# Математика для Data Science. Теория вероятностей. Шпаргалка

## Содержание

# Первая неделя. Вероятностное пространство

## Вероятностное пространство

*Вероятностное пространство* это тройка $(\Omega, F, P)$, где

1. $\Omega$ это любое множество. Оно называется *множеством элементарных исходов*, а его элементы – *элементарными исходами*. Мы пока будем заниматься только случаем, когда $\Omega$ это конечное множество. Оставшаяся часть определения написана только для этого случая.

2. Все подмножества $\Omega$ называются *событиями*. Множество событий обозначается $F$ и называется *алгеброй событий*.

   **Обозначение.** Для обозначения событий мы будем использовать заглавные буквы из начала латинского алфавита: $A, B, C$, и так далее.

3. *Вероятность* $P$ это функция из $F$ в $[0, 1]$. Другими словами, $P$ каждому событию сопоставляет число от 0 до 1. Это число называется *вероятностью* соответствующего события. Вероятность каждого события должна быть равна сумме вероятностей элементарных исходов, из которых состоит это событие. Сумма вероятностей всех элементарных исходов должна быть равна 1.

**Обозначение.** Если $A$ это конечное множество, то мы обозначаем число элементов $A$ через $|A|$.
Если число исходов в $\Omega$ равно $k$, то число событий равно $2^k$, то есть $|\Omega| = k, k \in \mathbb{N} \implies |F| = 2^k$.
Пусть дано событие $A$. Тогда событием $\bar{A}$ называется событие, состоящие из всех элементарных исходов, которые не входят в $A$. Можно произносить $\bar{A}$ как "не $A$."Верно следующее равенство: $P(\bar{A}) = 1 - P(A)$.

## Равновероятные исходы

Если вероятности всех элементарных исходов равны, то есть $P(a) = P(b)$ для всех $a, b \in \Omega$, то в этом случае мы говорим, что все элементарные исходы *равновероятны*.

Если все исходы равновероятны, то для любого события $A$ выполнено $P(A) = \dfrac{|A|}{|\Omega|}$. В частности, вероятность каждого элементарного исхода равна $\frac{1}{k}$.

## Условная вероятность и независимые события

Так как события являются множествами (состоящими из элементарных исходов), к ним можно применять все стандартные операции над множествами. В частности, события можно пересекать (символ $\cap$) и объединять (символ $\cup$).

Если $P(B) \neq 0$, то *условной вероятностью события $A$ при условии события $B$* называется дробь $\dfrac{P(A \cap B)}{P(B)}$.
Обозначение:

$$P(A|B) := \frac{P(A \cap B)}{P(B)}.$$

**Формула полной вероятности для двух событий:**

$$P(B) = P(B \cap A) + P(B \cap \bar{A}) \text{ или, равносильно,}$$
$$P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$$

## Независимые события

Пусть $P(B) \neq 0$. Событие $A$ называется *независимым от события $B$*, если и только если $P(A) = P(A|B)$.
Пусть $P(A) \neq 0$ и $P(B) \neq 0$. Тогда верна следующая цепочка эквивалентных утверждений:

$$\text{События } A \text{ и } B \textit{ независимы} \Leftrightarrow P(A) = P(A|B) \Leftrightarrow P(B) = P(B|A) \Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$$

Два события, которые не являются независимыми, называют *зависимыми*.

## Совместная независимость

События $A_1, \dots, A_n$ называются *совместно независимыми*, если для любого $k$ и любого набора индексов $1 \le i_1 < i_2 < \cdots < i_k \le n$ выполнено

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \cdots \cdot P(A_{i_k}).$$

События называются *попарно независимыми*, если любые два из них независимы.

Из совместной независимости событий следует их попарная независимость. Но в общем случае из попарной независимости событий не следует совместная независимость.