Расчет вероятности случайных событий

Виды случайных событий. Комбинаторика.Расчет вероятности. Формула Байеса

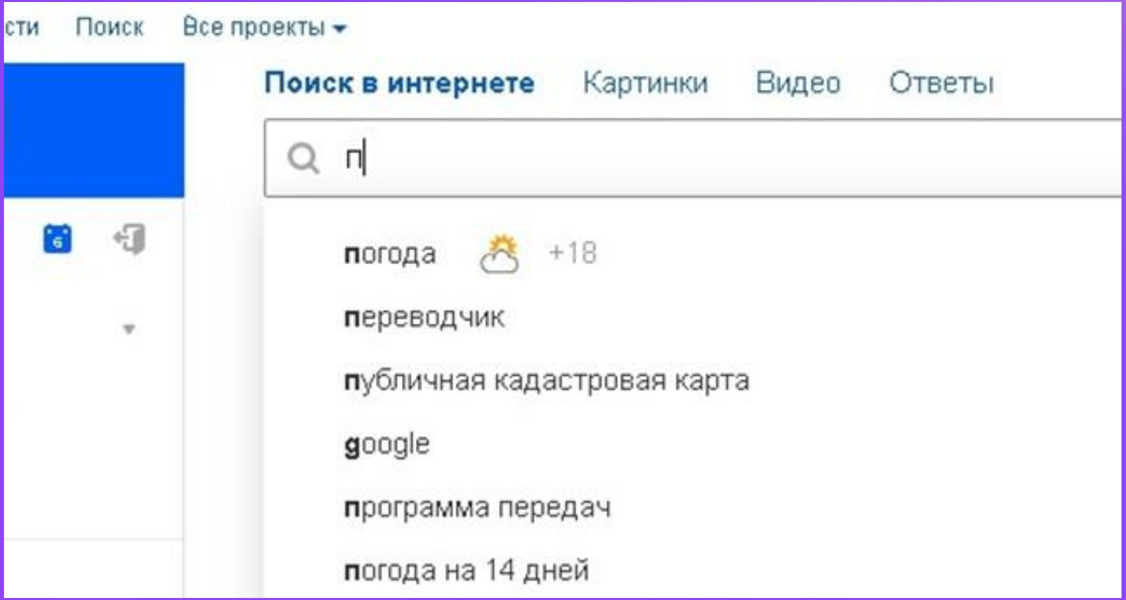
- ✓ Что такое случайное событие
- ✓ Статистическая вероятность
- ✓ Классическое определение вероятности
- ✓ Формулы комбинаторики
- ✓ Виды случайных событий
- ✓ Условная вероятность
- ✓ Формула полной вероятности
- ✓ Формула Байеса

План курса

1. Расчет вероятности случайных событий
2. Дискретные распределения вероятностей
3. Разведочный анализ (EDA)
4. Непрерывная случайная величина
5. Тестирование гипотез
6. Сравнение долей. Построение доверительного интервала
7. Непараметрические тесты
8. Корреляционный анализ
9. Регрессионный анализ
10. Дисперсионный анализ
11. Будущий урок
12. Будущий урок
13. Прошедший урок
14. Сегодняшний урок
15. Будущий урок
16. Будущий урок

Что будет на уроке сегодня

- Обсудим, зачем проходить этот курс?
- Научимся различать случайные события по видам
- Научимся применять комбинаторику для расчета вероятностей
- Изучим условную вероятность и способы ее расчета

