inline=symbol

Probability and Random Processes

Keivan Mallahi-Karai

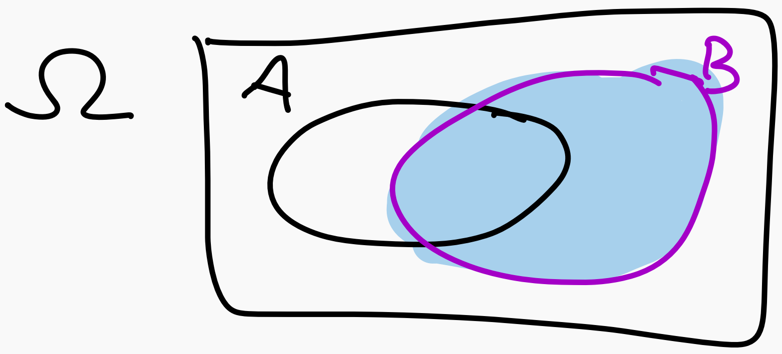
21.09.2022

Jacobs University

Definition of conditional probability

Definition
Suppose that A, B are two events and that $\mathbb{P}(B) \neq 0$. The conditional probability $\mathbb{P}(A|B)$ (read as A given B) is defined by

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$



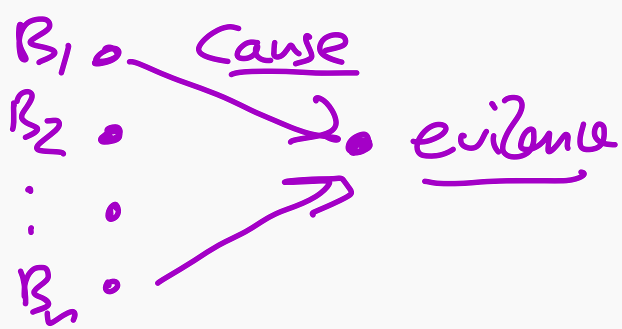
Definition of conditional probability

Definition
Suppose that A, B are two events and that $\mathbb{P}(B) \neq 0$. The conditional probability $\mathbb{P}(A|B)$ (read as A given B) is defined by

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

Remark:

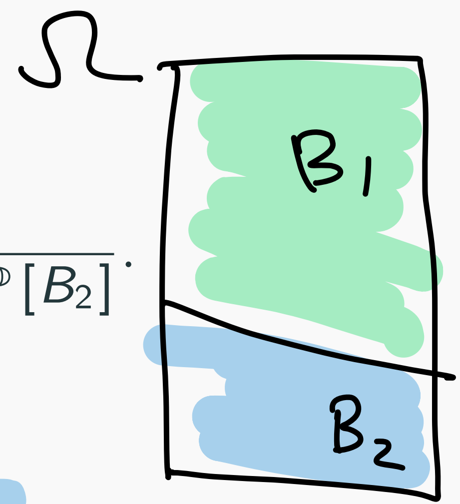
Note that for this definition to make sense, one needs to assume that $\mathbb{P}[B] \neq 0$.



Suppose $\Omega = B_1 \cup B_2$ is a partition of Ω.

$B_1 \cup B_2 = \Omega$

$B_1 \cap B_2 = \emptyset$



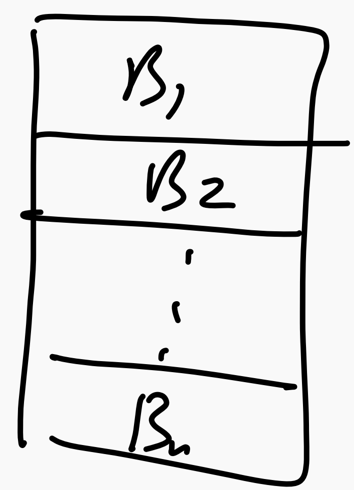
$$\mathbb{P}[B_1 | A] = \frac{\mathbb{P}[B_1 \cap A]}{\mathbb{P}[A]} = \frac{\mathbb{P}[A|B_1]\,\mathbb{P}[B_1]}{\mathbb{P}[A|B_1]\,\mathbb{P}[B_1] + \mathbb{P}[A|B_2]\,\mathbb{P}[B_2]}.$$

$$\mathbb{P}[B_1 | A] = \frac{\mathbb{P}(B_1 \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_1)\,\mathbb{P}(B_1)}{\mathbb{P}(A|B_1)\,\mathbb{P}(B_1) + \mathbb{P}(A|B_2)\,\mathbb{P}(B_2)}$$

$$\mathbb{P}(A|B_1) = \frac{\mathbb{P}(A \cap B_1)}{\mathbb{P}(B_1)}$$

$\Omega = B_1 \cup \dots \cup B_n$

$B_i \cap B_j = \emptyset$



Bayes' formula

Suppose $\Omega = B_1 \cup B_2$ is a partition of Ω.

