План второй недели

Вся первая неделя и первые пять уроков второй недели посвящены вероятностным пространствам и инструментам для работы с ними. Шестой урок этой недели открывает новую тему – случайные величины. Случайными величинами мы будем заниматься всю третью и четвёртую недели.

Во второй неделе шесть основных уроков и два дополнительных.

1. На первом уроке мы пройдём Теорему Байеса и формулу полной вероятности
2. На втором уроке мы научимся вычислять вероятность объединения событий
3. Затем познакомимся с комбинаторикой
4. Поймём, как комбинаторика связана с вычислениями вероятностей
5. Добавим к нашему арсеналу биномиальные коэффициенты
6. В конце недели мы изучим случайные величины и математическое ожидание
7. А на дополнительных уроках мы изучим треугольник Паскаля и потренируемся решать задачи.

Уроки 2.1 - 2.2. Теорема Байеса это полезный инструмент для работы с условными вероятностями, а правило суммы пригодится нам в дальнейших задачах при вычислении вероятностей.

Уроки 2.3 - 2.5. Конечная цель – разобраться с применением комбинаторики для вычислений вероятности. Довольно часто в таких вычислениях участвуют биномиальные коэффициенты. Они применяются для вычисления вероятностей в ситуациях, когда есть много похожих независимых событий (или случайных величин). Например, когда каждый из 1000 пользователей совершает или не совершает какое-то действие (в этих же ситуациях возникает *биномиальное распределение*, которое мы пройдём на третьей неделе).

На собеседованиях часто предлагают задачи на вычисления вероятностей. В этих уроках мы освоим приёмы, позволяющие решать такие задачи, а также потренируем общие навыки работы с вероятностями.

Уроки 2.6 - 2.7. В Data science регулярно используются случайные величины. Определение случайной величины на конечном вероятностном пространстве мы изучим на этом уроке. На следующей неделе мы углубим наше понимание, а также пройдём случайные величины на счётном вероятностном пространстве.

Уроки 2.8-2.9 – дополнительные. В них мы изучим треугольник Паскаля и потренируемся решать задачи, похожие на задачи из первых двух недель. За задачи этих уроков не даются баллы.

Теорема Байеса

На этом уроке мы выведем несложную, но очень часто применяемую формулу — *теорему Байеса*. Она бывает полезна при нахождении условных вероятностей.

Затем мы обобщим пройденную ранее формулу полной вероятности для двух событий на случай произвольного количества событий. Ранее мы уже видели, как эта формула применяется, а теперь научимся считать по этой формуле в более общем случае.

Теорема Байеса и DS

Мы используем теорему Байеса каждый раз, когда вычисляем, как новые наблюдения изменяют нашу модель мира. Другими словами теорема Байеса применяется в такой ситуации. Пусть есть интересное нам событие, вероятность которого мы считаем равным числу $c \in [0, 1]$. Теперь мы получили дополнительные данные, которые как-то связаны с нашим событием. Как изменится вероятность интересного нам события в свете новых наблюдений? Увеличится или уменьшится, и насколько?

Дополнительный материал. Видео про теорему Байеса на наглядных примерах (9 минут)

Теорема Байеса

Теорема Байеса. Для любых событий A и B, таких что $P(A), P(B) \neq 0$ выполнено

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

Теорема Байеса доказывается в одну строчку, и на одном из следующих шагов мы попросим вас это сделать.

Вот способ думать про теорему Байеса. Есть некоторое неизвестное событие A, вероятность которого мы знаем. Мы пронаблюдали событие B, которое как-то связано с A. Как нам теперь подправить вероятность события A в свете этого нового наблюдения? Другими словами, какова $P(A|B)$? На этот вопрос и отвечает теорема Байеса.

Более подробно этот взгляд на теорему Байеса описывается так:

- Событие A отвечает за какое-то скрытое явление. Например, A может быть "двигатель исправен", "змея ядовитая", "мой коллега меня недолюбливает".
- Событие B отвечает за наблюдение, которое как-то связано с A. Например, B может быть "двигатель громко шумит", "змея ярко оранжевая", "мой коллега принёс мне кофе".
- Мы по какой-то причине знаем $P(B|A)$. То есть мы знаем, насколько вероятно пронаблюдать B, если A случилось. Например, мы знаем, что исправные двигатели громко шумят только в 2 из 10 случаев.
- Мы знаем $P(B)$. Например, мы знаем, что двигатели громко шумят в 3 из 10 случаев.
- Мы знаем $P(A)$. Вероятность $P(A)$ называют *априорной* вероятностью, то есть вероятностью, которую мы приписывали событию A до наблюдения B. Например, 0.8 двигателей исправны.
- Теперь по теореме Байеса мы вычисляем $P(A|B)$, то есть вероятность события A после наблюдения события B. Поэтому вероятность $P(A|B)$ называют *апостериорной* вероятностью события A. Например, $P(A|B)$ будет вероятностью исправности двигателя после того, как мы пронаблюдали, что двигатель громко шумит.