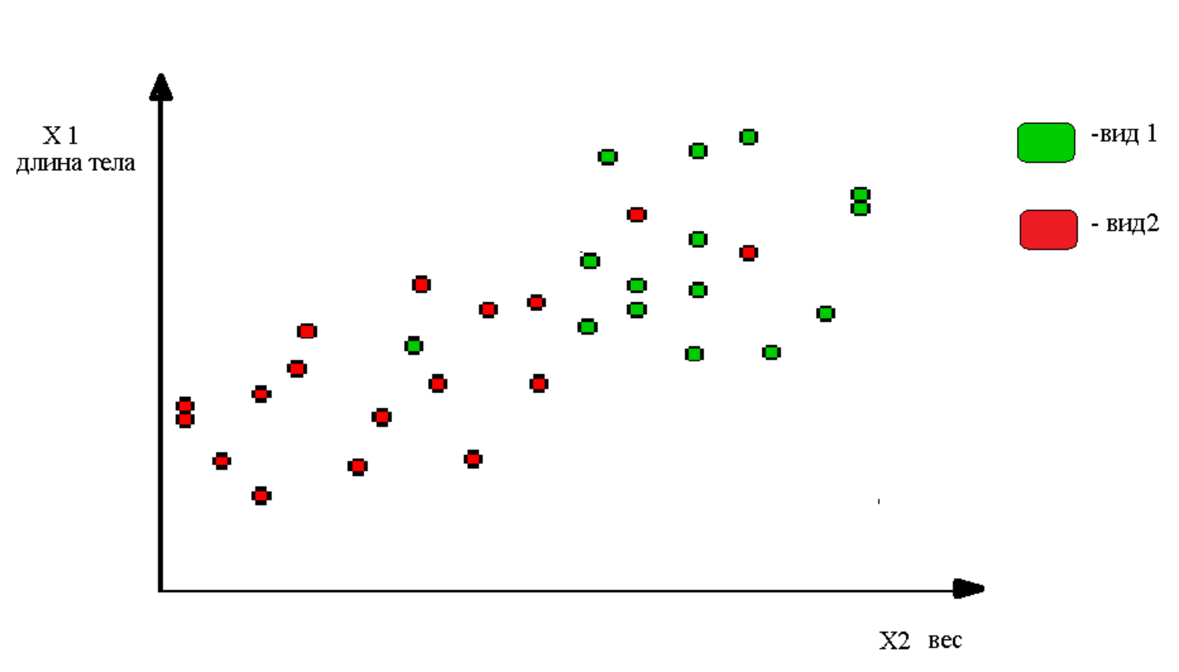


Система автозаполнения

Точность экспресс теста на ВИЧ :

Чувствительность > 99% P(+| болен)

Специфичность > 99% P(- |здоров)



Создание математических моделей и их оценка

Случайное событие- это событие, которое при определенных условиях может произойти либо не произойти.

Примеры случайного события

Пример 1

Средняя температура по Москве за последние 10 дней составила 25˚C

Пример 2

Клиент банка не вернул кредит

Пример 3

При стократном подбрасывании монеты решка выпала 55 раз

Достоверные случайные события

Событие можно назвать достоверным, если в результате испытания оно обязательно произойдет

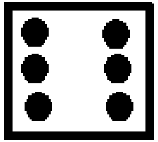
Невозможные случайные события

Событие называется невозможным, если в результате испытания оно никогда не произойдет

Примеры достоверных случайных событий

Пример 1

При броске игральной кости выпало число, не превышающее 6



Кость бы покрасивее ☺)

Пример 2

При подбрасывании монеты выпадает либо орел, либо решка

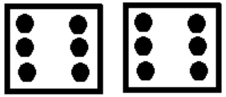
Пример 3

При стократном подбрасывании монеты решка выпала не более 100 раз

Примеры невозможных случайных событий

Пример 1

При однократном подбрасывании двух игральных костей сумма выпавших чисел составила 15

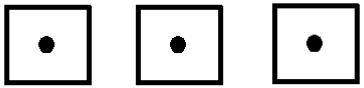


Пример 2

При стократном подбрасывании монеты решка выпала 55 раз, а орел – 56

Пример 3

При однократном подбрасывании трех игральных костей сумма выпавших чисел составила 2



Вероятность

Относительная частота

Экспериментальная характеристика

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

Статистическая вероятность

Экспериментальная характеристика

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

Классическое определение вероятности

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

W(A) - относительная частота события A,

m - число появления события A,

n - общее число испытаний.

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

```
import numpy as np
np.random.seed(1)
n= 60
b= np. random.randint (1, 7, n)
b
array([6, 4, 5, 1, 2, 4, 6, 1, 1, 2, 5, 6, 5, 2, 3, 5, 6, 3, 5, 4, 5, 3,
       5, 6, 3, 5, 2, 2, 1, 6, 2, 2, 6, 2, 2, 1, 5, 2, 1, 1, 6, 4, 3, 2,
       1, 4, 6, 2, 2, 4, 5, 1, 2, 4, 5, 3, 5, 1, 6, 4])
a = b [b == 3]
a
array([3, 3, 3, 3, 3, 3])
m = len ( a )
m
6
Теперь можем вычислить относительную частоту события А

W = m/n
W
0.1
```