$$\mathbb{P}\left[B_1 | A\right] = \frac{\mathbb{P}\left[B_1 \cap A\right]}{\mathbb{P}\left[A\right]} = \frac{\mathbb{P}\left[A | B_1\right] \mathbb{P}\left[B_1\right]}{\mathbb{P}\left[A | B_1\right] \mathbb{P}\left[B_1\right] + \mathbb{P}\left[A | B_2\right] \mathbb{P}\left[B_2\right]}.$$

More generally

Theorem (Bayes' Formula)
Let $\Omega = B_1 \cup B_2 \cup \cdots \cup B_n$ be a partitioning of the sample space Ω. Then we have

$$\mathbb{P}\left[B_i | A\right] = \frac{\mathbb{P}\left[A | B_i\right] \mathbb{P}\left[B_i\right]}{\sum_{j=1}^{n} \mathbb{P}\left[A | B_j\right] \mathbb{P}\left[B_j\right]}.$$

Example

Through a transmission channel two types of messages can be sent: 0 and 1. We assume that 40% of the time a 1 is transmitted. The probability that 0 is correctly received is 0.80 and the probability that a transmitted 1 is correctly received is 0.90. Determine

a) the probability of a 0 being received.

b) given a 1 received, the probability that 1 was transmitted.

I_1 = input is 1 | O_1 = output is 1

I_0 = input is 0 | O_0 = output is 0

$\mathbb{P}(I_1) = \frac{4}{10}$, $\mathbb{P}(I_0) = \frac{6}{10}$,

$\mathbb{P}(O_1 | I_1) = \frac{9}{10}$, $\mathbb{P}(O_0 | I_0) = \frac{8}{10}$,

$I_0 \cup I_1 = \Omega$

$$\mathbb{P}(O_0) = \mathbb{P}(O_0 | I_0)\, \mathbb{P}(I_0) + \underbrace{\mathbb{P}(O_0 | I_1)}_{= 1 - \mathbb{P}(O_1 | I_1)}\, \mathbb{P}(I_1)$$

$$= \frac{8}{10} \cdot \frac{6}{10} + \frac{1}{10} \cdot \frac{4}{10} = \frac{48}{100} + \frac{4}{100} = \frac{52}{100}$$

$$\mathbb{P}(I_1 \mid O_1) = \frac{\mathbb{P}(O_1 \mid I_1)\ \mathbb{P}(I_1)}{\mathbb{P}(O_1 \mid I_1)\,\mathbb{P}(I_1) + \mathbb{P}(O_1 \mid I_0)\,\mathbb{P}(I_0)}$$

$$= \frac{\frac{9}{10} \cdot \frac{4}{10}}{\frac{9}{10} \cdot \frac{4}{10} + \frac{2}{10} \cdot \frac{6}{10}}$$

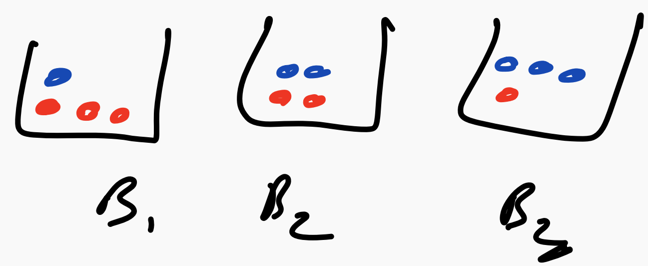
$$= \frac{\frac{36}{100}}{\frac{36}{100} + \frac{12}{100}} = \frac{36}{48} = \frac{9}{12} = \frac{3}{4} = 75\%$$

Example

Consider three buckets B_1, B_2, B_3 such that B_1 contains 3 red and 1 blue ball, B_2 contains 2 balls of each color, and B_3 contains 3 blue and 1 red ball. A random bucket is chosen and a random ball is picked out of the bucket.

