

Basic Probability Theory

Jiří Neubauer

Department of Econometrics FVL UO Brno
office 69a, tel. 973 442029
email: Jiri.Neubauer@unob.cz

Basic Probability Theory

- The random experiment and the random event
- The probability of the random experiment, the axioms of probability
- Definition of the probability (statistical, classical, geometrical)
- The conditional probability and independence
- The addition and multiplication law
- The law of total probability and Bayes' formula

Basic Probability Theory

- The random experiment and the random event
- The probability of the random experiment, the axioms of probability
- Definition of the probability (statistical, classical, geometrical)
- The conditional probability and independence
- The addition and multiplication law
- The law of total probability and Bayes' formula

Basic Probability Theory

- The random experiment and the random event
- The probability of the random experiment, the axioms of probability
- Definition of the probability (statistical, classical, geometrical)
- The conditional probability and independence
- The addition and multiplication law
- The law of total probability and Bayes' formula

Basic Probability Theory

- The random experiment and the random event
- The probability of the random experiment, the axioms of probability
- Definition of the probability (statistical, classical, geometrical)
- The conditional probability and independence
- The addition and multiplication law
- The law of total probability and Bayes' formula

Basic Probability Theory

- The random experiment and the random event
- The probability of the random experiment, the axioms of probability
- Definition of the probability (statistical, classical, geometrical)
- The conditional probability and independence
- The addition and multiplication law
- The law of total probability and Bayes' formula

Basic Probability Theory

- The random experiment and the random event
- The probability of the random experiment, the axioms of probability
- Definition of the probability (statistical, classical, geometrical)
- The conditional probability and independence
- The addition and multiplication law
- The law of total probability and Bayes' formula

The Random Experiment and the Random Event

- **The random experiment** is every activity which results is not unambiguously determined by ambient conditions (e.g. rolling a die (tossing a coin), measurement of length, 100 metres run, lottery, ...).
- The result of the random experiment we call **outcome** (e.g. in two tosses of a coin we get 2 heads, the length is 25.7 cm, time is 13.8 s, the winning lottery ticket is B265430, ...).

An experiment is the process of observing a phenomenon that has variation in its outcomes.

The Random Experiment and the Random Event

- **The random experiment** is every activity which results is not unambiguously determined by ambient conditions (e.g. rolling a die (tossing a coin), measurement of length, 100 metres run, lottery, ...).
- The result of the random experiment we call **outcome** (e.g. in two tosses of a coin we get 2 heads, the length is 25.7 cm, time is 13.8 s, the winning lottery ticket is B265430, ...).

An experiment is the process of observing a phenomenon that has variation in its outcomes.

The Random Experiment and the Random Event

- **The random experiment** is every activity which results is not unambiguously determined by ambient conditions (e.g. rolling a die (tossing a coin), measurement of length, 100 metres run, lottery, ...).
- The result of the random experiment we call **outcome** (e.g. in two tosses of a coin we get 2 heads, the length is 25.7 cm, time is 13.8 s, the winning lottery ticket is B265430, ...).

An experiment is the process of observing a phenomenon that has variation in its outcomes.

The Random Experiment and the Random Event

The **sample space** Ω associated with an experiment is the collection of all possible distinct outcomes of the experiment.

Each outcome is called an **elementary outcome**, a **simple event** or an **element of the sample space** – denote $\omega_1, \omega_2, \dots$.

An **event** is the set of elementary outcomes possessing a designated feature.

The sample space can

- discrete finite – roll a die and observe the number:
 $\Omega = \{1, 2, 3, 4, 5, 6\}$
- discrete infinite – roll a die repeatedly and count the number of rolls it takes until the first 6 appears $\Omega = \{1, 2, \dots\}$
- continuous – turn on a light bulb and measure its lifetime

$$\Omega = \langle 0, \infty \rangle$$

The Random Experiment and the Random Event

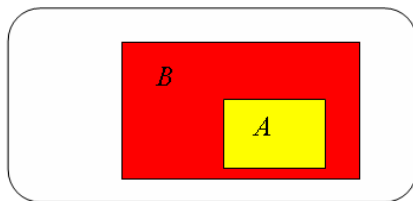
A subset of Ω , $A \subseteq \Omega$, is called an **event**.