Смоделируем ситуацию, когда бросают две игральные кости одновременно 360 раз. Посчитаем относительную частоту события, когда на одной кости выпадает 1 , а на другой 2

```
n = 360
np.random.seed(1)
c = np.random.randint (1, 7, size = n)
d = np.random.randint (1, 7, size = n)
c
array([6, 4, 5, 1, 2, 4, 6, 1, 1, 2, 5, 6, 5, 2, 3, 5, 6, 3, 5, 4, 5, 3,
       5, 6, 3, 5, 2, 2, 1, 6, 2, 2, 6, 2, 2, 1, 5, 2, 1, 1, 6, 4, 3, 2,
       1, 4, 6, 2, 2, 4, 5, 1, 2, 4, 5, 3, 5, 1, 6, 4, 2, 3, 1, 5, 2, 3,
       3, 2, 1, 2, 4, 6, 5, 4, 6, 2, 4, 1, 1, 3, 3, 2, 4, 5, 3, 1, 1, 2,
       2, 6, 4, 1, 1, 6, 6, 5, 6, 3, 5, 4, 6, 4, 6, 1, 4, 5, 4, 5, 5, 6,
       5, 2, 1, 5, 3, 1, 6, 3, 5, 2, 2, 1, 3, 5, 5, 1, 5, 2, 5, 2, 1, 3,
       4, 2, 3, 5, 5, 3, 3, 1, 2, 3, 3, 1, 6, 2, 3, 5, 1, 6, 6, 2, 3, 2,
       6, 5, 3, 1, 6, 1, 2, 6, 1, 2, 4, 2, 2, 6, 5, 5, 4, 6, 1, 4, 1, 4,
       2, 3, 6, 6, 5, 6, 1, 6, 1, 6, 4, 2, 2, 6, 1, 6, 1, 5, 3, 6, 4, 5,
       3, 1, 6, 4, 4, 6, 6, 2, 3, 5, 4, 1, 1, 6, 5, 3, 5, 3, 1, 6, 4, 1,
       1, 5, 6, 3, 2, 1, 5, 4, 1, 2, 3, 5, 5, 4, 4, 4, 4, 3, 4, 4, 6, 5,
       4, 3, 5, 5, 1, 4, 4, 1, 4, 6, 6, 2, 1, 3, 3, 3, 1, 3, 2, 5, 1, 5,
       5, 2, 4, 2, 5, 2, 3, 6, 2, 1, 1, 3, 5, 2, 1, 1, 4, 2, 1, 5, 4, 3,
       4, 5, 6, 5, 4, 1, 1, 1, 5, 6, 2, 6, 5, 2, 3, 6, 3, 6, 5, 6, 4, 5,
       5, 1, 4, 3, 5, 6, 4, 5, 3, 4, 1, 6, 6, 3, 2, 4, 3, 1, 6, 2, 5, 2,
       4, 4, 2, 3, 1, 3, 5, 1, 3, 5, 4, 5, 4, 1, 6, 6, 5, 3, 3, 5, 2, 3,
       6, 2, 2, 2, 1, 6, 6, 5])
```

```
d
array([5, 3, 3, 4, 2, 5, 1, 1, 4, 3, 5, 2, 4, 2, 2, 3, 6, 6, 5, 6, 1, 4,
       1, 5, 3, 4, 6, 2, 2, 5, 5, 1, 3, 2, 4, 1, 2, 1, 6, 3, 3, 5, 4, 3,
       3, 3, 1, 6, 3, 1, 5, 6, 2, 6, 1, 3, 4, 1, 5, 4, 4, 4, 1, 4, 2, 3,
       1, 2, 5, 3, 4, 5, 6, 5, 3, 2, 3, 6, 1, 6, 4, 6, 4, 3, 1, 1, 1, 1,
       3, 5, 1, 5, 2, 3, 2, 3, 5, 2, 4, 6, 2, 6, 2, 3, 5, 2, 1, 3, 6, 2,
       3, 1, 1, 6, 4, 5, 2, 6, 1, 5, 1, 4, 3, 5, 4, 3, 5, 3, 5, 1, 1, 6,
       5, 3, 3, 5, 3, 4, 6, 1, 1, 5, 4, 5, 4, 4, 5, 1, 6, 4, 6, 2, 5, 5,
       4, 3, 3, 3, 3, 3, 1, 3, 2, 6, 3, 4, 6, 1, 1, 6, 6, 6, 2, 2, 4, 4,
       6, 4, 2, 4, 4, 4, 2, 4, 1, 5, 6, 1, 6, 6, 3, 5, 5, 3, 1, 4, 6, 3,
       5, 6, 1, 5, 6, 3, 4, 5, 3, 5, 2, 4, 5, 4, 1, 4, 1, 5, 4, 1, 6, 4,
       2, 5, 5, 6, 3, 6, 6, 3, 5, 3, 2, 3, 4, 2, 6, 6, 4, 4, 1, 5, 4, 4,
       6, 4, 4, 1, 3, 4, 6, 2, 6, 4, 3, 6, 4, 3, 6, 3, 1, 5, 1, 2, 4, 1,
       1, 1, 2, 3, 5, 5, 3, 1, 2, 1, 1, 6, 3, 6, 5, 4, 1, 3, 2, 4, 4, 2,
       5, 5, 5, 6, 1, 2, 1, 2, 3, 3, 5, 6, 6, 3, 5, 6, 1, 6, 3, 6, 5, 5,
       2, 2, 2, 2, 6, 5, 6, 4, 5, 2, 2, 6, 2, 3, 5, 4, 3, 4, 4, 3, 4, 1,
       3, 1, 2, 6, 1, 6, 1, 4, 5, 1, 1, 6, 1, 1, 2, 1, 6, 2, 4, 4, 5, 6,
       1, 2, 6, 1, 4, 3, 2, 2])
a = c[( c == 1) & (d == 2)]
a
array([1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1])
m = len ( a )
m
12
W = m / n
W
0.03333333333333333
```

