Lectures 7. Random variables and random vectors. Joint and marginal distributions. Expectation and variance.

YULIA R. GEL

CS/SE/STAT 3341 Probability and Statistics in Computer Science and Software Engineering

February 2, 2017

A Random Variable

Definition. A **random variable** is a real-valued variable that takes on values by chance, or randomly.

Or more formally, a **random variable** (r.v.) X is a a real-valued function of outcome over the sample space Ω, i.e. $X = f(w)$. So,

$$X : \Omega \to \mathrm{Range},$$

where Range can be $(0, \infty)$ (job execution time), or $(-\infty, \infty)$ (gain/loss in forex trading), or the set of integers Z (number of bugs found in your program over certain period of time) etc.

A Random Variable: Coins

Flip a coin N times. Here the sample space is $\Omega = \{H, T\}^N$. We can then define the random variable $X \in \{0, 1, \ldots, N\}$ to be the number of tails.

Thus

$$X : \{H, T\}^N \to \{0, 1, \ldots, N\}.$$

A Random Variable: Coins

In an experiment involving a sequence of 6 tosses of a coin, the number of heads in the sequence is a random variable.

However, the 6-long sequence of heads and tails is not considered a random variable because it does not have an explicit numerical value.

A Random Variable: Dice and Message Transmission

In an experiment involving two rolls of a die, the following are examples of random variables:

- The max/min number over two rolls,
- The logarithm of the 5th roll,
- The number of 1s in the two rolls.

In an experiment involving the transmission of a message, the following are random variables:

- the time needed to transmit the message,
- the number of symbols received in error,
- the delay with which the message is received.

Distribution

Definitions:

- **Distribution** of r.v. X is a collection of all the probabilities related to X.
- The **probability mass function** (PMF) of r.v. X is
$$P(x) = \mathbf{P}\{X = x\}.$$
- The **cumulative distribution function** (CDF) of r.v. X is
$$F(x) = \mathbf{P}\{X \leq x\} = \sum_{y \leq x} \mathbf{P}(y).$$
Properties $F(-\infty) = 0$, $F(+\infty) = 1$,
$\mathbf{P}\{a < X \leq b\} = F(b) - F(a)$, and $F(x)$ is non-decreasing.
- The set of all possible values of X is called the **support** of the distribution F.

Remark.

Conventionally, capital letters denote random variables, while lower case letters denote particular values (or realizations).

PMF Calculation: Algorithm

Thus, the algorithm for calculation of the PMF of a Random Variable X can be summarized as follows.

For each possible value x of X:

1. Collect all the possible outcomes that give rise to the event $\{X = x\}$.
2. Add their probabilities to obtain $P(x)$.

Examples: Coins

For example, let the experiment consist of two independent tosses of a fair coin, and let X be the number of heads obtained.

Then the PMF of X is:

$$P(x) = \begin{cases} 0.25, \text{ if } x = 0 \text{ or } x = 2 \\ 0.5, \quad \text{if } x = 1 \\ 0, \text{ otherwise} \end{cases}$$

PMF and CDF Calculation: Rolling a Single Die

Let's roll a six-sided die. A die can land on any of its 6 faces $\{1, 2, 3, 4, 5, 6\}$, so that a single experiment has 6 possible outcomes.

For a "fair die", we expect to get each of the 6 results with an equal probability. I.e., if we repeat the same experiment many times, we would expect that, on average, the 6 possible events would occur with similar frequencies. (Later we will say that such events are uniformly distributed.)

The PMF is

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}.$$

PMF and CDF Calculation: Rolling a Single Die

Recall that CDF of a r. v. X is $F(x) = \mathbf{P}\{X \leq x\}$.

Hence, in the fair die experiment, we get

$$F(1) = \mathbf{P}\{X \leq 1\} = 1/6$$

$$F(2) = \mathbf{P}\{X \leq 2\} = 2/6$$

$$F(4) = \mathbf{P}\{X \leq 4\} = 4/6$$

$$F(6) = \mathbf{P}\{X \leq 6\} = 1$$

Discrete Random Variables

A r. v. is called **discrete** if its range (i.e. the set of values that it can take) is finite or at most countably infinite. E.g., tossing a coin, rolling a dice, getting a grade in STAT3341 can take at most a finite number of numerical values, and are therefore discrete.

Note that discrete r.v. do not need to be integer! E.g., proportion of defective components in a lot of 10 motherboards can be $1/10, 2/10, \ldots, 1$, i.e. countable but not integer outcomes.

Continuous Random Variables

A r. v. is called **continuous** if it assumes the whole interval of values, e.g. $(-\infty, \infty)$, (a, b), $[0, \infty]$ etc.

Continuous r. v. are usually measurements. Examples include height, weight, amount of sodium in a meal, program execution time etc.

"Mixed" Random Variables

Note that r.v. are neither discrete nor continuous.

Example. A packet arrives at a router in a communication network. If the input buffer is empty (happens with probability p), the packet is serviced immediately. Otherwise the packet must wait for a random amount of time.

Define the r.v. X to be the packet service time. Then X is neither discrete nor continuous.

For the next few lectures, we will focus solely on discrete random variables.

Motivation: A wheel of Fortune

Suppose you spin a wheel of fortune many times. At each spin, one of the numbers $m_1, m_2, \ldots, m_n$ comes up with corresponding probability $p_1, p_2, \ldots, p_n$, and this is your monetary reward from that spin.

What is the amount of money that you "expect" to get "per spin"?

The terms "expect" and "per spin" are a little ambiguous, but here is a reasonable interpretation.

Motivation: A wheel of Fortune

Suppose that you spin the wheel k times, and that k_i is the number of times that the outcome is m_i. Then, the total amount received is $m_1k_1 + m_2k_2 + \ldots + m_nk_n$. The amount received per spin is

$$M = \frac{m_1k_1 + m_2k_2 + \ldots + m_nk_n}{k}.$$

If the number of spins k is very large, and if we are willing to

interpret probabilities as relative frequencies, it is reasonable to anticipate that m_i comes up a fraction of times that is roughly equal to p_i:

$$p_i \approx \frac{k_i}{k}, \quad i = 1, \ldots, n.$$

Thus, the amount of money per spin that you expect to receive is

$$M = \frac{m_1k_1 + m_2k_2 + \ldots + m_nk_n}{k} \approx m_1p_1 + m_2p_2 + \ldots + m_np_n.$$

Expectation: Definition

Motivated by this example, we now introduce the notion of expectation.

Definition. We define the **expected value** (also called the **expectation** or the **mean**, or **averaged value**) of a random variable X, with PMF $P(x)$, by

$$E(X) = \sum_{x} xP(x).$$

The expected value is often denoted by μ.

Examples: Coins

Consider two independent coin tosses, each with a 0.75 probability of a head, and let X be the number of heads obtained. Its PMF is

Then the PMF of X is:

$$P(x) = \begin{cases} 0.25 \times 0.25, \text{ if } x = 0 \\ 0.25 \times 0.75 + 0.75 \times 0.25, & \text{if } x = 1 \\ 0.75 \times 0.75, & \text{if } x = 2 \end{cases}$$

Let us calculate its expected value:

$$\begin{aligned} \mu_X &= E(X) \\ &= 0 \times (0.25 \times 0.25) + 1 \times (0.25 \times 0.75 + 0.75 \times 0.25) \\ &+ 2 \times (0.75 \times 0.75) \\ &= \frac{3}{2}. \end{aligned}$$