



PROBABILITY

Adam Klepáč

February 15, 2024

PROBABILITY DISTRIBUTION

The bottom of the slide features a decorative design consisting of two large, dark red triangles that point towards each other, meeting at a central point. The background of the slide is white.

WHAT IS PROBABILITY DISTRIBUTION?

Probability distribution is a function that describes the probability of different *possible values* of a *random variable*.



WHAT IS PROBABILITY DISTRIBUTION?

Probability distribution is a function that describes the probability of different *possible values* of a *random variable*.
graphs.



WHAT IS PROBABILITY DISTRIBUTION?

Probability distribution is a function that describes the probability of different *possible values* of a *random variable*.

graphs.

For example, imagine the random variable X that describes the outcome of a coin toss, that is $X \in \{\text{heads}, \text{tails}\}$.

WHAT IS PROBABILITY DISTRIBUTION?

Probability distribution is a function that describes the probability of different *possible values* of a *random variable*.

graphs.

For example, imagine the random variable X that describes the outcome of a coin toss, that is $X \in \{\text{heads}, \text{tails}\}$.

If the coin is fair then $P(X = \text{heads}) = P(X = \text{tails}) = 1/2$.

WHAT IS PROBABILITY DISTRIBUTION?

Probability distribution is a function that describes the probability of different *possible values* of a *random variable*.

graphs.

For example, imagine the random variable X that describes the outcome of a coin toss, that is $X \in \{\text{heads}, \text{tails}\}$.

If the coin is fair then $P(X = \text{heads}) = P(X = \text{tails}) = 1/2$.

The **probability distribution** of this random variable is a function $f : \{\text{heads}, \text{tails}\} \rightarrow [0, 1]$ which assigns to the element 'heads' the probability $P(X = \text{heads})$ and to 'tails' the probability $P(X = \text{tails})$.

WHAT IS PROBABILITY DISTRIBUTION?

Probability distribution is a function that describes the probability of different *possible values* of a *random variable*.

graphs.

For example, imagine the random variable X that describes the outcome of a coin toss, that is $X \in \{\text{heads}, \text{tails}\}$.

If the coin is fair then $P(X = \text{heads}) = P(X = \text{tails}) = 1/2$.

The **probability distribution** of this random variable is a function $f : \{\text{heads}, \text{tails}\} \rightarrow [0, 1]$ which assigns to the element 'heads' the probability $P(X = \text{heads})$ and to 'tails' the probability $P(X = \text{tails})$.

In other words, $f(\text{heads}) = f(\text{tails}) = 1/2$.

PROBABILITY DISTRIBUTION – EXAMPLES

- The **probability distribution** of a random variable representing the value of a dice roll is a function

$$f : \{1, 2, 3, 4, 5, 6\} \rightarrow [0, 1]$$

such that $f(k) = 1/6$ for all numbers $k \in \{1, 2, 3, 4, 5, 6\}$.

PROBABILITY DISTRIBUTION – EXAMPLES

- The **probability distribution** of a random variable representing the value of a dice roll is a function

$$f : \{1, 2, 3, 4, 5, 6\} \rightarrow [0, 1]$$

such that $f(k) = 1/6$ for all numbers $k \in \{1, 2, 3, 4, 5, 6\}$.

- The **probability distribution** of a random variable representing the rank of a randomly chosen playing card is a function

PROBABILITY DISTRIBUTION – EXAMPLES

- The **probability distribution** of a random variable representing the value of a dice roll is a function

$$f : \{1, 2, 3, 4, 5, 6\} \rightarrow [0, 1]$$

such that $f(k) = 1/6$ for all numbers $k \in \{1, 2, 3, 4, 5, 6\}$.

- The **probability distribution** of a random variable representing the rank of a randomly chosen playing card is a function

$$f : \{2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A\} \rightarrow [0, 1]$$

such that $f(r) = 4/52$ where r is a rank of a playing card.

VISUALIZING PROBABILITY DISTRIBUTIONS – TABLES

Discrete **probability distributions** (meaning distributions of a *discrete* random variable) can be easily represented using **tables**.

VISUALIZING PROBABILITY DISTRIBUTIONS – TABLES

Discrete **probability distributions** (meaning distributions of a *discrete* random variable) can be easily represented using **tables**.

For example, the probability distribution of a dice roll is given simply by

Roll	1	2	3	4	5	6
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$.

VISUALIZING PROBABILITY DISTRIBUTIONS – TABLES

For a more abstract example, if X can attain any of the four values a, b, c, d with probabilities $P(X = a) = 3/10, P(X = b) = 5/10, P(X = c) = 1/10, P(X = d) = 1/10$, then its probability distribution is

Value	a	b	c	d
Probability	$\frac{3}{10}$	$\frac{5}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

VISUALIZING PROBABILITY DISTRIBUTIONS – GRAPHS



Probability distributions (both *discrete* and *continuous*) can be represented as graphs. These are your typical function graphs which draw inputs on the x -axis and outputs on the y -axis.