Пример

Давайте уже досчитаем про двигатели. Событие A – двигатель исправен, событие B – двигатель громко шумит. Мы заводим какой-то конкретный двигатель и обнаруживаем, что он громко шумит. Мы хотим понять, какова вероятность исправности двигателя в свете этого наблюдения. То есть мы хотим найти $P(A|B)$. Пусть вероятности $P(A) = 0.8$, $P(B) = 0.3$, $P(B|A) = 0.2$ Тогда:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{0.2 \ \cdot 0.8}{0.3} = \frac{0.16}{0.3} \approx 0.533$$

Заметьте, что исправные двигатели громко шумят в 2 из 10 случаев($P(B|A) = 0.2$), а все двигатели громко шумят в 3 из 10 случаев ($P(B) = 0.3$). Напрашивается вывод, что неисправные двигатели шумят чаще, чем исправные. То есть громкий шум двигателя свидетельствует в пользу того, что двигатель неисправен. Действительно, $P(A) = 0.8 > 0.533 = P(A|B)$. То есть вероятность исправности двигателя стала ниже после того, как мы услышали, что он громко шумит.

Заключение

Теорема Байеса отвечает за изменение наших представлений о мире в свете новых наблюдений. То есть за переход от априорной вероятности $P(A)$ (до наблюдения) к апостериорной вероятности $P(A|B)$ (после наблюдения). В обычной жизни люди редко используют теорему Байеса и переходят от $P(A)$ к $P(A|B)$ на интуитивном уровне. Часто это приводит к неверным результатам – то есть к неверной интерпретации того, насколько должна измениться вероятность A после наблюдения B.

Теорема Байеса названа "теоремой" не за сложность, а за то, что она часто применяется. В теории вероятностей довольно часто, когда мы проводим вычисления с условными вероятностями, мы используем теорему Байеса (или неявно доказываем её заново). Часто мы просто не акцентируем внимание на применении теоремы Байеса. Например, последнюю задачу этого урока можно решить, ни слова не сказав про теорему Байеса, но если приглядеться, то видно, что теорема Байеса там применяется.

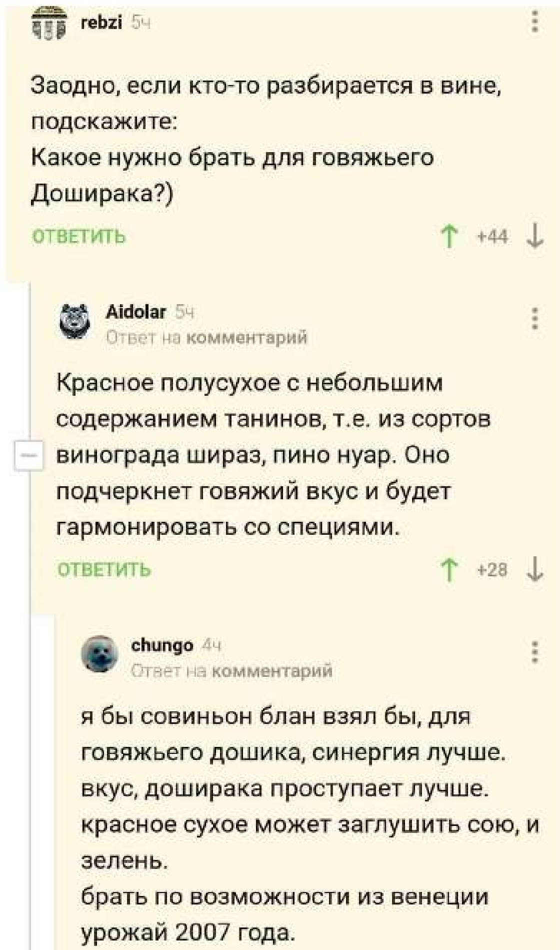
Запоминать формулу из теоремы Байеса не стоит – вы всегда сможете вывести её заново.

Задача. Теорема Байеса 1

Докажите теорему Байеса: $P(A|B) = \dfrac{P(B|A) \cdot P(A)}{P(B)}$ при $P(A), P(B) \neq 0$.

Мы предлагаем сомелье Василию попробовать разные вина. Часть вин была произведена на Сицилии. А остальные вина были куплены в ближайшем магазине по 250 рублей за коробку.

