mathbb{E}(X - \mathbb{E}\,X)^2 \tag{13}$$

Example: two coin tosses

Continuous random variables

- Take values in the reals
- $\mathrm{P}(X = x) = 0$ for all x.
- Characterised by probability density function

Uniform random variable

X, a uniform random variable on $[0, 1]$, is continuous

Cumulative distribution function

The (cumulative) distribution function of X is written F_X and is defined by:

$$F_X(x) = \mathrm{P}(X \leq x) \tag{14}$$

Uniform random variable

If X is uniform on $[0, 1]$, then

$$F_X(x) = \mathrm{P}(X \leq x) = x \tag{15}$$

for $0 \leq x \leq 1$.

Cumulative distribution function

The distribution function allows us to compute event probabilities

Uniform random variable

Let X be uniform on $[0, 1]$, and let's compute $\mathrm{P}(1/4 \leq X \leq 3/4)$.
We should get $\mathrm{P}(1/4 \leq X \leq 3/4) = 1/2$.

Probability density function

For X continuous, we have

$$P(X = x) = 0 \tag{16}$$

Probability density function

For X continuous, we have

$$P(X = x) = 0 \tag{16}$$

Instead, let's try

$$P(x - h/2 \leq X \leq x + h/2) = F_X(x + h/2) - F_X(x - h/2) \tag{17}$$

Now, normalize and take the limit, define the p.d.f. $f_X(x)$:

$$f_X(x) = F'_X(x) \tag{18}$$

Probability density function

Often, we will be given only the density f_X. How can we compute probabilities with the density?

Fundamental theorem of calculus

$$\int_a^b f_X(x)dx = F_X(b) - F_X(a) \tag{19}$$

Probability density function

Often, we will be given only the density f_X. How can we compute probabilities with the density?

Fundamental theorem of calculus

$$\int_a^b f_X(x)dx = F_X(b) - F_X(a) \tag{19}$$

To compute the probability of an event, integrate:

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f_X(x)dx \tag{20}$$

Probability density function

Some properties of the density function.

- The density is non-negative: $f_X(x) \geq 0$
- The density is normalized.

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \tag{21}$$

Expectation (bis)

The p.d.f. allows us to define the expectation of a continuous random variable X.

$$\mathbb{E}\,X = \int x f_X(x) dx \tag{22}$$

Expectation (bis)

The p.d.f. allows us to define the expectation of a continuous random variable X.

$$\mathbb{E}\,X = \int x f_X(x) dx \tag{22}$$

Uniform random variable

If X is uniform on $[0, 1]$, then

$$\mathbb{E}\,X = \frac{1}{2} \tag{23}$$

Expectation (bis)

Similarly, can define expectation of functions of X.

$$\mathbb{E}\, g(X) = \int g(x) f_X(x) dx \tag{24}$$

Expectation (bis)

Similarly, can define expectation of functions of X.

$$\mathbb{E}\, g(X) = \int g(x) f_X(x) dx \tag{24}$$

Uniform random variable

If X is uniform on $[0, 1]$, then

$$\mathbb{E}\, X^2 = \frac{1}{3} \tag{25}$$

Variance (bis)

Define the variance of a continuous random variable X as

$$\mathbb{E}(X - \mathbb{E}\,X)^2 \tag{26}$$

Uniform random variable

If X is uniform $[0, 1]$, then

$$\sigma^2 = \mathbb{E}(X - \mathbb{E}\,X)^2 = \frac{1}{12} \tag{27}$$