1. What is the probability that the chosen ball is blue?
2. Given that the chosen ball is blue, what is the probability that it is picked from each one of the buckets?

$$
\begin{aligned}
\mathbb{P}(\text{blue}) = \; & \mathbb{P}(\text{blue} \mid B_1)\, \mathbb{P}(B_1) \\
& + \mathbb{P}(\text{blue} \mid B_2) \cdot \mathbb{P}(B_2) \\
& + \mathbb{P}(\text{blue} \mid B_3)\, \mathbb{P}(B_2)
\end{aligned}
$$



$$
= \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{12} + \frac{2}{12} + \frac{3}{12} = \frac{6}{12} = \frac{1}{2}.
$$

$$
\mathbb{P}(B_1 \mid \text{blue}) = \frac{\mathbb{P}(\text{blue} \mid B_1)\, \mathbb{P}(B_1)}{\mathbb{P}(\text{blue} \mid B_1)\, \mathbb{P}(B_1) + \mathbb{P}(\text{blue} \mid B_2)\, \mathbb{P}(B_2) + \mathbb{P}(\text{blue} \mid B_2)\, \mathbb{P}(B_3)}
$$

$$= \frac{\frac{1}{4} \cdot \frac{1}{3}}{\frac{1}{4} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} + \frac{3}{4} \cdot \frac{1}{3}} = \frac{\frac{1}{12}}{\frac{1}{2}} = \frac{1}{6}$$

$$\mathbb{P}(B_2 \mid \text{blue}) = \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

$$\mathbb{P}(B_3 \mid \text{blue}) = \frac{\frac{3}{4} \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}.$$

False positives

Suppose that a particular test for Corona is 90% sensitive, that is to 90% of people with Corona will test positive. Assume, further, that the test is also 80% specific, that is, it has 80% true negatives. Assume that currently 5 percent of the population is infected with Corona. Find the probability that a random person who tests positive is indeed infected.

I = infected   + positive

N not infected   − negative

$$\mathbb{P}(+|I) = \frac{9}{10}, \quad \mathbb{P}(-|I) = \frac{1}{10} \qquad \mathbb{P}(I) = \frac{5}{100}.$$

$$\mathbb{P}(-|N) = \frac{8}{10}, \quad \mathbb{P}(+|N) = \frac{2}{10}$$

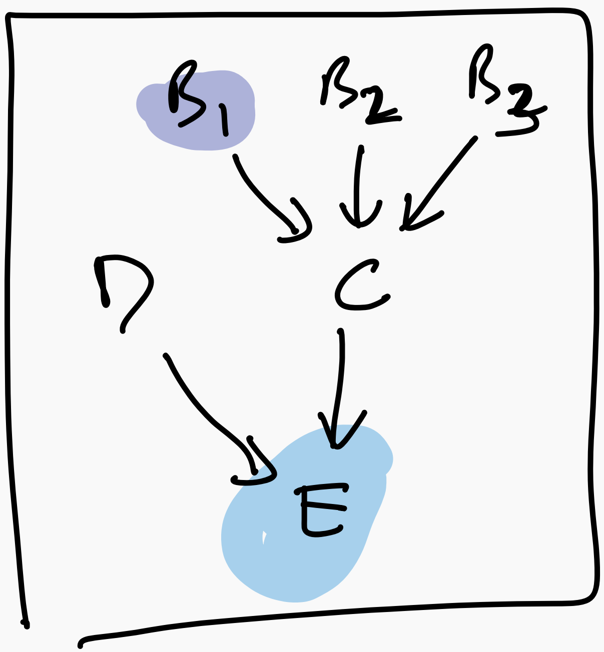
$$\mathbb{P}(I|+) = \frac{\mathbb{P}(+|I)\,\mathbb{P}(I)}{\mathbb{P}(+|I)\,\mathbb{P}(I) + \mathbb{P}(+|N)\cdot\mathbb{P}(N)}$$

$$= \frac{\frac{9}{10}\cdot\frac{5}{100}}{\frac{9}{10}\cdot\frac{5}{100} + \frac{2}{10}\cdot\frac{95}{100}}$$

$$= \frac{\frac{45}{1000}}{\frac{45}{1000} + \frac{190}{1000}} = \frac{45}{45+190} = \frac{45}{235} \approx 0.19$$

