

Случайные события. Вероятность события. Условная вероятность. Формула Байеса

Теория вероятностей
и математическая статистика / Урок 1



Случайные события

Случайное событие — любой исход эксперимента, который может произойти или не произойти.

Случайное событие — любой исход эксперимента, который может произойти или не произойти.

Например,

- ❶ При броске двух игральных костей на одной выпало число 1, а на другой — 2.
- ❷ Клиент банка не вернул кредит.
- ❸ Температура воздуха в Москве за последние десять дней не превышала 29 градусов по Цельсию.

Пусть A и B — случайные события.

- Сумма событий $A + B$ соответствует наступлению хотя бы одного из событий A и B . Такое событие иногда называют объединением.
- Произведение $A \cdot B$ соответствует наступлению событий A и B одновременно. Такое событие ещё называется совместным.
- Отрицание \bar{A} соответствует тому, что событие A не наступило. Такое событие также называется дополнением.

Событие называется **достоверным**, если в результате испытания оно обязательно произойдёт.

Событие называется **достоверным**, если в результате испытания оно обязательно произойдёт.

Например,

- ❶ При броске игральной кости выпало число, не превышающее 6.
- ❷ Подбросили монету, и выпал либо орёл, либо решка.
- ❸ Монету подбросили стократно, и решка выпала не более 100 раз.

Невозможным событием мы называем событие, которое никогда не произойдёт.

Невозможным событием мы называем событие, которое никогда не произойдёт.

Например,

- ① Две игральные кости бросили один раз, и сумма выпавших чисел составила 15.
- ② Монету подбросили стократно, и решка выпала 55 раз, а орёл — 56.

Совместными называются события, которые могут произойти вместе.

Соответственно, несовместными называются события, которые вместе случиться не могут.

Совместными называются события, которые могут произойти вместе. Соответственно, **несовместными** называются события, которые вместе случиться не могут.

Например,

- 1 При броске монеты не могут одновременно выпасть орёл и решка.
- 2 При броске дротика в круглую мишень можно попасть одновременно в правую половину мишени и в нижнюю половину.

Статистическая вероятность

Относительная частота случайного события — это отношение количества испытаний, в которых данное событие состоялось, к общему числу испытаний:

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

где

- m — число испытаний, в результате которых произошло событие A ,
- n — общее число испытаний.

Как правило, чем больше испытаний мы делаем, тем больше значение частоты «стабилизируется», т.е. приближается к какому-то конкретному значению.

Статистической вероятностью события A называется его относительная частота при достаточно большом («бесконечном») количестве опытов. Вероятность события A обозначается $P(A)$.

Как правило, чем больше испытаний мы делаем, тем больше значение частоты «стабилизируется», т.е. приближается к какому-то конкретному значению.

Статистической вероятностью события A называется его относительная частота при достаточно большом («бесконечном») количестве опытов. Вероятность события A обозначается $P(A)$.

Например, при многократном повторении бросков монеты относительная частота выпадения орла может различаться, однако, вероятность выпадения орла равна 0.5.

- $0 \leq P(A) \leq 1$ для любого события A .
- $P(\emptyset) = 0$, $P(\Omega) = 1$, где \emptyset — невозможное событие, Ω — достоверное событие.
- $P(A + B) = P(A) + P(B) - P(AB)$, где $A + B$ — объединение событий (происходит хотя бы одно), а AB — совместное событие (происходят оба).
- В частности, для несовместных событий: $P(A + B) = P(A) + P(B)$.
- $P(A) + P(\bar{A}) = 1$ для любого события A .

Замечание. Ещё раз подчеркнём, что вероятности произвольных событий суммировать нельзя, сперва необходимо установить несовместность событий.

Комбинаторика

Комбинаторика — раздел математики, в котором изучаются задачи выбора элементов из заданного множества и расположения их в группы по заданным правилам.

Комбинаторика — раздел математики, в котором изучаются задачи выбора элементов из заданного множества и расположения их в группы по заданным правилам.

Мы изучим:

- ① размещения,
- ② перестановки,
- ③ сочетания.

Размещение из n элементов по k элементов — это упорядоченный набор из k элементов, выбранных из множества, содержащего n элементов.

Замечание: здесь k и n — натуральные числа и $0 \leq k \leq n$.

Размещение из n элементов по k элементов — это упорядоченный набор из k элементов, выбранных из множества, содержащего n элементов.

Замечание: здесь k и n — натуральные числа и $0 \leq k \leq n$.

Например, набор $(1, 3, 5)$ является размещением из множества $\{1, 2, 3, 4, 5\}$.

Размещение из n элементов по k элементов — это упорядоченный набор из k элементов, выбранных из множества, содержащего n элементов.

Замечание: здесь k и n — натуральные числа и $0 \leq k \leq n$.

Например, набор $(1, 3, 5)$ является размещением из множества $\{1, 2, 3, 4, 5\}$.

В размещениях важен порядок. Так, $(1, 3, 5)$ и $(5, 1, 3)$ — разные размещения.

Договорённость: будем обозначать круглыми скобками упорядоченные наборы, а фигурными — неупорядоченные.