If we roll a die and observe the number, two possible events are that we get odd outcome and that we get at least 4.

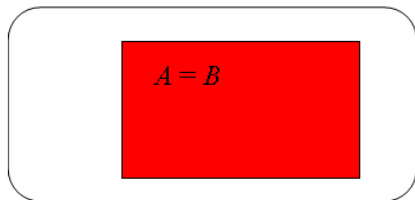
$$A = \{1, 3, 5\} \quad B = \{4, 5, 6\}$$

Basic set operations and:

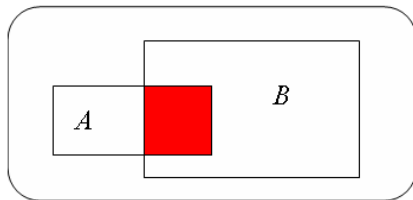
- an event A is a part of an event B ($A \subset B$)
if A occurs then B occurs



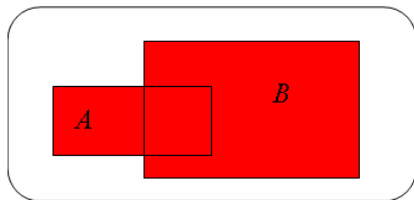
- events A and B are equivalent ($A = B$)
 A occurs if and only if B occurs



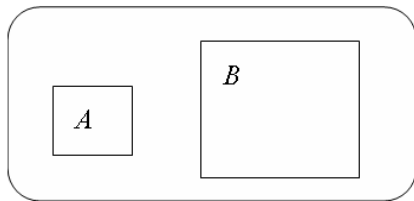
- the intersection of A and B ($A \cap B$)
both A and B occur



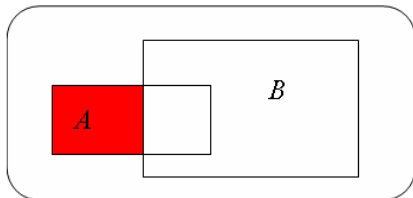
- the union of A and B ($A \cup B$)
 A or B (or both) occurs



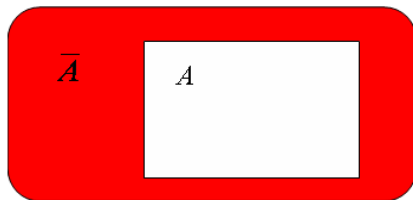
- A and B disjoint or mutually exclusive if they cannot both occur ($A \cap B = \emptyset$), where \emptyset is an impossible event



- difference between A and B ($A - B$)
 A occurs but not B

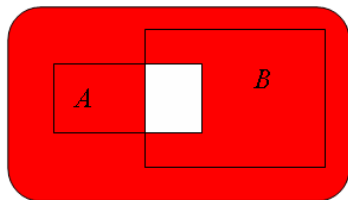


- the complement of A (denote \bar{A})
A does not occur

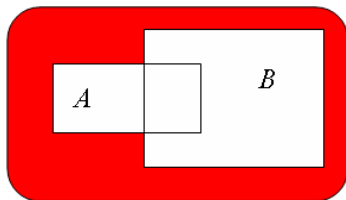


$$A \cup \bar{A} = \Omega, A \cap \bar{A} = \emptyset.$$

It easy to see that



$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

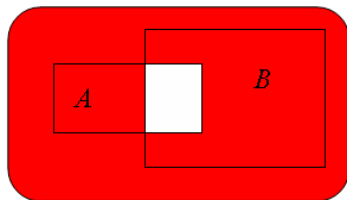


$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

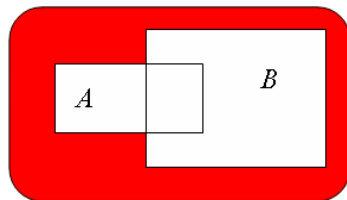
For more than two events we can write

$$\overline{\bigcap_i A_i} = \bigcup_i \bar{A}_i \quad \text{and} \quad \overline{\bigcup_i A_i} = \bigcap_i \bar{A}_i.$$