Bayes's Theorem

Causality Judea Pearl



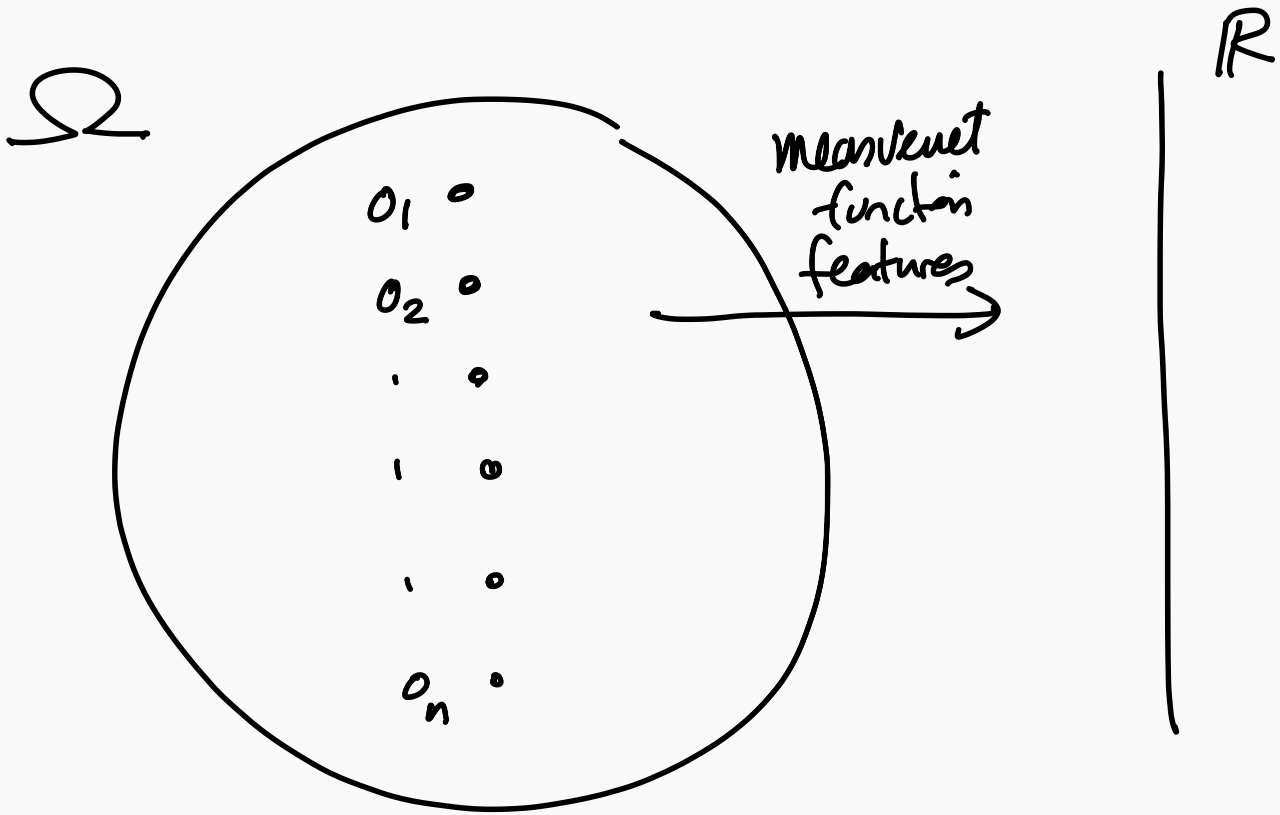
DAG

directed arrow diagram

Statistic

- Bayesian Inference

Random variables: definition



A random variable

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

Random variables: definition

Definition
Let $(\Omega, \mathbb{P})$ be a probability space. A function

$$X : \Omega \to \mathbb{R}$$

is called a real valued *random variable*. Similarly, a function $X : \Omega \to \mathbb{R}^n$ is called a vector-valued random variable.

