

# Теория Вероятностей и Математическая Статистика

## ФИИТ, 2 курс, 4 семестр

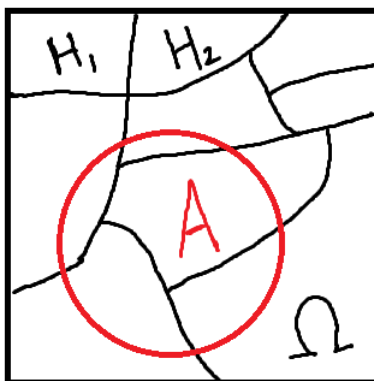
### Лекция 3

28 февраля 2021 г.

## 1 Формула полной вероятности

Наша основная тема на сегодня – формула полной вероятности. Мы будем рассматривать ситуации, когда опыт производится сложным образом и содержит слишком много составляющих.

Итак, допустим мы имеем изначально множество исходов опыта  $\Omega$ . Это множество можно разделить на некоторые части ( $H_1, H_2, \dots$ ), а мы хотим рассмотреть некоторое событие  $A$ , которое так же является подмножеством  $\Omega$  (по определению).



Мы можем составить набор событий (которые мы будем называть **гипотезы**):  $H_1, H_2, \dots, H_n, \dots$ . Этот набор удовлетворяет следующим условиям:

1. События попарно:  $H_i H_j = \emptyset \quad (i \neq j)$
2. События образуют полную группу:  $H_1 + H_2 + \dots + H_n + \dots = \Omega$

Таким образом мы получаем набор гипотез – полная группа попарно несовместных событий, которая построена до опыта – априорти. А наше некоторое событие  $A$  можно представить как объединение этих групп:

$$A = A\Omega = A(H_1 + H_2 + \dots) = AH_1 + AH_2 + \dots$$

Очевидно, то слагаемые будут, как и гипотезы, попарно несовместны. Тогда можно перейти к нахождению вероятности через сумму вероятностей:

$$P(A) = P(A\Omega) = P(A(H_1 + H_2 + \dots)) = P(AH_1) + P(AH_2) + \dots$$

А далее к каждому слагаемому применим теорему умножения:

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

Таким образом мы получили **формулу полной вероятности**:

$$P(A) = \sum_{i=1}^{\infty} P(H_i)P(A|H_i)$$

## 1.1 Пример

Три станка изготавливают одинаковые детали. Вероятность брака составляет 0.01, 0.02, 0.03 соответственно. На 1-м станке изготовлено 2000 деталей, на 2-м – 3000, а на 3-м 5000. ОТК берет на контроль одну из всех выпущенных деталей. Найти вероятность того, что эта деталь – годная.

Первое, что стоит отметить, что задача – многонаправленная (и производство есть, и отбор). Если не брать в расчет годная ли деталь, можно отметить, что любая деталь будет изготовлена на одном из станков. Значит можно выделить следующие события:

$$\begin{aligned} H_1 &= \{ \text{взятая деталь изготовлена на 1-м станке} \} \\ H_2 &= \{ \text{взятая деталь изготовлена на 2-м станке} \} \\ H_3 &= \{ \text{взятая деталь изготовлена на 3-м станке} \} \end{aligned}$$

При этом общее количество деталей равно:  $N = 10000$ . А мы можем вычислить вероятности описанных событий (дроби сокращать не будем, так как одинаковые знаменатели могут в будущем упростить жизнь):

$$P(H_1) = \frac{2000}{10000} = \frac{2}{10} \quad P(H_2) = \frac{3000}{10000} = \frac{3}{10} \quad P(H_3) = \frac{5000}{10000} = \frac{5}{10}$$

На что можно обратить внимание: если мы построили события правильно, то они должны составлять полную группу, а значит и сумма вероятностей должна быть равна единице (что у нас работает).

