Теория вероятностей и математическая статистика

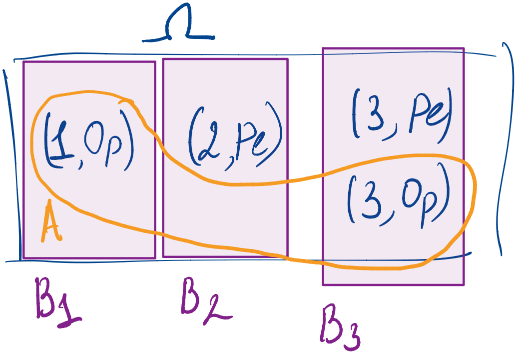
Полная вероятность. Теорема Байеса. Практическая сессия.

Глеб Карпов

ФКН ВШЭ

Пример 0

- Есть три монеты: у одной с обеих сторон орел, у второй с обеих сторон решка, у третьей орел с одной стороны и решка с другой. Мы выбираем монету наугад, подбрасываем её, и выпадает орел. Какова вероятность того, что на противоположной стороне решка?



$$P(B_i) = \frac{1}{3}$$

$A =$ "орёл"

$$A = (A \cap B_1) \cup \overset{\varnothing}{\cancel{(A \cap B_2)}} \cup (A \cap B_3)$$

$$P(A) = P(A \cap B_1) + 0 + P(A \cap B_3) =$$

$$= P(A|B_1) \cdot P(B_1) + P(A|B_3) \cdot P(B_3) =$$

$$= 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = 1/2$$

Пример 0

- Есть три монеты: у одной с обеих сторон орел, у второй с обеих сторон решка, у третьей орел с одной стороны и решка с другой. Мы выбираем монету наугад, подбрасываем её, и выпадает орел. Какова вероятность того, что на противоположной стороне решка?
- B_1, B_2, B_3 – события, соответствующие выбору монеты согласно условиям задачи (разбиение Ω). A – событие, соответствующее выпадению орла.

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) = 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 0.5 \cdot \frac{1}{3} = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

Согласно условию, нам нужно узнать вероятность того, что была выбрана третья монета.

$$P(B_3|A) = \frac{P(A|B_3)P(B_3)}{P(A)} = \frac{0.5 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

Пример 1

- Подбросили две игральные кости. Введем следующие события: A — на первой кости выпала тройка, B — сумма очков является четным числом, C — на второй кости выпало больше очков, чем на первой.

1. Найдите вероятность каждого из событий A, B и C.
2. Найдите условную вероятность $P(A \mid C)$.
3. Проверьте, есть ли среди событий A, C и $B \cap C$ пары независимых событий.

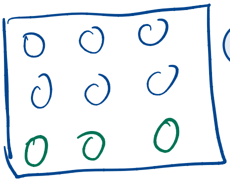
Пример 2

- Брат и сестра путешествуют на поезде. У них обоих нет билетов, и контролёр поймал их. Он уполномочен применить особое наказание за это правонарушение. У него есть коробка с девятью внешне одинаковыми шоколадными конфетами, три из которых содержат снотворное моментального действия, такое, что человек проспит месяц. Контролёр заставляет каждого из нарушителей по очереди выбрать и немедленно съесть одну конфету.

1. Если брат выбирает перед сестрой, какова вероятность того, что он уснет?
2. Если брат выбирает первым и не впадает в кому, какова вероятность того, что сестра останется бодрствующей?
3. Если брат выбирает первым и впадает в кому, какова вероятность того, что сестра останется бодрствующей?
4. В интересах ли сестры убедить брата выбирать первым? I.e. есть ли разница в вероятности засыпания в зависимости от того, кто выбирает первым?

Слайд для записей

B - брать 2-м и не уснуть



① A - брать 1-м и не уснуть

$$P(A) = \frac{6}{9}$$

② $P(B|A) = \frac{5}{8}$

③ $P(B|\bar{A}) = \frac{6}{8}$

④ $$P(B) = P(B|A)\cdot P(A) + P(B|\bar{A})\cdot P(\bar{A}) = \frac{5}{8}\cdot\frac{6}{9} + \frac{6}{8}\cdot\frac{3}{9} =$$

$$= \frac{48}{72} = \frac{6}{9}$$

Пример 3

False positives

- Редкое, но опасное заболевание имеет распространенность 1 на 10^5 в общей популяции. Существует диагностический тест, но он несовершенен. Если у человека есть заболевание, тест положительный с вероятностью 0.9; если нет - тест положительный с вероятностью $\frac{1}{20}$. Тест дал положительный результат. Какова вероятность того, что у человека **действительно** есть заболевание?