Random variables: definition

Definition
Let $(\Omega, \mathbb{P})$ be a probability space. A function

$$X : \Omega \to \mathbb{R}$$

is called a real valued *random variable*. Similarly, a function $X : \Omega \to \mathbb{R}^n$ is called a vector-valued random variable.

$p = 1/2$

Example

Experiment: n independent flipping of a coin can result in heads with probability p and in tails with probability $1-p$,

$n=3$

$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\} \xrightarrow{X} \mathbb{R}$

X = # of heads

$HHH \to 3$, $HHT \to 2$, $HTH \to 2$, $HTT \to 1$, $THH \to 2$, THT, TTH, $TTT \to 0$

Possible values	0	1	2	3
Prob.	1/8	3/8	3/8	1/8

$HHH \to 3$
$HHT \to 2$
$HTH \to 1$
$HTT \to 1$
$THH \to 2$
$THT \to 1$
$TTH \to 1$
$TTT \to 0$

longest streak of heads

0	1	2	3
1/8	4/8	2/8	1/8

first time that H shows up

$HHH \to 1$
$HHT \to 1$
$HTH \to 1$
HTT
$THH \to 2$
THT
$TTH \to 3$
$TTT \to 0$

Examples

first line that
H shows up

	0	1	2	3
	1/8	4/8	2/8	1/8

Example

Experiment: n independent flipping of a coin can result in heads with probability p and in tails with probability $1 - p$,

Sample space:

$$\Omega = \{(x_1, x_2, \ldots, x_n) \mid x_i = 0 \text{ or } 1\}.$$

Example
Experiment: n independent flipping of a coin can result in heads with probability p and in tails with probability $1 - p$,

Sample space:

$$\Omega = \{(x_1, x_2, \ldots, x_n) \mid x_i = 0 \text{ or } 1\}.$$

Each outcome of this experiment corresponds to an element $\omega \in \Omega$. Let

$$X_1(\omega) = \{\text{first head}\},$$

$$X_2(\omega) = \{\text{first tail}\},$$

$$X_3(\omega) = \{\text{total number of heads}\},$$

$n=3$

Example
Experiment: n independent throws of a faire die

Ω has $6^3 = 216$ possible outcomes.

X_1 sum of outcomes

$125 \longrightarrow 8$ 0

$613 \longrightarrow 10$ 0

$331 \longrightarrow 7$ 0

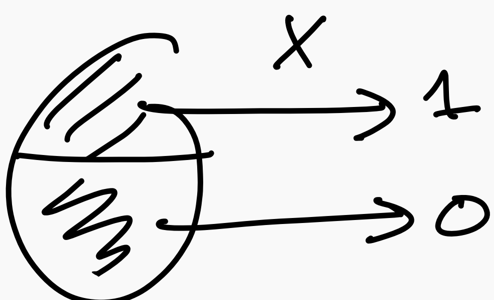
X_2 2 ones in row

X_3 product of outcomes

$\underline{11}2 \longrightarrow 1$

$111 \longrightarrow 2$

get of digits



$125 \rightsquigarrow 0$

$521 \rightsquigarrow 1$

Bernoulli

Examples

Example
Experiment: n independent throws of a faire die

Sample space:

$$\Omega = \{(x_1, x_2, \ldots, x_n) : 1 \leq x_i \leq 6\}.$$

Example
Experiment: n independent throws of a faire die

Sample space:

$$\Omega = \{(x_1, x_2, \ldots, x_n) : 1 \leq x_i \leq 6\}.$$

Each outcome of this experiment corresponds to an element $\omega \in \Omega$. Let

Example
Experiment: n independent throws of a faire die

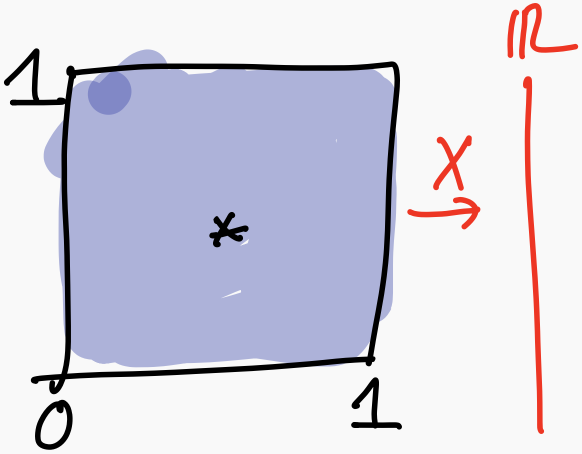
Sample space:

$$\Omega = \{(x_1, x_2, \ldots, x_n) : 1 \leq x_i \leq 6\}.$$

Each outcome of this experiment corresponds to an element $\omega \in \Omega$. Let

$$X_1 = \{\text{sum of the scores}\},$$
$$X_2 = \{\text{the smallest score}\},$$
$$X_3 = \{\text{the second largest score}\},$$