Далее поинтересуемся, с каким качеством изготовлена деталь. Построим следующее событие:

$$A = \{ \text{изготовленная деталь годная} \}$$

Тогда можем построить условные вероятности:

$$P(A|H_1) = 0,99 \quad P(A|H_2) = 0,98 \quad P(A|H_3) = 0,97$$

По условию нам надо найти просто полную вероятность, так что воспользуемся **формулой полной вероятности**:

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) = \frac{2}{10} \cdot 0.99 + \frac{3}{10} \cdot 0.98 + \frac{5}{10} \cdot 0.97 = \dots$$

## 2 Формула Байеса

Допустим, мы проделали следующие шаги:

1. До опыта построены гипотезы
2. Провели опыт (событие  $A$  реализовалось, либо нет) // Для себя примем, что произошло событие  $A$  произошло (если нет, то произошло  $\bar{A}$ )

Возникает вопрос: как изменятся вероятности гипотез после проведения опыта. Рассмотрим следующую вероятность некоторой гипотезы  $H_k$ :

$$P(H_k|A) = \frac{P(H_k A)}{P(A)} = \frac{P(H_k)P(A|H_k)}{\sum_{i=1}^{\infty} P(H_i)P(A|H_i)}, k = 1, 2, \dots$$

Эта запись и называется Формулой Байеса. Она позволяет учесть в опыте произошедшее событие  $A$ .

## 2.1 Пример

Из 30 экзаменационных билетов студент выучил 10. Какова вероятность, что ему попадет нужный билет?

Если он вышел первый ( $A$ ), тогда вероятность достать билет:  $P(A) = \frac{10}{30} = \frac{1}{3}$ .

Если он вышел второй, но все выученные билеты все еще лежат на столе, то получаем условную вероятность:  $P(A|B) = \frac{10}{29}$

...

## 2.2 Пример (продолжение 1.1)

Взятая ОТК деталь годная. Найти вероятность того, что эта деталь изготовлена на 3-м станке.

У нас имеются следующие события:

$H_3 = \{ \text{взятая деталь изготовлена на 3-м станке} \}$

$A = \{ \text{изготовленная деталь годная} \}$

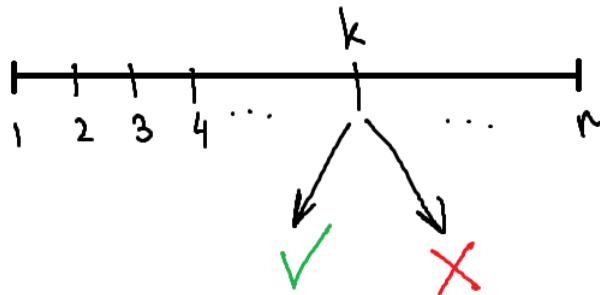
Тогда нам надо найти следующую вероятность, воспользуясь Формулой Байеса:

$$P(H_3|A) = \frac{P(H_3)P(A|H_3)}{P(A)} = \frac{\frac{5}{10} \cdot 0.97}{\frac{2}{10} \cdot 0.99 + \frac{3}{10} \cdot 0.98 + \frac{5}{10} \cdot 0.97} = \frac{5 \cdot 97}{2 \cdot 99 + 3 \cdot 98 + 5 \cdot 97} = \dots$$

## 3 Схема независимых опытов / Схема Бернулли

Одна из самых важных частей Теории Вероятностей – Схема Бернулли.

В одних и тех же условиях один и тот же опыт повторяется  $n$  раз. В каждом из случаев мы имеем один из двух исходов:  $A$  реализовалось, либо  $\bar{A}$  реализовалось.

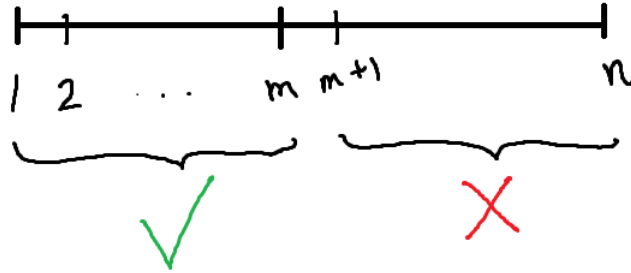


Если событие  $A$  реализовалось, мы говорим, что в одном отдельно взятом опыте произошел успех. Вероятность обозначим следующим образом:  $P(A) = p$ .

Вероятность реализации события  $\bar{A}$  (неуспех) обозначим следующим образом:  $P(\bar{A}) = 1 - p = q$

Нас же будет интересовать следующее: какова вероятность того, что в последовательности из  $n$  опытов "успех" будет зафиксирован  $m$  раз (где  $0 \leq m \leq n$ ).