При достаточно большом числе испытаний за статистическую вероятность принимают относительную частоту

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

При условии, что заранее известны все вероятные исходы и они равновозможные можно воспользоваться классическим определением вероятности:

Вероятность события — это отношение числа элементарных исходов, благоприятствующих данному событию, к числу всех равновозможных исходов опыта, в котором оно может появиться

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

Комбинаторика

1

Сочетания

Порядок не важен

2

Размещение

Порядок важен

Участвуют не все элементы

3

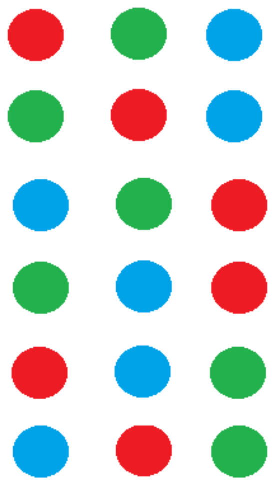
Перестановка

Прядок важен

Усаствуют все элементы

Сочетание - набор, состоящий из k элементов, выбранных из множества, содержащего n различных элементов.

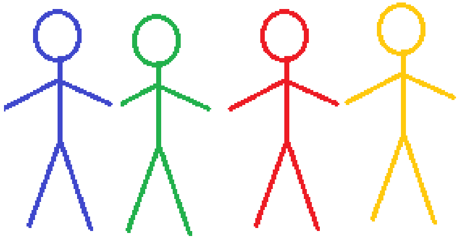
Перестановки - комбинации из n элементов, отличающиеся их порядком.

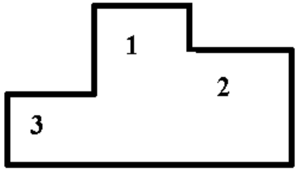


Размещения из k элементов,

выбранных из множества n –

это такие комбинации, которые отличаются либо самими элементами, либо порядком их расположения.





Формулы комбинаторики

1

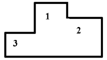
Сочетание

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

2

Размещение

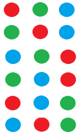
$$A_n^k = \frac{n!}{(n-k)!}$$



3

Перестановка

$$P_n = n!$$



Определить сочетания, размещения или перестановки используются для решения этой задачи.