Discrete random variables



$$\Omega = \left\{ (x,y) : \begin{array}{l} 0 \le x \le 1 \\ 0 \le y \le 1 \end{array} \right\}$$

sample space

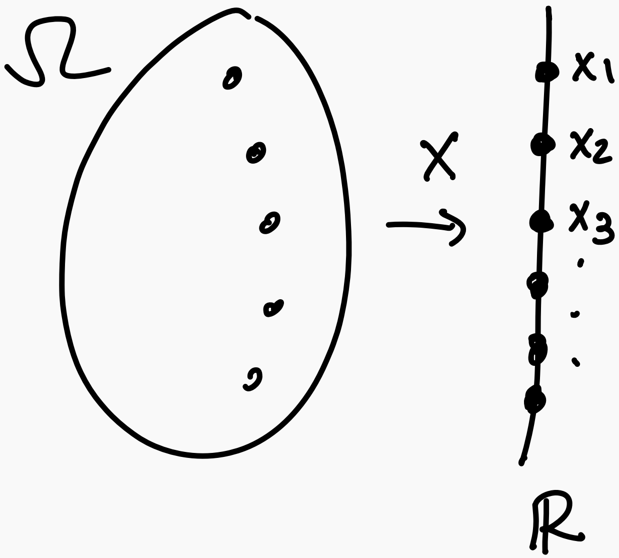
X

$|x-y|$

Definition
The probability mass function of a random variable X with sets of values $x_1, \dots$ is defined by

$$p(x) = \mathbb{P}[X = x].$$

Note that for every i, we have $p(x_i) = p_i > 0$.



Possible values $x_1, x_2, x_3, \dots$

Care about

$$p_i = \mathbb{P}[X = x_i]$$

Bernoulli random variables

Values	0	1
	$1-p$	p

$$p(x) = \mathbb{P}(X = x)$$

$$p(0) = \mathbb{P}(X=0) = 1-p$$

$$p(1) = \mathbb{P}(X=1) = p$$

Ω = 2 throws of a die

$$X = \begin{cases} 1 & \text{if scores are decending} \\ 0 & \text{otherwise} \end{cases}$$

65, 64, 63, 62, 61
54, 53, 52, 51
43, 42, 41
32, 31
21

$$\mathbb{P}(X=1) = \frac{15}{36}$$

$$p = \frac{15}{36}.$$

Bernoulli random variables

The simplest discrete random variables are Bernoulli random variables.

Bernoulli random variables

The simplest discrete random variables are Bernoulli random variables.

Definition
A random variable X is called the *Bernoulli* random variable with parameter p if it only takes values 0 and 1, and

$$\mathbb{P}[X = 1] = p, \qquad \mathbb{P}[X = 0] = 1 - p.$$

Bernoulli random variables

The simplest discrete random variables are Bernoulli random variables.

Definition
A random variable X is called the *Bernoulli* random variable with parameter p if it only takes values 0 and 1, and

$$\mathbb{P}[X = 1] = p, \qquad \mathbb{P}[X = 0] = 1 - p.$$

A Bernoulli random variable X tells us whether something happened or not. The probability of happening $\mathbb{P}[X = 1]$ is called the parameter of X.

Bernoulli random variables

The simplest discrete random variables are Bernoulli random variables.

Definition
A random variable X is called the *Bernoulli* random variable with parameter p if it only takes values 0 and 1, and

$$\mathbb{P}[X = 1] = p, \qquad \mathbb{P}[X = 0] = 1 - p.$$

A Bernoulli random variable X tells us whether something happened or not. The probability of happening $\mathbb{P}[X = 1]$ is called the parameter of X.

Example

A die is rolled. Let X be the random variable that tells us whether the outcome is larger than 4 or not.

Binomial distribution