Рассмотрим следующую ситуацию. Пусть у нас имеется  $m$  успехов, и все они расположены в начале друг за другом: (1)



Тогда для нашей картины найти вероятность достаточно просто в силу независимости опытов:

$$P(A \dots A \cdot \bar{A} \dots \bar{A}) = p^m \cdot q^{n-m}$$

Но в нашем вопросе нас будет интересовать количество успехов, а не место их расположения. То есть требуется выбрать не первые  $m$  мест, а произвольные  $m$  мест. Тогда количество таких способов мы тоже знаем, как подсчитывается:  $C_n^m$  (сочетания из  $n$  по  $m$ ). Таким образом, получаем следующую вероятность " $m$  успехов среди  $n$  опытов":

$$P_n(m) = C_n^m p^m \cdot q^{n-m}$$

Полученное выражение является формулой Бернулли.

### 3.1 Пример

Бросают 3 монеты. Найти вероятность выпадения одного герба.

Имеем событие:

$A = \{ \text{выпал 1 герб} \}$  – один "успех" в серии из 3-х опытов:

$$P_3(1) = C_3^1 \cdot \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^{3-1} = 3 \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{3}{8}$$

### 3.2 Пример

Что вероятнее: выиграть у равносильного партнера 3 партии из 4, или 5 партий из 8. (Без ничьих!).

"успех" 3 из 4 (75%)

"успех" 5 из 8 (62.5%)

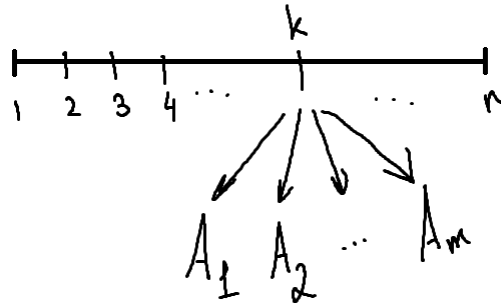
Мы же склоняемся к тому, что более "скромный" успех достигим. Тогда сравним следующие вероятности:  $P_3(4)$  и  $P_8(5)$ . Построим отношение:

$$\frac{P_4(3)}{P_8(5)} = \frac{C_4^3 \cdot (\frac{1}{2})^3 (\frac{1}{2})^{4-3}}{C_8^5 \cdot (\frac{1}{2})^5 (\frac{1}{2})^{8-5}} = \frac{4 \cdot (\frac{1}{2})^4}{\frac{2 \cdot 8 \cdot 7 \cdot 6}{6} \cdot (\frac{1}{2})^8} = \frac{16}{14} = \frac{8}{7} > 1$$

Таким образом, мы явно показали, что "скромный" результат вероятен меньше. Однако, это вполне логично. Ту же ситуацию мы можем посмотреть в другой ситуации: легче добиться успеха в стометровке или 40 километровом марафоне? Тут уже вполне очевидно, что в стометровке просто надо выложиться один раз в короткий промежуток, в отличии от марафона, где даже в легком режиме можно просто не дойти.

## 4 Полиномиальная схема

Это расширение схемы Бернулли. Теперь у нас ситуация выглядит похожим образом: имеем последовательность независимых опытов, и в каком-то опыте  $k$  у нас может произойти событие  $A_1$ , или событие  $A_2$ , или событие  $A_m$ . При этом эти события образуют полную группу попарно несовместных событий.



$$P(A_1) = p_1, \quad P(A_2) = p_2, \quad \dots \quad P(A_m) = p_m$$

$$p_1 + p_2 + \dots + p_m = 1$$

И для разбора таких ситуаций имеем:

$$P_n(n_1, n_2, \dots, n_m) = \frac{n!}{n_1! n_2! \dots n_m!} P_1^{n_1} P_2^{n_2} \dots P_m^{n_m}$$

Эта запись и есть та самая полиномиальная формула.