Сколькими способами можно выбрать из колоды, состоящей из 36 карт, 4 карты?

```
from math import factorial

def combinations (n,k):
  return np.math.factorial(n) // (np.math.factorial(k) * np.math.factorial(n - k))

combinations ( 36, 4)
58905
```

$$C_{36}^{4} = \frac{36!}{4!(36-4)!} = \frac{36!}{4! \cdot 32!} = \frac{33 \cdot 34 \cdot 35 \cdot 36}{4!} = \frac{33 \cdot 34 \cdot 35 \cdot 36}{2 \cdot 3 \cdot 4} = 58905$$

Определить сочетания, размещения или перестановки используются для решения этой задачи.

В магазине 20 покупателей. Сколькими способами они могут образовать очередь из 5 человек?

```
def arrangements ( n, k ):
  return np.math.factorial ( n ) // np.math.factorial ( n - k )

arrangements(20, 5)
1860480
```

Определить сочетания, размещения или перестановки используются для решения этой задачи.

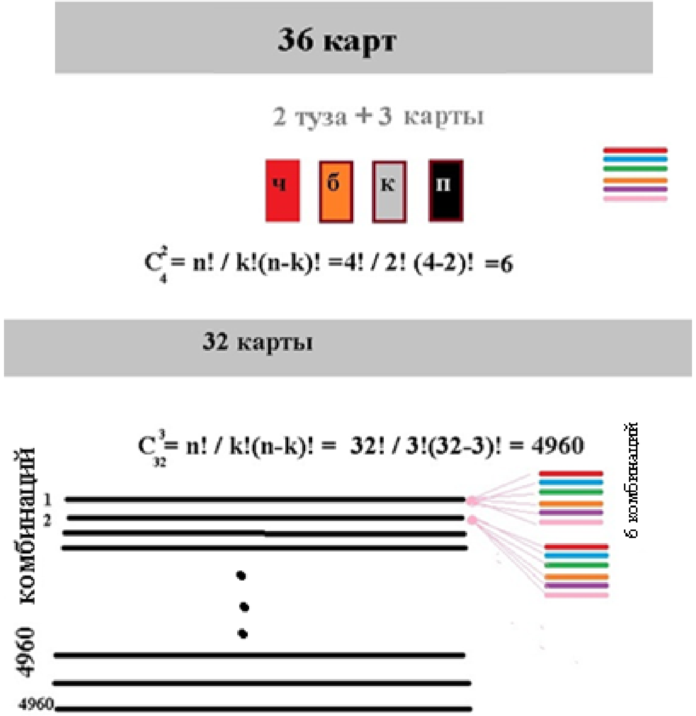
Сколькими способами 5 покупателей могут образовать очередь?

```
def permutations ( n ):
      return np.math.factorial(n)

 permutations ( 5 )
120
```

Расчет возможных комбинаций

Из колоды, состоящей из 36 карт, случайным образом выбраны 5. Сколькими способами можно выбрать эти карты так, чтобы среди них оказалось 2 туза?



1) Берем подмножество тузов

$$\frac{4!}{2!(4-2)!} = 6$$ сочетаний из 2 тузов

2) Берем подмножество без тузов

$$\frac{32!}{3!(32-3)!} = 4960$$ сочений из 3 нетузов

C = 4960 * 6 =29 760 сочетаний из 5 карт, где 2 туза

Расчет вероятности с помощью комбинаторики

Из колоды 36 карт случайным образом берут 5 карт. Найти вероятность того, среди 5 карт будет 2 туза.

$$P = \frac{\text{число исходов, благоприятствующих событию}}{\text{общее число исходов}}$$

Число благоприятствующих исходов $C = 29\ 760$

$$C_{36}^{5} = \frac{36!}{5! * 31!} = \frac{\cancel{31!} * 32 * 33 * 34 * 35 * 36}{120 * \cancel{31!}} = 376\ 992$$

P= 29760 / 376992 = 0,0789

Ответ: вероятность взять 5 карт с 2 тузами 0,0789 или около 8%

Зависимые и независимые события

Независимые события

появление одного не влияет на появление другого.

Вероятность **одновременного** появления двух независимых событий :

$$P(AB) = P(A) * P(B)$$

Зависимые события

появление одного влияет на появление другого.

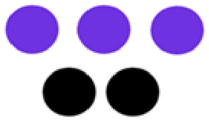
Вероятность наступления двух зависимых событий:

$$P(A * B) = P(A) * P(B|A)$$

Выражение P(B|A) означает вероятность наступления события B при том, что событие A уже наступило

Зависимые события.