It easy to see that



$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$



$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

For more than two events we can write

$$\overline{\bigcap_i A_i} = \bigcup_i \bar{A}_i \quad \text{and} \quad \overline{\bigcup_i A_i} = \bigcap_i \bar{A}_i.$$

The Axioms of Probability

- The probability of an event A is a non-negative number ($P(A) \geq 0$).
- The probability of Ω is equal to 1 ($P(\Omega) = 1$).
- If A_1, A_2, \dots is a sequence of pairwise disjoint events, that is, if $i \neq j$ then $A_i \cap A_j = \emptyset$, then

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

The Axioms of Probability

- The probability of an event A is a non-negative number ($P(A) \geq 0$).
- The probability of Ω is equal to 1 ($P(\Omega) = 1$).
- If A_1, A_2, \dots is a sequence of pairwise disjoint events, that is, if $i \neq j$ then $A_i \cap A_j = \emptyset$, then

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

The Axioms of Probability

- The probability of an event A is a non-negative number ($P(A) \geq 0$).
- The probability of Ω is equal to 1 ($P(\Omega) = 1$).
- If A_1, A_2, \dots is a sequence of pairwise disjoint events, that is, if $i \neq j$ then $A_i \cap A_j = \emptyset$, then

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

We can derive for the axioms another properties of probability:

- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$

(Ω and \emptyset are disjoint events and $\Omega \cup \emptyset = \Omega$, we can write $P(\Omega) + P(\emptyset) = 1$, $1 + P(\emptyset) = 1 \Rightarrow P(\emptyset) = 0$.)

- $P(\bar{A}) = 1 - P(A)$

(A, \bar{A} are disjoint events and $P(A) \cup P(\bar{A}) = 1$, then $P(A) + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - P(A)$.)

We can derive for the axioms another properties of probability:

- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$

(Ω and \emptyset are disjoint events and $\Omega \cup \emptyset = \Omega$, we can write $P(\Omega) + P(\emptyset) = 1$, $1 + P(\emptyset) = 1 \Rightarrow P(\emptyset) = 0$.)

- $P(\bar{A}) = 1 - P(A)$

(A, \bar{A} are disjoint events and $P(A) \cup P(\bar{A}) = 1$, then $P(A) + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - P(A)$.)

We can derive for the axioms another properties of probability:

- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$

(Ω and \emptyset are disjoint events and $\Omega \cup \emptyset = \Omega$, we can write $P(\Omega) + P(\emptyset) = 1$, $1 + P(\emptyset) = 1 \Rightarrow P(\emptyset) = 0$.)

- $P(\bar{A}) = 1 - P(A)$

(A, \bar{A} are disjoint events and $P(A) \cup P(\bar{A}) = 1$, then $P(A) + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - P(A)$.)

We can derive for the axioms another properties of probability:

- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$

(Ω and \emptyset are disjoint events and $\Omega \cup \emptyset = \Omega$, we can write $P(\Omega) + P(\emptyset) = 1$, $1 + P(\emptyset) = 1 \Rightarrow P(\emptyset) = 0$.)

- $P(\bar{A}) = 1 - P(A)$

(A, \bar{A} are disjoint events and $P(A) \cup P(\bar{A}) = 1$, then $P(A) + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - P(A)$.)