Посчитаем количество размещений из n по k . Представим себе k пустых ячеек. В первой ячейке может быть любой из n элементов. Во второй ячейке может быть что угодно кроме элемента из первой ячейки, т.е. всего $n - 1$ элементов. В третьей ячейке, аналогично, может быть любой из $n - 2$ элементов, и т.д.

Теперь чтобы получить число всевозможных размещений, нужно перемножить все эти числа. Итак, количество размещений из n по k :

$$A_n^k = n \cdot (n - 1) \cdot (n - 2) \cdots (n - k + 1) = \frac{n!}{(n - k)!},$$

где $n! = n \cdot (n - 1) \cdot (n - 2) \cdots 2 \cdot 1$ — факториал.

Перестановкой из n элементов называется размещение из n элементов по n элементов.

Чтобы посчитать количество перестановок, достаточно знать, что $0! = 1$. Итак, количество перестановок из n элементов:

$$P_n = \frac{n!}{0!} = n!$$

Сочетание из n элементов по k элементов — это неупорядоченный набор из k элементов, выбранных из множества, содержащего n элементов.

Например, набор $\{1, 3, 5\}$ является сочетанием из множества $\{1, 2, 3, 4, 5\}$. При этом, $\{1, 3, 5\}$ и $\{5, 1, 3\}$ — одно и то же сочетание.

Сочетание из n элементов по k элементов — это неупорядоченный набор из k элементов, выбранных из множества, содержащего n элементов.

Например, набор $\{1, 3, 5\}$ является сочетанием из множества $\{1, 2, 3, 4, 5\}$. При этом, $\{1, 3, 5\}$ и $\{5, 1, 3\}$ — одно и то же сочетание.

Сочетаний из n по k меньше, чем размещений. Насколько меньше? Из каждого сочетания размера k можно получить ровно $k!$ различных размещений (переставляя элементы из сочетания всевозможными способами). Итак, число сочетаний из n по k :

$$C_n^k = \frac{A_n^k}{P_k} = \frac{n!}{k!(n-k)!}$$

Классическое определение вероятности

Сформулируем классическое определение вероятности.

Предположим, проводится опыт с n возможными исходами, причём все эти исходы *равновозможны* и *несовместны*. Такие исходы называются элементарными событиями.

Например,

- 1 Игральный кубик бросается однажды. Его выпадение каждой из 6 сторон — все элементарные события.
- 2 Кубик бросается дважды. Элементарные события — все пары его значений.

Рассмотрим событие A , которое можно «собрать» из элементарных событий (т.е. указать, какие элементарные события повлекут за собой событие A , а какие — нет).

Например, выпадение кубика стороной, значение которой не превышает 3, включает в себя три элементарных события.

В этом случае вероятность события A :

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

Здесь n — общее число исходов, а m — число исходов, которые влекут за собой событие A .

Условная вероятность.
Независимые события

Наступление одного события может влиять на наступление другого. Например, вероятность того, что за день хоть раз выпадет снег, выше зимой.

Условная вероятность $P(A|B)$ — это вероятность наступления события A при условии, что произошло событие B :

$$P(A|B) = \frac{P(AB)}{P(B)}$$

Замечание. Такое определение интуитивно напоминает классическое определение вероятности, данное выше: «доля» вероятности совместного события AB относительно вероятности события B .

События A и B называются **независимыми**, если $P(A|B) = P(A)$, т.е. если наступление события B не влияет на вероятность события A , и наоборот. Например, при многократном броске кубика результаты одного броска никак не влияют на результаты других бросков. Напротив, события «зима» и «снег» из примера выше, являются зависимыми.

Замечание. Зависимость событий не означает, что одно гарантированно влечёт другое. Она лишь означает, что наступление одного **меняет вероятность** наступления другого.

Для независимых событий:

$$P(AB) = P(A) \cdot P(B)$$

Замечание. Аналогично суммированию, вероятности произвольных событий перемножать нельзя, сперва необходимо установить независимость событий.

Формула полной вероятности.
Формула Байеса

Говорят, что события H_1, H_2, \dots, H_n образуют полную группу событий, если они несовместны, и в ходе любого испытания одно из этих событий обязательно произойдёт.

Формула полной вероятности для таких событий и произвольного события A :

$$P(A) = P(A|H_1) \cdot P(H_1) + P(A|H_2) \cdot P(H_2) + \dots + P(A|H_n) \cdot P(H_n).$$

В частности, для произвольных событий A и B :

$$P(A) = P(A|B) \cdot P(B) + P(A|\overline{B}) \cdot P(\overline{B})$$

Формула Байеса позволяет «развернуть» условную вероятность $P(A|B)$, т.е. выразить её через $P(B|A)$. По определению условной вероятности:

$$P(A|B) = \frac{P(AB)}{P(B)}$$


Совместную вероятность $P(AB)$ можно теперь выразить в обратном порядке:
 $P(AB) = P(B|A) \cdot P(A)$

Итак, формула Байеса:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Формула Байеса является основой для одной из моделей машинного обучения, решающих задачу классификации — Наивного Байесовского классификатора (Naive Bayes classifier).

Как работает эта модель — читайте в дополнительных материалах к уроку.



Спасибо за внимание