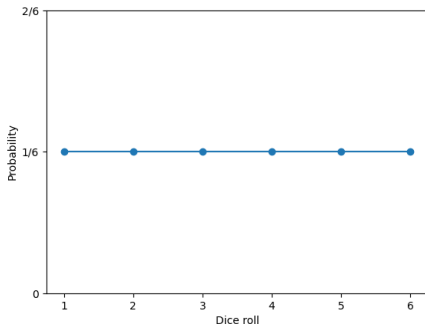


VISUALIZING PROBABILITY DISTRIBUTIONS – GRAPHS



Probability distributions (both *discrete* and *continuous*) can be represented as graphs. These are your typical function graphs which draw inputs on the x -axis and outputs on the y -axis.

The probability distribution of a dice roll looks like this

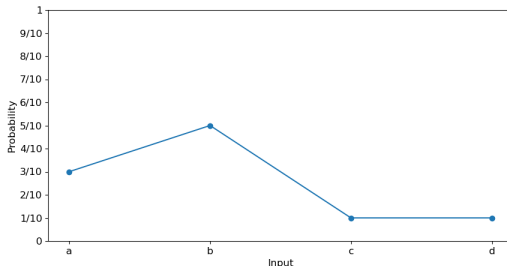


VISUALIZING PROBABILITY DISTRIBUTIONS – GRAPHS

The probability distribution from this table

Value	a	b	c	d
Probability	$\frac{3}{10}$	$\frac{5}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

looks like this



1

DISCRETE PROBABILITY DISTRIBUTIONS



DISCRETE PROBABILITY DISTRIBUTION

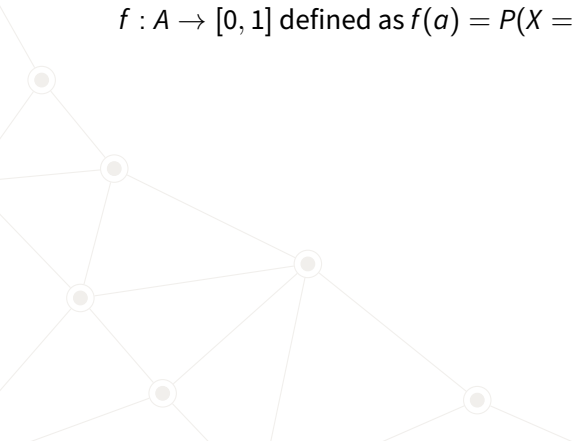
Let X be a random variable taking values from a set A .



DISCRETE PROBABILITY DISTRIBUTION

Let X be a random variable taking values from a set A .

- The **probability distribution** (or also **probability mass function**) of X is the function $f : A \rightarrow [0, 1]$ defined as $f(a) = P(X = a)$ for $a \in A$.



DISCRETE PROBABILITY DISTRIBUTION

Let X be a random variable taking values from a set A .

- The **probability distribution** (or also **probability mass function**) of X is the function $f : A \rightarrow [0, 1]$ defined as $f(a) = P(X = a)$ for $a \in A$.
- The **cumulative distribution function** of X gives the probability that a random variable is **less than a certain value**. It is defined as $F(a) = P(X \leq a)$ for $a \in A$.

DISCRETE PROBABILITY DISTRIBUTION

Let X be a random variable taking values from a set A .

- The **probability distribution** (or also **probability mass function**) of X is the function $f : A \rightarrow [0, 1]$ defined as $f(a) = P(X = a)$ for $a \in A$.
- The **cumulative distribution function** of X gives the probability that a random variable is **less than a certain value**. It is defined as $F(a) = P(X \leq a)$ for $a \in A$.
- The **mean** of X is defined as $E(X) = \sum_{a \in A} a \cdot P(X = a)$. It represents the 'expected' value of X .

DISCRETE PROBABILITY DISTRIBUTION

Let X be a random variable taking values from a set A .

- The **probability distribution** (or also **probability mass function**) of X is the function $f : A \rightarrow [0, 1]$ defined as $f(a) = P(X = a)$ for $a \in A$.
- The **cumulative distribution function** of X gives the probability that a random variable is **less than a certain value**. It is defined as $F(a) = P(X \leq a)$ for $a \in A$.
- The **mean** of X is defined as $E(X) = \sum_{a \in A} a \cdot P(X = a)$. It represents the 'expected' value of X .
- The **variance** (describing the *dispersion* of the distribution around the mean) of X is defined as

$$\text{Var}(X) = \sum_{a \in A} (a - E(X))^2 \cdot P(X = a).$$

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Let's see what these concepts mean in a simple statistical experiment.



DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Let's see what these concepts mean in a simple statistical experiment.