We can derive for the axioms another properties of probability:

- $0 \leq P(A) \leq 1$
- $P(\emptyset) = 0$

(Ω and \emptyset are disjoint events and $\Omega \cup \emptyset = \Omega$, we can write $P(\Omega) + P(\emptyset) = 1$, $1 + P(\emptyset) = 1 \Rightarrow P(\emptyset) = 0$.)

- $P(\bar{A}) = 1 - P(A)$

(A, \bar{A} are disjoint events and $P(A) \cup P(\bar{A}) = 1$, then $P(A) + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - P(A)$.)

We can derive for the axioms another properties of probability:

- If $A \subset B$ then $0 \leq P(A) \leq P(B)$.

(If $A \subset B$, then $A \cap B = A$ and for an event B we can write $B = \Omega \cap B = (A \cup \bar{A}) \cap B = (A \cap B) \cup (\bar{A} \cap B) = A \cup (\bar{A} \cap B)$. A and $\bar{A} \cap B$ are disjoint which implies $P(B) = P(A) + P(\bar{A} \cap B) \geq P(A)$, because $P(\bar{A} \cap B) \geq 0$.)

- If $A \subset B$ then $P(B - A) = P(B) - P(A)$.

($B - A = \bar{A} \cap B$, using previous results we get

$$P(B) = P(A) + P(B - A) \Rightarrow P(B - A) = P(B) - P(A).$$

We can derive for the axioms another properties of probability:

- If $A \subset B$ then $0 \leq P(A) \leq P(B)$.

(If $A \subset B$, then $A \cap B = A$ and for an event B we can write $B = \Omega \cap B = (A \cup \bar{A}) \cap B = (A \cap B) \cup (\bar{A} \cap B) = A \cup (\bar{A} \cap B)$. A and $\bar{A} \cap B$ are disjoint which implies $P(B) = P(A) + P(\bar{A} \cap B) \geq P(A)$, because $P(\bar{A} \cap B) \geq 0$.)

- If $A \subset B$ then $P(B - A) = P(B) - P(A)$.

($B - A = \bar{A} \cap B$, using previous results we get

$$P(B) = P(A) + P(B - A) \Rightarrow P(B - A) = P(B) - P(A).$$

We can derive for the axioms another properties of probability:

- If $A \subset B$ then $0 \leq P(A) \leq P(B)$.

(If $A \subset B$, then $A \cap B = A$ and for an event B we can write $B = \Omega \cap B = (A \cup \bar{A}) \cap B = (A \cap B) \cup (\bar{A} \cap B) = A \cup (\bar{A} \cap B)$. A and $\bar{A} \cap B$ are disjoint which implies $P(B) = P(A) + P(\bar{A} \cap B) \geq P(A)$, because $P(\bar{A} \cap B) \geq 0$.)

- If $A \subset B$ then $P(B - A) = P(B) - P(A)$.

($B - A = \bar{A} \cap B$, using previous results we get

$$P(B) = P(A) + P(B - A) \Rightarrow P(B - A) = P(B) - P(A).$$

We can derive for the axioms another properties of probability:

- If $A \subset B$ then $0 \leq P(A) \leq P(B)$.

(If $A \subset B$, then $A \cap B = A$ and for an event B we can write $B = \Omega \cap B = (A \cup \bar{A}) \cap B = (A \cap B) \cup (\bar{A} \cap B) = A \cup (\bar{A} \cap B)$. A and $\bar{A} \cap B$ are disjoint which implies $P(B) = P(A) + P(\bar{A} \cap B) \geq P(A)$, because $P(\bar{A} \cap B) \geq 0$.)

- If $A \subset B$ then $P(B - A) = P(B) - P(A)$.

($B - A = \bar{A} \cap B$, using previous results we get

$$P(B) = P(A) + P(B - A) \Rightarrow P(B - A) = P(B) - P(A).$$

Statistical Definition of Probability

Definition

Statistical definition of probability is given by the formula

$$P(A) \approx \frac{n(A)}{n},$$

where n is a number of observations, $n(A)$ denotes a frequency of an event A .

