

Элементарная теория вероятности. Формула Байеса

Руслан Назирович Мокаев

Математико-механический факультет,
Санкт-Петербургский государственный университет

Санкт-Петербург, 21.11.2023

Закон полной вероятности. Формула Байеса

Утверждение: $Pr((A \cap B)|C) = Pr(A|(B \cap C)) \cdot Pr(B|C)$.

Доказательство: По определению условной вероятности.

Определение: Множество событий $\{B_i\}_{i=1}^n$ – **полная система событий**, если оно не более чем счётно, события попарно несовместны и их объединение образует пространство элементарных событий ($Pr(B_i) > 0, B_1 \cup \dots \cup B_n = S$)

Закон полной вероятности:

$A \subset S, \{B_i\}_{i=1}^n$ – полная система событий.

Тогда $Pr(A) = \sum_{i=1}^n Pr(A | B_i) \cdot Pr(B_i)$.

Доказательство: По определению условной вероятности.

Формула Байеса: $Pr(B|A) \cdot Pr(A) = Pr(A \cap B) = Pr(A|B) \cdot Pr(B)$.

Доказательство: По определению условной вероятности.

Утверждение:

$Pr(A_1 \cap \dots \cap A_n) = Pr(A_1|(A_2 \cap \dots \cap A_n)) \cdot \dots \cdot Pr(A_{n-1}|A_n) \cdot Pr(A_n)$.

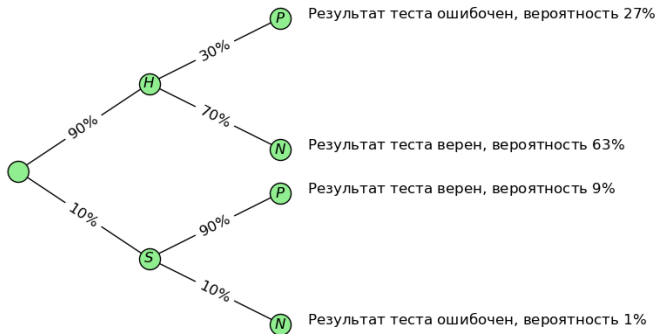
Доказательство: По индукции.

Задача о медицинском обследовании

Пусть существует некая болезнь, которой подвержены 10% населения. Есть тест на выявление этой болезни, который:

- ▶ Имеет 10% вероятность ложного отрицания (вероятность отрицательного (N) результата теста, если человек болен (S));
- ▶ Имеет 30% вероятность ложного подтверждения (вероятность положительного (P) результата теста, если человек здоров (H)).

Построим дерево вариантов.



Задача о медицинском обследовании

Найдём несколько вероятностей:

1. $Pr\{\text{Тест положительный}\} = Pr(\{HP, SP\}) = 9\% + 27\% = 36\%.$
2. $Pr\{\text{Человек болен, если тест положительный}\} =$
 $= Pr(\{SP, SN\}|\{HP, SP\}) = \frac{Pr(\{SP\})}{Pr(\{HP, SP\})} = \frac{9\%}{36\%} = 25\%.$
3. $Pr\{\text{Тест работает верно}\} = Pr(\{HN, SP\}) = 63\% + 9\% = 72\%.$

Кажется, что тест не очень эффективен. Но так ли это?

Рассмотрим инновационный тест, который всегда выдает отрицательный результат.

Для такого теста:

$$Pr\{\text{Тест работает верно}\} = Pr(\{HN, SP\}) = 90\% + 0\% = 90\%.$$

Мораль: нужно аккуратно работать с вероятностями!

Задача об угадывании чисел в конвертах

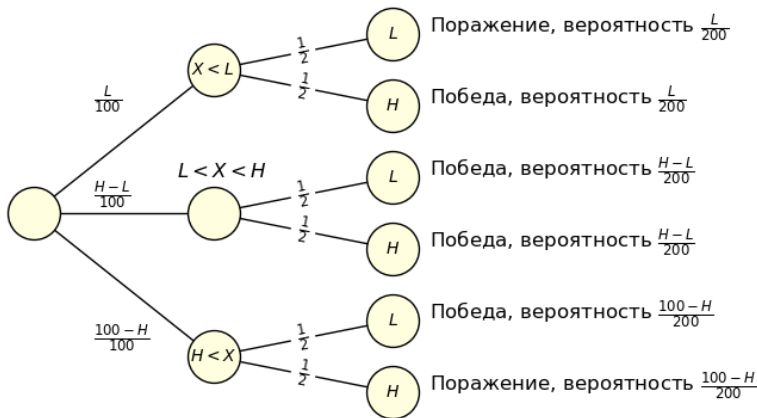
Пусть есть два различных числа от 0 до 100 в закрытых конвертах. Вам предложено, открыв один из конвертов (наугад, то есть, конверты ничем не отличаются), попытаться угадать, открыли вы конверт с большим из двух чисел (H) или меньшим (L). Необходимо научиться угадывать с вероятностью хотя бы чуть-чуть лучшей, чем 50%.

Давайте попробуем найти такое число X , что $L < X < H$. По имеющемуся X и числу из открытого конверта можно очень легко понять, какое из двух чисел меньше (если число из открытого конверта меньше X , то это L , иначе – H).

Чтобы избежать ситуаций, когда X совпадает с L или H , будем выбирать X из множества $A = \{\frac{1}{2}, \dots, 99\frac{1}{2}\}$. Вопрос в том, как выбрать хорошее число X ?

Задача об угадывании чисел в конвертах

Попробуем наугад! Нарисуем дерево вариантов.



Итак, посчитаем вероятность победы с использованием данной стратегии.

$$\begin{aligned} Pr\{\text{победить}\} &= \frac{L}{200} + \frac{H-L}{200} + \frac{H-L}{200} + \frac{100-H}{200} = \frac{100-H+H-L+L}{200} + \frac{H-L}{200} = \\ &= \frac{1}{2} + \frac{H-L}{200}. \end{aligned}$$