Пусть событие A – "вино было произведено на Сицилии", событие B – "сомелье Василий дал вину положительную оценку". При этом $P(A) = 0.3, P(B) = 0.4$ и $P(B|A) = 0.6$. Найдите вероятность $P(A|B)$. То есть вероятность того, что вино, которому Василий дал положительную оценку, было произведено на Сицилии.



Введите численный ответ

Введите число

Обозначим за A событие "змея ядовитая", а за B событие "змея оранжевая".

Известно, что $P(A) = 0.2$ и $P(B) = 0.3$. При этом $P(B|A) = 0.9$ то есть если змея ядовитая, то она почти наверняка оранжевая.

Найдите $P(A|B)$, то есть вероятность того, что оранжевая змея ядовитая.

Введите численный ответ

Введите число

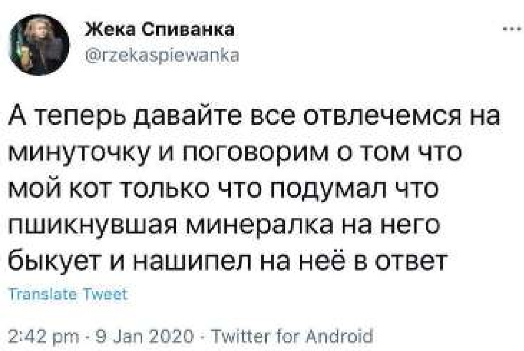
Усложним задачу – не будем явно давать $P(B)$.

В этой задаче мы будем предсказывать эмоциональный профиль кота по его детским привычкам.

Взрослые коты бывают двух видов – спокойные и беспокойные. Известно, что некоторые котята шумят по ночам, а некоторые не шумят. По котёнку нельзя однозначно понять, вырастет он спокойным или беспокойным.

- 0.8 беспокойных котов в детстве шумели по ночам.
- 0.3 спокойных котов в детстве шумели по ночам.
- Среди взрослых котов 0.4 котов беспокойные, а 0.6 спокойные.

Мы смотрим на котёнка, который шумит по ночам. С какой вероятностью он вырастет беспокойным котом?



Заполните пропуски

Обозначим за событие A то, что кот беспокойный, за событие B – наблюдение, что котёток шумел по ночам. Тогда $P(B|A) =$ ______ , $P(B|\bar{A}) =$ ______ .

Значит, по формуле полной вероятности $P(B) = P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A}) =$ ______

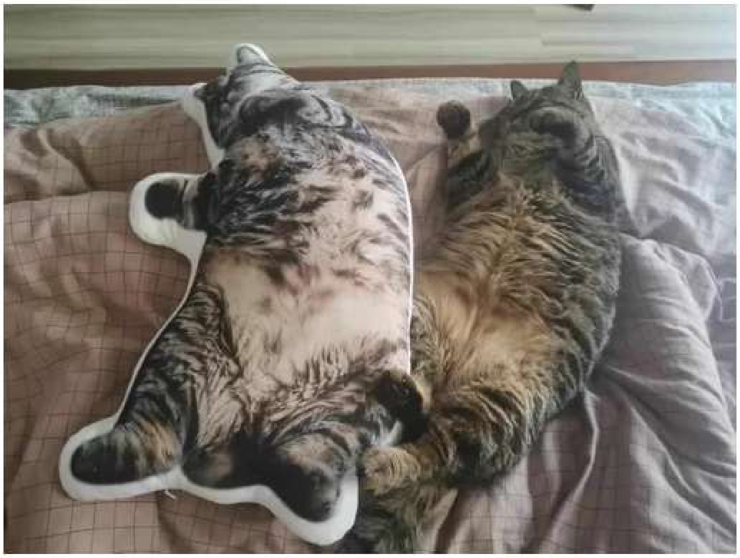
Отсюда следует, что $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} =$ ______

Задача с проверкой. Теорема Байеса 2

1. Мы держим в руках монетку. С вероятностью $\frac{99}{100}$ она честная. С вероятностью $\frac{1}{100}$ это нечестная монетка, которая всегда выпадает орлом вверх. Мы подбросили монетку 5 раз, и все разы она выпала орлом вверх. Ясно, что это наблюдение свидетельствует в пользу нечестности монетки. Какова вероятность того, что монетка нечестная?
2. Тот же вопрос, что и в Пункте 1, но мы подбросили монетку 20 раз и все разы выпал орёл. Сравните ответ с ответом из Пункта 1.

Попробуйте интерпретировать эту задачу через теорему Байеса. Что будет скрытым состоянием? Что будет наблюдением?

Проверка. Заполните пропуски, записав ответы с точностью до **четвёртого** знака после запятой.