The ratio $\frac{n(A)}{n}$ we call a relative frequency. If the number of observation is big, the relative frequency of the event A is close to the probability of A .

Classical Definition of Probability

Definition

If the sample space $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ is finite and all elementary events have the same probability $P(\omega_i = \frac{1}{n})$ for $i = 1, 2, \dots, n$, then the probability of an event A is

$$P(A) = \frac{m}{n},$$

where m is a number of favorable outcomes of A , where n is a number of possible outcomes of A .

Example 1

We have 32 cards (8 red, 8 black, 8 green and 8 blue), 4 cards are randomly chosen. What is the probability that we get 3 red cards?

Number of possible outcomes is

$$n = \binom{32}{4},$$

Number of favorable outcomes is

$$m = \binom{8}{3} \binom{24}{1},$$

$$P(A) = \frac{m}{n} = \frac{\binom{8}{3} \binom{24}{1}}{\binom{32}{4}} = 0.037.$$

Example 1

We have 32 cards (8 red, 8 black, 8 green and 8 blue), 4 cards are randomly chosen. What is the probability that we get 3 red cards?

Number of possible outcomes is

$$n = \binom{32}{4},$$

Number of favorable outcomes is

$$m = \binom{8}{3} \binom{24}{1},$$

$$P(A) = \frac{m}{n} = \frac{\binom{8}{3} \binom{24}{1}}{\binom{32}{4}} = 0.037.$$

Example 1

We have 32 cards (8 red, 8 black, 8 green and 8 blue), 4 cards are randomly chosen. What is the probability that we get 3 red cards?

Number of possible outcomes is

$$n = \binom{32}{4},$$

Number of favorable outcomes is

$$m = \binom{8}{3} \binom{24}{1},$$

$$P(A) = \frac{m}{n} = \frac{\binom{8}{3} \binom{24}{1}}{\binom{32}{4}} = 0.037.$$

Example 1

We have 32 cards (8 red, 8 black, 8 green and 8 blue), 4 cards are randomly chosen. What is the probability that we get 3 red cards?

Number of possible outcomes is

$$n = \binom{32}{4},$$

Number of favorable outcomes is

$$m = \binom{8}{3} \binom{24}{1},$$

$$P(A) = \frac{m}{n} = \frac{\binom{8}{3} \binom{24}{1}}{\binom{32}{4}} = 0.037.$$

Example 2

We roll 2 dice at once. What is the probability that the sum of numbers is equal to 3?

There are only 2 possibilities: 1+2 or 2+1.

We get the probability

$$P(A) = \frac{m}{n} = \frac{2}{36} = \frac{1}{18} = 0.056.$$

Example 2

We roll 2 dice at once. What is the probability that the sum of numbers is equal to 3?

There are only 2 possibilities: 1+2 or 2+1.

We get the probability

$$P(A) = \frac{m}{n} = \frac{2}{36} = \frac{1}{18} = 0.056.$$

Example 2

We roll 2 dice at once. What is the probability that the sum of numbers is equal to 3?

There are only 2 possibilities: 1+2 or 2+1.

We get the probability

$$P(A) = \frac{m}{n} = \frac{2}{36} = \frac{1}{18} = 0.056.$$

Example 3

It is known that among 100 products are 5 waste products. We choose randomly 3 products (not returning). What is the probability that we get 1 waste product?

Number of possible outcomes is

$$n = \binom{100}{3},$$

Number of favorable outcomes is

$$m = \binom{5}{1} \binom{95}{2},$$

$$P(A) = \frac{m}{n} = \frac{\binom{5}{1} \binom{95}{2}}{\binom{100}{3}} = 0.138$$

Example 3

It is known that among 100 products are 5 waste products. We choose randomly 3 products (not returning). What is the probability that we get 1 waste product?