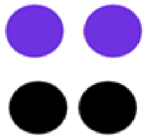
В ящике лежат 3 фиолетовых и 2 черных шара. Найти вероятность, что подряд вытащат фиолетовый и черный шар.



Вероятность вытащить первый фиолетовый шар:

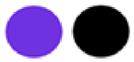
$$P(A) = \frac{3}{5}$$

Событие «вытащить» черный мяч имеет условие.
Оно наступает после события «вытащить фиолетовый шар»



$$P(B|A) = \frac{2}{4} = \frac{1}{2}$$

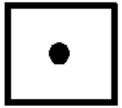
Вероятность что подряд вытащат фиолетовый и черный шар



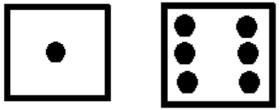
$$P(A * B) = P(A) * P(B|A) = \frac{3}{5} * \frac{1}{2} = \frac{3}{10}$$

Совместные и несовместные события

События называются несовместными ,если появление
одного из них исключает появление другого



События называются совместными, если появление одного из них не исключает появления другого



Произведение событий. События A и B наступают одновременно.

Несовместные события

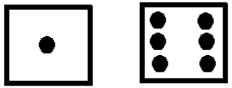
Событие A – выпадает 2
Событие B – выпадает 1

$$P\ (C)\ =\ P(A)\ *\ P(B)\ =\ 0$$

Совместные события

Событие A – выпадает 2
Событие B – выпадает 1

$$P\ (C) = \ P(A) * \ P(B) = \frac{1}{6} * \frac{1}{6} = \frac{1}{36}$$



Сумма событий. Строго или событие A, или событие B

$$P\ (C) = \ P(A) + \ P(B) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$P\ (C) = \ P(A) + \ P(B)\ - P(AB) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$$

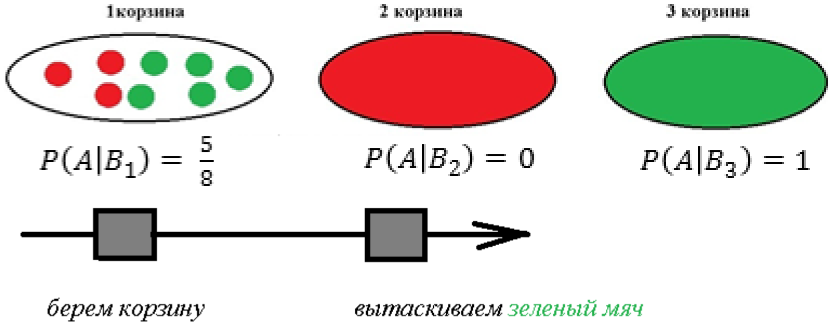
Если событие A может наступить только при наступлении событий B_1, B_2, …, B_n, образующих полную группу событий *, то вероятность события A вычисляется по формуле:

$$P(A) = P(B_1) \cdot P(A \mid B_1) + P(B_2) \cdot P(A \mid B_2) + \ldots + P(B_n) \cdot P(A \mid B_n)$$

* полная группа событий означает, что при отдельном испытании обязательно произойдет одно из них

Полная вероятность

Случайным образом выбирается корзина и случайным образом из нее берут мяч
Какова вероятность, что мяч окажется зеленым?



A - вытащить зеленый мяч

B_1 – взять 1-ю корзину
B_2 – взять 2-ю корзину
B_3 – взять 3-ю корзину

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$$

$$\mathrm{P(A)} = \mathrm{P}(B_1) * P(A|B_1) + P(B_2) * P(A|B_2) + P(B_3) * P(A|B_3) = \frac{1}{3} * \frac{5}{8} + \frac{1}{3} * 0 + \frac{1}{3} * 1 = \frac{13}{24}$$

Чтобы определить вероятность события B при условии, что событие A уже произошло, используют формулу Байеса.

Здесь P(B) – априорная вероятность (определяется до испытания)

P(B|A) - апостериорная вероятность

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

Формула Байеса

Случайным образом выбирается корзина и случайным образом из нее берут зеленый мяч. Какова вероятность, что зеленый мяч окажется из 3-ей корзины?