Suppose we measure the height of a randomly picked 20-year-old males. We might get something akin to the following table

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Let's see what these concepts mean in a simple statistical experiment.

Suppose we measure the height of a randomly picked 20-year-old males. We might get something akin to the following table

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

We can easily calculate the mean and standard deviation of this data.

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

Using the formula for the arithmetic mean, we get

$$\bar{x} = \frac{175 \cdot 13 + 176 \cdot 20 + \dots + 183 \cdot 3}{13 + 20 + \dots + 3} = 178.1$$

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

The standard deviation is then

$$\sigma = \sqrt{\frac{13 \cdot (175 - 178.1)^2 + 20 \cdot (176 - 178.1)^2 + \dots + 3 \cdot (183 - 178.1)^2}{13 + 20 + \dots + 3}} = 8.203.$$

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

Let's now define a random variable X which can be any of those heights in the table above.

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

Let's now define a random variable X which can be any of those heights in the table above.

We define the probabilities that X is a particular height based on the counts above. That gives the following table

Height	175	176	177	178	179	180	181	182	183
Probability	$\frac{13}{100}$	$\frac{20}{100}$	$\frac{11}{100}$	$\frac{17}{100}$	$\frac{11}{100}$	$\frac{8}{100}$	$\frac{10}{100}$	$\frac{7}{100}$	$\frac{3}{100}$

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Count	13	20	11	17	11	8	10	7	3

Let's now define a random variable X which can be any of those heights in the table above.

We define the probabilities that X is a particular height based on the counts above. That gives the following table

Height	175	176	177	178	179	180	181	182	183
Probability	$\frac{13}{100}$	$\frac{20}{100}$	$\frac{11}{100}$	$\frac{17}{100}$	$\frac{11}{100}$	$\frac{8}{100}$	$\frac{10}{100}$	$\frac{7}{100}$	$\frac{3}{100}$

In other words, this gives a **distribution function** f of X where the set

$A = \{175, 176, 177, 178, 179, 180, 181, 182, 183\}$ and the outputs of f on each of these numbers are given by the table above.

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Probability	$\frac{13}{100}$	$\frac{20}{100}$	$\frac{11}{100}$	$\frac{17}{100}$	$\frac{11}{100}$	$\frac{8}{100}$	$\frac{10}{100}$	$\frac{7}{100}$	$\frac{3}{100}$

- The **cumulative distribution function** F describes the probability that a randomly chosen person from the group has height *less than* a particular number.

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Probability	$\frac{13}{100}$	$\frac{20}{100}$	$\frac{11}{100}$	$\frac{17}{100}$	$\frac{11}{100}$	$\frac{8}{100}$	$\frac{10}{100}$	$\frac{7}{100}$	$\frac{3}{100}$

- The **cumulative distribution function** F describes the probability that a randomly chosen person from the group has height *less than* a particular number. For example,

$$\begin{aligned}
 F(178) &= P(X \leq 178) = P(X = 175) + P(X = 176) + P(X = 177) + P(X = 178) \\
 &= \frac{13}{100} + \frac{20}{100} + \frac{11}{100} + \frac{17}{100} = \frac{61}{100}.
 \end{aligned}$$

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Probability	$\frac{13}{100}$	$\frac{20}{100}$	$\frac{11}{100}$	$\frac{17}{100}$	$\frac{11}{100}$	$\frac{8}{100}$	$\frac{10}{100}$	$\frac{7}{100}$	$\frac{3}{100}$

- The **mean** of X is the same as the arithmetic mean of the data. Indeed,

$$\begin{aligned}
 E(X) &= \sum_{a \in A} a \cdot P(X = a) \\
 &= 175 \cdot P(X = 175) + 176 \cdot P(X = 176) + \dots + 183 \cdot P(X = 183) \\
 &= 175 \cdot \frac{13}{100} + 176 \cdot \frac{20}{100} + \dots + 183 \cdot \frac{3}{100} = 178.1.
 \end{aligned}$$

DISCRETE PROBABILITY DISTRIBUTION – EXAMPLE

Height	175	176	177	178	179	180	181	182	183
Probability	$\frac{13}{100}$	$\frac{20}{100}$	$\frac{11}{100}$	$\frac{17}{100}$	$\frac{11}{100}$	$\frac{8}{100}$	$\frac{10}{100}$	$\frac{7}{100}$	$\frac{3}{100}$

- The **variance** of X is the same as the standard deviation *squared* (that is, $\text{Var}(X) = \sigma^2$). Indeed,

$$\begin{aligned}
 \text{Var}(X) &= \sum_{a \in A} (a - E(X))^2 \cdot P(X = a) \\
 &= (175 - 178.1)^2 \cdot P(X = 175) + \dots + (183 - 178.1)^2 \cdot P(X = 183) \\
 &= 67.29 = 8.203^2.
 \end{aligned}$$