Number of possible outcomes is

$$n = \binom{100}{3},$$

Number of favorable outcomes is

$$m = \binom{5}{1} \binom{95}{2},$$

$$P(A) = \frac{m}{n} = \frac{\binom{5}{1} \binom{95}{2}}{\binom{100}{3}} = 0.138$$

Example 3

It is known that among 100 products are 5 waste products. We choose randomly 3 products (not returning). What is the probability that we get 1 waste product?

Number of possible outcomes is

$$n = \binom{100}{3},$$

Number of favorable outcomes is

$$m = \binom{5}{1} \binom{95}{2},$$

$$P(A) = \frac{m}{n} = \frac{\binom{5}{1} \binom{95}{2}}{\binom{100}{3}} = 0.138$$

Example 3

It is known that among 100 products are 5 waste products. We choose randomly 3 products (not returning). What is the probability that we get 1 waste product?

Number of possible outcomes is

$$n = \binom{100}{3},$$

Number of favorable outcomes is

$$m = \binom{5}{1} \binom{95}{2},$$

$$P(A) = \frac{m}{n} = \frac{\binom{5}{1} \binom{95}{2}}{\binom{100}{3}} = 0.138$$

Geometrical Definition of Probability

If sample space Ω is continuous it is not possible to calculate probability of an event A by the formula of classical probability (number of possible outcomes is infinite)

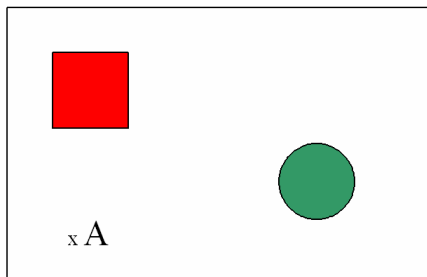
Definition

The sample space Ω is continuous creating some area which is bounded and closed and it's size is $V(\Omega)$ (described by the length, eventually the area or the volume). If an event $A \subset \Omega$ creates area with the size $V(A)$ then the probability of A is given by the formula

$$P(A) = \frac{V(A)}{V(\Omega)}.$$

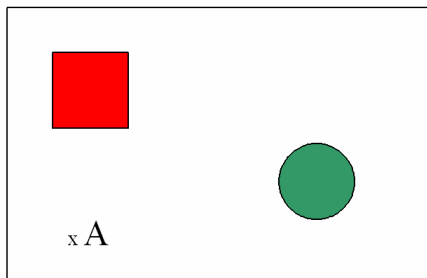
Example 1

Let us have a rectangle 10×5 cm. There are a square (2×2 cm), a circle (with diameter 2 cm) and a point A inside the rectangle. What is the probability that the randomly chosen point inside the rectangle is inside the square, inside the circle and is the point A ?



Example 1

Let us have a rectangle 10×5 cm. There are a square (2×2 cm), a circle (with diameter 2 cm) and a point A inside the rectangle. What is the probability that the randomly chosen point inside the rectangle is inside the square, inside the circle and is the point A ?



Example 1

The probability that the randomly chosen point is inside the square is:

- $V(A) = 2 \cdot 2 = 4 \text{ cm}^2$
- $V(\Omega) = 10 \cdot 5 = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{4}{50} = 0.08$$

Example 1

The probability that the randomly chosen point is inside the square is:

- $V(A) = 2 \cdot 2 = 4 \text{ cm}^2$
- $V(\Omega) = 10 \cdot 5 = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{4}{50} = 0.08$$

Example 1

The probability that the randomly chosen point is inside the square is:

- $V(A) = 2 \cdot 2 = 4 \text{ cm}^2$
- $V(\Omega) = 10 \cdot 5 = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{4}{50} = 0.08$$

Example 1

The probability that the randomly chosen point is inside the circle is:

- $V(A) = \pi \text{ cm}^2$
- $V(\Omega) = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{\pi}{50} = 0.0628$$