Заполните пропуски

1. Вероятность того, что монетка нечестная, если орёл выпал 5 раз подряд, равна

2. Вероятность того, что монетка нечестная, если орёл выпал 20 раз подряд, равна

Придумайте свой пример из жизни про скрытое состояние A и наблюдение B. Придумайте реалистичные вероятности $P(A), P(B)$ и $P(B|A)$, и найдите $P(A|B)$.

(задача не оценивается)

Больше двух вариантов

Пусть есть скрытое состояние A, вероятность которого мы хотим подправить на основании наблюдения события B. По сути, здесь мы работаем с выбором между A и $\bar{A}$ – то есть между двумя возможными скрытыми состояниями. В примере из четвёртого шага этого урока эти состояния это "двигатель исправен" и "двигатель неисправен".

В более общей постановке вопроса возможных вариантов скрытых состояний может быть не 2, а больше. Рассмотрим это на примере наблюдения B "у человека есть кашель".

1. Вариант двух скрытых состояний. Событие A – у человека есть грипп, событие $\bar{A}$ – человек здоров.
2. Вариант нескольких состояний. Событие A_1 – у человека грипп, A_2 – астма, A_3 – бронхит, A_4 – человек здоров (здесь мы предполагаем, что человек не может быть болен двумя из этих болезней одновременно)

Формула полной вероятности

В таких ситуациях бывает полезна *формула полной вероятности*. Формулу полной вероятности для двух событий (случая из пункта 1 предыдущего параграфа) мы её уже разобрали на этом шаге. В общем случае получится так:

Даны события $A_1, A_2, \dots, A_n$, такие что каждый элементарный исход из Ω лежит ровно в одном из этих событий. Другими словами,

- эти события не пересекаются друг с другом, то есть $A_i \cap A_j = \emptyset$ для любых $i \neq j$
- и их объединение равно Ω, то есть $\bigcup\limits_{i=1}^{n} A_i = \Omega$.

Рассмотрим любое событие B. Ясно, что B можно представить в виде объединения непресекающихся подмножеств $B \cap A_1,\ B \cap A_2, \dots,\ B \cap A_n$. Поэтому выполнено:

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i).$$

Воспользовавшись формулой, связывающей пересечение и условную вероятность: $P(B \cap A_i) = P(B|A_i) \cdot P(A_i)$, мы можем переписать последнее равенство как

$$P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i).$$

Неформально. Интерпретировать это можно так. Условия на $A_1, \dots, A_n$ говорят, что может произойти ровно одно событие из $A_1, \dots, A_n$. То есть не может произойти несколько из них одновременно, и не может случиться, что ни одно из них не произошло. Тогда вероятность события B это сумма таких выражений по всем возможным i: берём вероятность события B при условии, что A_i произошло, и умножаем на вероятность A_i. Каждое такое выражение равно вероятности того, что B и A_i произошли одновременно.

Формулы $P(B) = \sum\limits_{i=1}^{n} P(B \cap A_i)$ и $P(B) = \sum\limits_{i=1}^{n} P(B|A_i) \cdot P(A_i)$ называются *формулой полной вероятности*.

На следующем шаге мы увидим, как применять формулу полной вероятности вместе с теоремой Байеса.

Задача с проверкой. Теорема Байеса 3

Вы покупаете одинаковые аккумуляторы у трёх поставщиков: X, Y и Z. На основании предыдущих покупок у этих поставщиков вы знаете, какова доля брака в продукции каждого из них.

- Среди всех ваших аккумуляторов доля купленных у X аккумуляторов равна 0.6. Среди аккумуляторов, которые вы покупаете у X, доля брака 0.1.
- Среди всех ваших аккумуляторов доля купленных у Y аккумуляторов равна 0.3. Среди аккумуляторов, которые вы покупаете у Y, доля брака 0.2.
- Среди всех ваших аккумуляторов доля купленных у Z аккумуляторов равна 0.1. Среди аккумуляторов, которые вы покупаете у Z, доля брака 0.05.

Задача. Вы смотрите на один из ваших аккумуляторов, который оказался бракованным. Какова вероятность, что его вам продал Y?

Попробуйте интерпретировать эту задачу через теорему Байеса. Что будет скрытым состоянием? Что будет наблюдением?

Проверка. Введите ответ.

Введите численный ответ

Введите число

Что мы прошли на этом уроке

На этом уроке мы

- познакомились со взглядом на мир через теорему Байеса:

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

- доказали формулу полной вероятности: для событий $A_1, A_2, \ldots, A_n$, таких что $A_i \cap A_j = \emptyset$ для любых $i \neq j$ и $\bigcup\limits_{i=1}^{n} A_i = \Omega$ выполнено

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$$

Что нас ждёт на следующем уроке

На следующем уроке мы

- познакомимся с правилом суммы вероятностей
- научимся в общем случае считать вероятность объединения событий