A – вытащить зеленый мяч

B_1, B_2, B_3 – взять 1, 2, 3–ю корзину соответственно

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$$

$$P(A|B_1) = \frac{5}{8} \qquad P(A|B_2) = 0 \qquad P(A|B_3) = 1$$

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

$$P\ (B_3|A) = \frac{P(A|B_3) * P(B_3)}{P(A)} = \frac{1 * \frac{1}{3}}{\frac{13}{24}} = \frac{24}{39} = \frac{8}{13}$$

Формула Байеса

Пациент приходит к врачу, чтобы сделать тест на конкретное заболевание.
Вероятность положительного теста при том, что болезнь есть 90%.
Вероятность отрицательного теста, при том, что человек здоров 95 %.
Вероятность заболевания 2%.
Используя теорему Байеса, найти вероятность, что человек болен, при условии , что тест положительный?

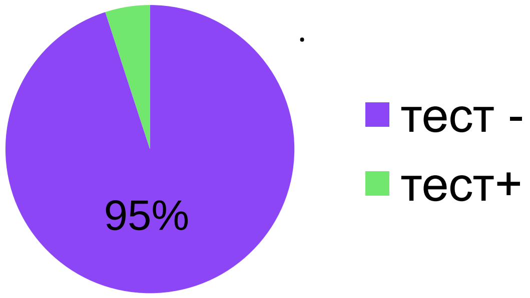
Определимся:

- ✓ Что будет являться событием А
- ✓ Что будет являться событием В

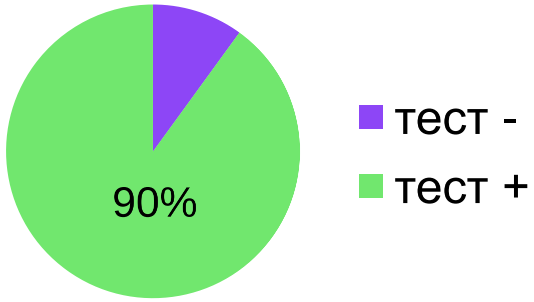
$$P(B) = ?$$
$$P(A|B) = ?$$
$$P(A) = ?$$

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

здоровые пациенты



больные пациенты



$$P\ (B) = P(\text{болезнь}) = 2\% = 0,02$$

$$P\ (A|B) = P(+|\text{болен}) = 90\% = 0,9$$

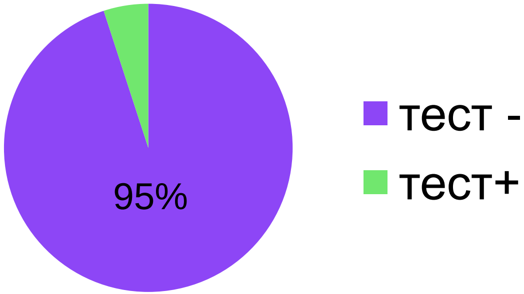
$$P(A) = P(+|\text{болен}) * P(\text{болезнь}) + P(+|\text{здоров}) * P(\text{здоров}) =$$

$$= 0,9 * 0,02 + (1 - 0,95) * (1 - 0,02) = 0,067$$

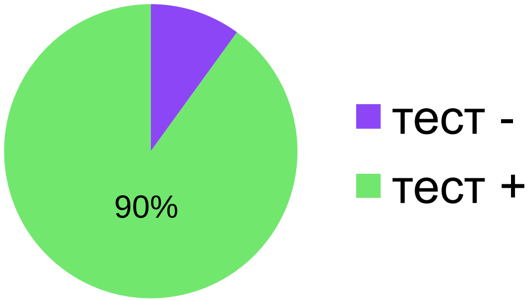
$$P\ (B|A) = \frac{P(A|B) * P(B)}{P(A)}$$

$$P\ (\text{болен}|+) = \frac{0,9 * 0,02}{0,067} = 0,269 = 26,9\%$$

здоровые пациенты



больные пациенты



- ✓ Изучили виды случайных событий:

  - возможные/невозможные
  - зависимые / независимые
  - совместные/несовместные

- ✓ Изучили комбинаторику: сочетания, размещения (порядок важен) , перестановки (порядок важен)

- ✓ Рассмотрели разницу между статистической вероятностью и классическим определением вероятности

- ✓ Запомнили, что вероятность – это доля

- ✓ Научились применять комбинаторику для расчета вероятности

- ✓ Теперь знаем, что для расчета вероятности события B, при условии, что A уже произошло, используется формула Байеса.