Example 1

The probability that the randomly chosen point is inside the circle is:

- $V(A) = \pi \text{ cm}^2$
- $V(\Omega) = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{\pi}{50} = 0.0628$$

Example 1

The probability that the randomly chosen point is inside the circle is:

- $V(A) = \pi \text{ cm}^2$
- $V(\Omega) = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{\pi}{50} = 0.0628$$

Example 1

The probability that the randomly chosen point is the point A is:

- $V(A) = 0 \text{ cm}^2$
- $V(\Omega) = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{0}{50} = 0$$

Example 1

The probability that the randomly chosen point is the point A is:

- $V(A) = 0 \text{ cm}^2$
- $V(\Omega) = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{0}{50} = 0$$

Example 1

The probability that the randomly chosen point is the point A is:

- $V(A) = 0 \text{ cm}^2$
- $V(\Omega) = 50 \text{ cm}^2$
-

$$P(A) = \frac{V(A)}{V(\Omega)} = \frac{0}{50} = 0$$

Example 2

Let us assume that x, y are real numbers in the interval $[0; 1]$. What is the probability that their sum is smaller than 1 and their product is not bigger than $2/9$ at the same time?

- the sum smaller than 1

$$x + y < 1 \Leftrightarrow y < 1 - x$$

- the product smaller than $2/9$

$$xy < \frac{2}{9} \Leftrightarrow y < \frac{2}{9x}$$

Example 2

Let us assume that x, y are real numbers in the interval $[0; 1]$. What is the probability that their sum is smaller than 1 and their product is not bigger than $2/9$ at the same time?

- the sum smaller than 1

$$x + y < 1 \Leftrightarrow y < 1 - x$$

- the product smaller than $2/9$

$$xy < \frac{2}{9} \Leftrightarrow y < \frac{2}{9x}$$

Example 2

Let us assume that x, y are real numbers in the interval $[0; 1]$. What is the probability that their sum is smaller than 1 and their product is not bigger than $2/9$ at the same time?

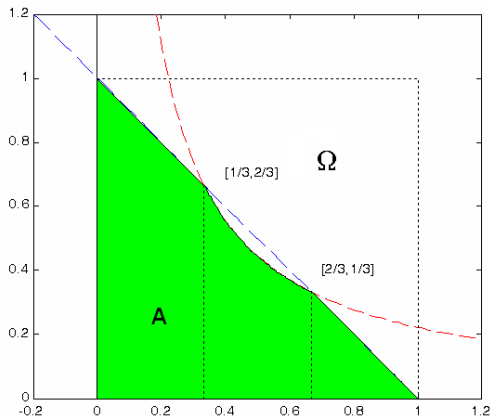
- the sum smaller than 1

$$x + y < 1 \Leftrightarrow y < 1 - x$$

- the product smaller than $2/9$

$$xy < \frac{2}{9} \Leftrightarrow y < \frac{2}{9x}$$

Example 2



Example 2



$$V(\Omega) = 1$$



$$V(A) = \int_0^{1/3} (1-x) dx + \int_{1/3}^{2/3} \frac{2}{9x} dx + \int_{2/3}^1 (1-x) dx = 0.487$$



$$P(A) = \frac{V(A)}{V(\Omega)} = 0.487$$

Example 2



$$V(\Omega) = 1$$



$$V(A) = \int_0^{1/3} (1-x) dx + \int_{1/3}^{2/3} \frac{2}{9x} dx + \int_{2/3}^1 (1-x) dx = 0.487$$



$$P(A) = \frac{V(A)}{V(\Omega)} = 0.487$$

Example 2



$$V(\Omega) = 1$$



$$V(A) = \int_0^{1/3} (1-x) dx + \int_{1/3}^{2/3} \frac{2}{9x} dx + \int_{2/3}^1 (1-x) dx = 0.487$$



$$P(A) = \frac{V(A)}{V(\Omega)} = 0.487$$