Lecture 3

Outline

1. Independence of many events
   - Bernoulli scheme
2. Random variables
   - Probability mass function
   - Examples

Independence of many events

We can extend the definition of independence to multiple events (more than two events).

Definition (Independence of n events, $n \geq 2$)

Let $(\Omega, \mathbb{P})$ be a probability space. Events $A_1, \ldots, A_n$ are **independent**, if and only if

$$\mathbb{P}\left(\bigcap_{i \in K} A_i\right) = \prod_{i \in K} \mathbb{P}(A_i),$$

for any subset $K \subset \{1, 2, \ldots, n\}$.

What does it mean for three events A, B and C?

The following four conditions are to be satisfied:

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(A)\mathbb{P}(B), \\ \mathbb{P}(A \cap C) &= \mathbb{P}(A)\mathbb{P}(C), \\ \mathbb{P}(B \cap C) &= \mathbb{P}(B)\mathbb{P}(C), \\ \mathbb{P}(A \cap B \cap C) &= \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C). \end{aligned}$$

- The first three conditions imply that any two events are independent. This property is known as pairwise independence.
- The fourth condition does not follow from the first three.

Example (Pairwise independence $\not\Longrightarrow$ independence)

Consider two tosses of a fair coin. Let A be the event that the first comes up heads, B the event that the second comes up tails and C the event that both flips have the same result. Are A, B and C independent?

Solution

Let $A = \{$*the first comes up heads*$\}$, $B = \{$*the second comes up tails*$\}$, $C = \{$*both flips have the same result*$\}$.

It is easy to verify that:
$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$ *and* $\mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = \frac{1}{4}$
$\Longrightarrow$ *any two events (out of the set* $\{A, B, C\}$*) are independent*

On the other hand:

$$\mathbb{P}(A \cap B \cap C) = 0 \text{ and } \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{8}.$$

$\Longrightarrow$ *A, B and C are not independent.*

The following example shows that the fourth condition:

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$$

is not enough for independence.

Example

Let $(\Omega, \mathbb{P})$ be a probability space, $\Omega = [0,1]^2$ and $\mathbb{P}$ is a geometric probability on Ω. Let

$$A = \left\{(x,y) : x \leq \frac{1}{2}\right\},$$

$$B = \left\{(x,y) : y \leq \frac{1}{2}\right\},$$

$$C = \{(x,y) : y \leq x\}.$$

Are A, B and C independent?

Solution

Using geometrical probability model one can check that:

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2} \text{ and } \mathbb{P}(A \cap B \cap C) = \frac{1}{8}.$$

On the other hand

$$A \cap C = \left\{(x, y) : y \leq x \leq \frac{1}{2}\right\} \implies \mathbb{P}(A \cap C) = |A \cap C| = \frac{1}{8},$$

so

$$\mathbb{P}(A \cap C) \neq \mathbb{P}(A)\mathbb{P}(C).$$

Bernoulli scheme

Consider an experiment that involves a sequence of trials satisfying the following conditions (**Bernoulli trials**):

- each trial results in one of two possible outcomes: success(S) and failure (F),
- the probability of success (p) is the same at each trial,
- the trials are independent.

Such sequence is called the **Bernoulli scheme**.

Example

A biased coin (the probability of "heads" is p) is tossed n times (each toss is independent of the others).

- Probability of any particular sequence $(a_1, \ldots, a_n)$, $a_i \in \{\text{heads}, \text{tails}\}$, $i = 1, 2, \ldots, n$, containing k heads and $n - k$ tails is

$$p^k (1 - p)^{n-k}.$$

- $\mathbb{P}(\text{k heads come up in } n \text{ trials}) = \binom{n}{k} p^k (1 - p)^{n-k}$.
- $\mathbb{P}(\text{the first success (head) comes out in kth trial}) = (1 - p)^{k-1} p$.

Example

An internet service provider has installed c modems to serve needs of a population of n customers, It is estimated that at a given time, each customer will need a connection with probability p, independently of the others. What is the probability that there are more customers needing a connection than there are modems?

Solution

Let $A = \{\textit{more than c customers need a connection at the same time}\}$ *and* $B_i = \{\text{"}i\text{" customers need a connection at the same time}\}$, $i = 1, \ldots, n$. *Thus* $A = \bigcup_{i=c+1}^{n} B_i$, *and using the Bernoulli scheme*

$$\mathbb{P}(B_i) = \binom{n}{i} p^i (1-p)^{n-i}.$$

Therefore

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=c+1}^{n} B_i\right) = \sum_{i=c+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}.$$

Random variables

A function that assignes a numerical value to the outcome of an experiment is called a **random variable**.

Definition (Random variable)

Let $(\Omega, \mathbb{P})$ be a probability space. A **random variable** X is a real-valued function of the experimental outcome:

$$X : \Omega \longrightarrow \mathbb{R}.$$

X assignes a value $X(\omega)$ to each outcome $\omega \in \Omega$.