### 4.1 Пример

На интервале  $(1; 5)$  наугад выбирают 5 точек. Нужно найти вероятность

1. В интервале  $(1; 2)$  и  $(3; 5)$  окажется 2 и 3 точки соответственно
2. В интервале  $(1; 3)$  и  $(2; 5)$  окажется равное количество точек

1) Можно отметить, что вероятности мы рассчитываем геометрически. То есть попасть именно в точку 1 можно с нулевой вероятностью. А попасть в интервал можно с вероятностью отношения длины интервала к всему интегралу. Тогда получается, что вероятность попасть в интервал  $(1; 2)$  равна  $\frac{1}{4}$ , а в интервал  $(3; 5)$  равна  $\frac{2}{4}$ .

Однако наши вероятности не образуют полную группу (сумма вероятностей не равна 1). Можем решить эту проблему добавлением интервала  $(2; 3)$ , для которого вероятность будет  $\frac{1}{4}$ . Так мы образуем полную группу.

Нас интересует следующее распределение точек:

$$P_5(2, 0, 3) = \frac{5!}{2!0!3!} \left(\frac{1}{4}\right)^2 \left(\frac{1}{4}\right)^0 \left(\frac{2}{4}\right)^3 = \frac{5 \cdot 4}{2} \cdot \frac{2^3}{4^5} = \dots$$

2) Для этого пункта у нас ситуация немного другая. Интервалы получились (1; 3) и (2; 5). И длина у них разная, количество точек должно быть равное.

Однако, если учесть, что колво точек всего у нас нечетное (5), а интервалы пересекаются, то в пересечении может быть 1 точка, 3 точки или 5 точек. Таким образом, получаем следующие случаи:

$$P_5(2, 1, 2) \quad P_5(1, 3, 1) \quad P_5(0, 5, 0)$$

Таким образом, надо найти сумму вышеприведенных вероятностей:

$$P_5(2, 1, 2) + P_5(1, 3, 1) + P_5(0, 5, 0) = \left[ \frac{5!}{2!1!2!} \left(\frac{1}{4}\right)^2 \left(\frac{1}{4}\right)^1 \left(\frac{2}{4}\right)^2 \right] + \left[ \frac{5!}{1!3!1!} \left(\frac{1}{4}\right)^1 \left(\frac{1}{4}\right)^3 \left(\frac{2}{4}\right)^1 \right] + \left[ \frac{5!}{0!5!0!} \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^5 \left(\frac{2}{4}\right)^0 \right] = \dots$$

## 5 Пример (для формулы полной вероятности)

Имеется 2 урны: первая содержит 2 белых и 3 черных шара, а вторая 3 белых и 4 черных.

Ситуация следующая: наугад выбрана урна из нее наугад взят шар, оказавшийся белым. Нужно найти вероятность того, что второй шар взятый из той же урны окажется так же белым.

Самое существенное в данном условии то, что взятый шар оказался белым. Это означает, что в нашем опыте есть событие, которое произошло, поэтому вероятность, которую мы ищем, точно условная.

Посмотрим, какие события мы можем здесь рассмотреть:

$$\begin{aligned} H_1 &= \{ \text{шар взят из первой урны} \} \\ H_2 &= \{ \text{шар взят из второй урны} \} \\ P(H_1) &= P(H_2) = \frac{1}{2} \text{ (до проведения опыта)} \end{aligned}$$

Также имеем:

$$\begin{aligned} A &= \{ \text{первый шар белый} \} \\ B &= \{ \text{второй шар белый} \} \end{aligned}$$

**Примечание.** Как всегда мы сначала описываем задачу. Нахождением, вычислениями и остальные размышления мы делаем на следующем этапе.

Требуемая вероятность:

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

По формуле полной вероятности получим:

$$= \frac{P(H_1)P(AB|H_1) + P(H_2)P(AB|H_2)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{\frac{1}{2} \cdot \frac{2}{5} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{7} \cdot \frac{2}{6}}{\frac{1}{2} \cdot \frac{2}{5} + \frac{1}{2} \cdot \frac{3}{7}} = \dots$$

Вообще говоря, можно было бы применить формулу полной вероятности напрямую, но это было бы достаточно сложно вычислять, так как полученные вероятности были бы условными, в том числе условные вероятности гипотез:

$$P(B|A) = P(H_1|A) \cdot P(B|AH_1) + P(H_2|A) \cdot P(B|AH_2)$$