$P(ill) = \frac{1}{10^5}$

$P(+|ill) = 0.9$

$P(+|healthy) = \frac{1}{20}$

$P(ill|+) - ?$

Слайд для записей

$$P(+) = P(+|ill)\cdot P(ill) + P(+|healthy)\cdot P(healthy)$$

$$P(+) = 0.9\cdot\frac{1}{10^5} + \frac{1}{20}\cdot\frac{(10^5-1)}{10^5}$$

$$= 0.9\cdot 10^{-5} - 0.05\cdot 10^{-5} + 0.05 =$$

$$= 0.85\cdot 10^{-5} + 0.05 =$$

$$= \underline{0.05} + 0.85\cdot 10^{-5}$$

$$P(ill|+) = \frac{P(+|ill)\cdot P(ill)}{P(+)} = \frac{0.9\cdot 10^{-5}}{0.0500085} =$$

$$= \approx 0.00018$$

Пример 3

False positives

1. Редкое, но опасное заболевание имеет распространенность 1 на 10^5 в общей популяции. Существует диагностический тест, но он несовершенен. Если у человека есть заболевание, тест положительный с вероятностью 0.9; если нет — тест положительный с вероятностью $\frac{1}{20}$. Тест дал положительный результат. Какова вероятность того, что у человека **действительно** есть заболевание?
2. Обозначим вероятности: (H = healthy, I = ill, + for detected positive, − for detected negative) $P(I) = 10^{-5}$, $P(+|I) = 0.9$, $P(+|H) = \frac{1}{20}$. Мы заинтересованы в $P(I|+)$.
3. $P(+) = P(+|I)P(I) + P(+|H)P(H) = 0.9 \cdot 10^{-5} + 0.05 \cdot (1 - 10^{-5}) = 0.0500085$
4. $$P(I|+) = \frac{P(+|I)P(I)}{P(+)} = \frac{0.9 \cdot 10^{-5}}{0.0500085} \approx 0.0002$$

Пример 4

Компания А собирается выводить на рынок новую игровую приставку. Ее конкуренты – компании В и С, работающие независимо друг от друга, каждая из которых может опередить А с вероятностью 0.6. Если у приставки А на момент выхода в продажу не будет конкурентов, то она окажется успешной с вероятностью 0.9, если к этому моменту свою приставку успеет выпустить только один конкурент, то вероятность успеха равна 0.7, если же на рынке будут оба конкурента – вероятность успеха А равна 0.4.

1. Найти вероятность того, что продукт окажется успешным.
2. Известно, что продукт оказался успешным. Найти вероятность того, что это произошло при отсутствии конкурентов.
3. Известно, что продукт не оказался успешным. Найти вероятность того, что только один из конкурентов успел обогнать А.

A – приставка успешна

B_1 – никто не опередил

B_2 – опередила 1 компания

B_3 – опередили 2

Слайд для записей

$$P(B_1) = 0.4 \cdot 0.4 = 0.16$$

$$P(B_2) = \{(B_{\text{успела}}, C\text{-нет}), (B\text{-нет}, C\text{-успела})\} = 0.6 \cdot 0.4 + 0.4 \cdot 0.6 =$$
$$= 2 \cdot 0.4 \cdot 0.6 = 0.48$$

$$P(B_3) = 0.6 \cdot 0.6 = 0.36$$

$$P(A) = \underbrace{P(A \mid B_1)}_{0.9} \cdot P(B_1) + \underbrace{P(A \mid B_2)}_{0.7} \cdot P(B_2) + \underbrace{P(A \mid B_3)}_{0.4} \cdot P(B_3) = 0.624$$

$$P(B_1 \mid A) = \frac{P(A \mid B_1) \cdot P(B_1)}{P(A)} = \frac{0.9 \cdot 0.16}{0.624} = 0.23$$

Слайд для записей

$$P(B_2 \mid \bar{A}) = \frac{P(\bar{A} \mid B_2) \cdot P(B_2)}{P(\bar{A})} =$$

$$P(A \mid B_2) + P(\bar{A} \mid B_2) = 1$$

$$= \frac{0.3 \cdot 0.48}{1 - 0.624} = \frac{0.3 \cdot 0.48}{0.376} = \approx 0.383$$

Пример 5

В некоторой компании работают IT-специалисты трех профилей: разработчики фронтенда, разработчики бэкенда и тестировщики. Среди всех IT-сотрудников доля фронтендеров составляет 0.3; доля тестировщиков — 0.2; доля бэкендеров — 0.5.

Специфика IT-специалистов в компании такова, что в случае если в какой-то момент один из сотрудников вдруг временно выбывает (например, заболел) —заменить его может только тот, у кого есть подходящие навыки: тестировщика может заменить любой специалист; бэкенд-разработчика может заменить только другой бэкенд-разработчик; фронтенд-разработчика может заменить фронтенд- или бэкенд-разработчик.

Технический руководитель хочет просчитать все риски перед очередным важным релизом. Предположим, один из IT-специалистов компании заболел. И пусть известно, что случайно взятый другой IT-специалист может его заменить. Какова вероятность, что еще один, другой случайно взятый IT-специалист сможет тоже заменить этого заболевшего сотрудника?

Компанию считать достаточно большой, чтобы доли специалистов не менялись от выбывания нескольких человек. Выбор сотрудника для замены считать абсолютно случайным.

Пример 5

B – условие

A – 2ой сл. взятый заменит

Пространство исходов

Обозначения: Ф — фронтенд, Б — бэкенд, Т — тестировщик.
Тройка = (заболевший, 1-й выбр., 2-й выбр.).

заболевший	Ф			Б			Т		
Ф	ФФФ	ФФБ	ФФТ	ФБФ	ФББ	ФБТ	ФТФ	ФТБ	ФТТ
Б	БФФ	БФБ	БФТ	ББФ	БББ	ББТ	БТФ	БТБ	БТТ
Т	ТФФ	ТФБ	ТФТ	ТБФ	ТББ	ТБТ	ТТФ	ТТБ	ТТТ

(1-й выбр., 2-й выбр.)

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(\bigcirc)}{P(\square)}$$

$$P(ТБФ) = 0.2\cdot 0.5\cdot 0.3$$