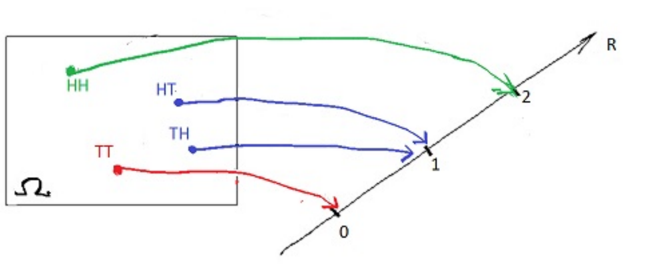
Example (Flipping a coin)

Flip a coin twice. The sample space Ω consists of four possible outcomes:

$$\Omega = \{HH, HT, TH, TT\}.$$

Let us define a random variable X on this space as a number of heads:

$$\begin{aligned} X(\{HH\}) &= 2, \\ X(\{HT\}) &= 1, \\ X(\{TH\}) &= 1, \\ X(\{TT\}) &= 0. \end{aligned}$$



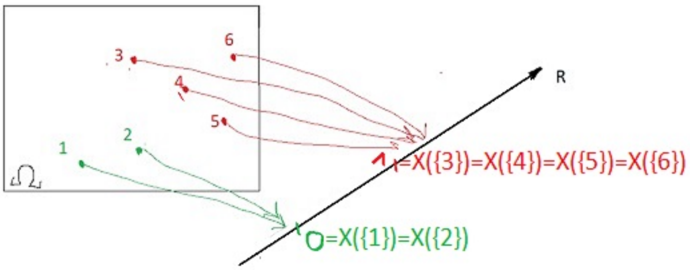
Example (Rolling a die)

Roll a die once. The sample space Ω consists of six possible outcomes:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

and let us define a random variable $X : \Omega \to \mathbb{R}$ in the following way:

$X(\{1\}) = X(\{2\}) = 0$ and $X(\{3\}) = X(\{4\}) = X(\{5\}) = X(\{6\}) = 1.$



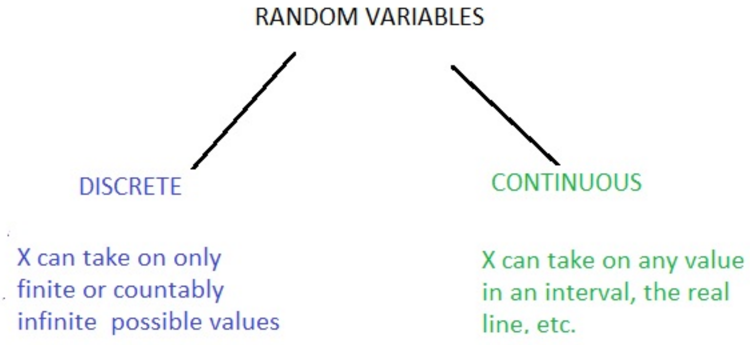
Before the experiment is performed, we don't know the exact value of X but we can ask some natural questions like:

- compute the probability that the random variable will take on a given value,
- calculate the probability that the random variable will fall into a given range.

To answer these questions we need to know the distribution of the random variable X. It provides us with probabilities of the events like, for example:

- $X = 5$,
- $X \in (0, 3/2)$,
- $X > 10$, etc.

We distinguish to types of random variables: discrete and continuous.



In this lecture, we will focus only on the discrete type (the continuous one will be discussed in the next lecture).

How can we decribe the distribution of a discrete random variable?

Example (Flipping a coin twice)

X - the number of heads, $\Omega = \{HH, HT, TH, TT\}$,
$S_X = \{0, 1, 2\}$ - possible values that X can take on,

- $\mathbb{P}(X = 2) := \mathbb{P}(\{\omega \in \Omega : X(\omega) = 2\}) = \mathbb{P}(\{HH\}) = \frac{1}{4}$,
- $\mathbb{P}(X = 1) := \mathbb{P}(\{\omega \in \Omega : X(\omega) = 1\}) = \mathbb{P}(\{HT\} \cup \{TH\}) = \frac{1}{2}$,
- $\mathbb{P}(X = 0) := \mathbb{P}(\{\omega \in \Omega : X(\omega) = 0\}) = \mathbb{P}(\{TT\}) = \frac{1}{4}$.

Example (Rolling a dice once)

$\Omega = \{1, 2, 3, 4, 5, 6\}$,
define X as follows:

- $X(\{1\}) = X(\{2\}) = 0$,
- $X(\{3\}) = X(\{4\}) = X(\{5\}) = X(\{6\}) = 1$,

$S_X = \{0, 1\}$ and

- $\mathbb{P}(X = 0) := \mathbb{P}(\{\omega \in \Omega : X(\omega) = 0\}) = \mathbb{P}(\{1\} \cup \{2\}) = \frac{2}{6}$,
- $\mathbb{P}(X = 1) := \mathbb{P}(\{\omega \in \Omega : X(\omega) = 1\}) = \mathbb{P}(\{3\} \cup \{4\} \cup \{5\} \cup \{6\}) = \frac{4}{6}$.

In the above examples, we assigned probabilities to each value that the random variable can take on. In this way we described the distributions of the random variables.

We will know summerize our considerations in a more formal and general way.

Discrete random variables

Discrete random variables take values in a **countable** set $S_X = \{x_1, x_2, \ldots\}$ (the **support** of X).

Definition

A random variable X is said to be **discrete** if there exists a countable set $S_X \subset \mathbb{R}$ (called **support**) such that

- $\mathbb{P}(X = x) > 0$, $\forall x \in S_X$,
- $\sum_{x \in S_X} \mathbb{P}(X = x) = 1$.

Probability mass function

To specify distribution of a discrete random variable we use **the probability mass function (PMF)**:

Definition

If X is a discrete random variable, the function $p_X : \mathbb{R} \to [0, 1]$ defined by

$$p_X(x) = \mathbb{P}(X = x),$$

for every $x \in \mathbb{R}$, is called **the probability mass function** of X.

Example

1 Roll a four-sided dice twice. Let X be the sum of the faces (of the first and the second roll). Find the probability mass function of X.
Solution: What are the possible values that X can take on? We need to determine the support S_X.
The sample space for this experiment: $\Omega = \{(i, j) : i, j \in \{1, 2, 3, 4\}\}$. X is defined as the sum of the faces, so

$$X(i, j) = i + j.$$

Therefore, the possible values that X can take on are given by $S_X = \{2, 3, 4, 5, 6, 7, 8\}$. Now, we kneed to assign probabilities to each element of the support S_X:

$$\begin{aligned} p_X(2) &= \mathbb{P}(X = 2) = \mathbb{P}(\{(i, j) : X(i + j) = 2\}) = \mathbb{P}((1, 1)) = \frac{1}{16}, \\ p_X(3) &= \mathbb{P}(X = 3) = \mathbb{P}(\{(i, j) : X(i + j) = 3\}) = \mathbb{P}((1, 2) \cup \{(2, 1)\}) = \frac{2}{16}, \\ p_X(4) &= \mathbb{P}(\{(1, 3)\} \cup \{(3, 1)\} \cup \{(2, 2)\}) = \frac{3}{16}, \end{aligned}$$

In the similar we can compute: $p_X(5) = \frac{4}{16}$, $p_X(6) = \frac{3}{16}$, $p_X(7) = \frac{2}{16}$, $p_X(8) = \frac{1}{16}$.

Example

2 Flip a coin until the fist H appears. Let X be the number of tosses needed. Find the probability mass function of X.
Solution: $S_X = \{1, 2, 3, \ldots\}$, we need to determine probability mass function of X (PMF):

$$p_X(1) = \mathbb{P}(X = 1) = \frac{1}{2}$$

$$p_X(k) = \mathbb{P}(X = k) = \left(\frac{1}{2}\right)^k, \quad k \geq 1.$$

Example

3 The first urn contains 2 white and 3 black balls and the second one contains 3 white and 1 black balls. Pick one urn at random and draw two balls out of it. Let X be the number of black balls drawn. Find the probability mass function of X.

Solution: $S_X = \{0, 1, 2\}$ - number of possible black balls drawn from a random urn.

- H_1 - the first urn was chosen, H_2 - the second urn was chosen and $\mathbb{P}(H_1) = \mathbb{P}(H_2) = \frac{1}{2}$.
- To find the PMF of X, we need to use the Total Probability Rule:

$$p_X(0) = \mathbb{P}(X = 0) = \mathbb{P}(X = 0|H_1)\mathbb{P}(H_1) + \mathbb{P}(X = 0|H_2)\mathbb{P}(H_2),$$

where $\mathbb{P}(X = 0|H_1) = \frac{\binom{3}{0}\binom{2}{2}}{\binom{5}{2}}$ and $\mathbb{P}(X = 0|H_2) = \frac{\binom{1}{0}\binom{3}{2}}{\binom{4}{2}}$,

$$p_X(1) = \mathbb{P}(X = 1) = \mathbb{P}(X = 1|H_1)\mathbb{P}(H_1) + \mathbb{P}(X = 1|H_2)\mathbb{P}(H_2),$$

where $\mathbb{P}(X = 1|H_1) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}}$ and $\mathbb{P}(X = 1|H_2) = \frac{\binom{1}{1}\binom{3}{1}}{\binom{4}{2}}$, and

$$p_X(2) = \mathbb{P}(X = 2) = \mathbb{P}(X = 2|H_1)\mathbb{P}(H_1) + \mathbb{P}(X = 2|H_2)\mathbb{P}(H_2),$$

where $\mathbb{P}(X = 2|H_1) = \frac{\binom{3}{2}\binom{2}{0}}{\binom{5}{2}}$ and $\mathbb{P}(X = 2|H_2) = 0$.