50 идей,
о которых нужно знать

МАТЕМАТИКА

Тони Крилли

32 Теория Байеса

О ранних годах жизни преподобного Томаса Байеса нам мало что известно. Он родился на юго-востоке Англии примерно в 1702 году и стал священником-нонконформистом, а заодно заработал репутацию математика и был избран в 1742 году в Лондонское королевское общество. Знаменитая работа Байеса «Очерки к решению задач в учении о случае» была опубликована в 1763-м — через два года после смерти автора. В ней приводилась формула для нахождения обратной вероятности, вероятности «наоборот», и с ее помощью была развита ключевая концепция байесовской философии — условная вероятность.

Томас Байес дал имя байесовцам — приверженцам особой разновидности статистики, в противовес традиционным статистикам, или «частотникам». Сторонники частотного подхода к вероятностям придерживаются взглядов на вероятность, основанных на упрямых численных данных. Байесовский подход базируется на знаменитой формуле Байеса и на принципе, что с субъективной степенью уверенности можно обращаться, как с математической вероятностью.

Условная вероятность Вообразите: бравому доктору Почему необходимо диагностировать корь у своих пациентов. Сыпь — показатель, но диагноз по ней впрямую не ставят. У пациента может быть корь без сыпи, а у некоторых — сыпь без кори. Вероятность того, что у пациента сыпь, *при условии* наличия у него кори, и есть условная вероятность. Байесовцы в своих записях изображают это самое «при условии» вертикальной линией:

$$P(\text{у пациента сыпь} \mid \text{у пациента корь}),$$

и это означает вероятность, что у пациента сыпь при условии, что у него корь. Значение P(у пациента сыпь | у пациента корь) не то же самое,

СТРЕЛА ВРЕМЕНИ

1763
Изданы очерки Томаса Байеса о вероятности

1937
Бруно де Финетти настаивает на субъективном подходе к вероятности как на альтернативе «частотному» подходу

что P(у пациента корь | у пациента сыпь). Относительно друг друга эти вероятности обратны. Формула Байеса предназначена для расчетов одной вероятности через другую.

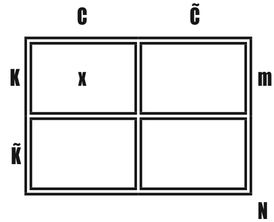

Диаграмма Венна, показывающая логическую структуру появления сыпи и наличия кори

Математики обожают условные обозначения реальных вещей. Скажем, событие наличия кори обозначается K, а событие наличия у пациента кори – C. Символ $\tilde{C}$ – событие, при котором у пациента *нет* сыпи, и $\tilde{K}$ – событие, при котором у пациента *нет* кори. Все это отражено в диаграмме Венна.

Из этой диаграммы доктор Почему видит, что есть x пациентов с корью и сыпью, m пациентов с корью, а общее число пациентов – N. Ему также понятно следующее: вероятность того, что некто болеет корью и покрыт сыпью, – попросту x/N, а вероятность того, что некто болеет корью, – m/N. Условная вероятность – вероятность того, что у кого-то сыпь, при условии, что у этого человека корь, записанная $P(C \mid K)$, равна x/m. Сведя все вместе, доктор Почему получит вероятность того, что некто и болеет корью, и покрыт сыпью:

$$P(K \text{ и } C) = \frac{x}{N} - \frac{x}{m} \times \frac{m}{N}$$

или

$$P(K \text{ и } C) = P(C \mid K) \times P(K)$$

и, сходно,

$$P(K \text{ и } C) = P(K \mid C) \times P(C).$$

Формула Байеса Приравнивая выражения для $P(K \text{ и } C)$, получаем формулу Байеса – отношение между условной вероятностью и ее противоположностью. Доктор Почему получит уверенное представление о $P(K \mid C)$ – вероятности наличия сыпи при наличии кори. Условная вероятность обратного события – как раз того, которое интересует доктора Почему больше, – есть ли у пациента корь, если у него сыпь. Установить это означает решить обратную задачу, и именно ее решал Байес в своих работах. Чтобы вычислить вероятность, нужно вбросить кое-какие цифры. Они будут субъективными, но нам важно увидеть, как это работает. Вероятность того, что, если у пациента корь, у него есть сыпь, $P(C \mid K)$, – довольно высока; допустим, это 0,9; если у пациента нет кори, вероятность сыпи – $P(C \mid \tilde{K})$ – окажется низкой; допустим, 0,15.

$$P(K \mid C) = \frac{P(K)}{P(C)} \times P(C \mid K)$$

Формула Байеса

1950 Джимми Сэвидж и Деннис Линдли возглавляют современное байесовское движение

1950-е Термин «байесовский» входит в оборот

1992 Основано Международное общество байесовского анализа

В обеих ситуациях доктор Почему сможет с приличной точностью оценить значения этих вероятностей. Бравый доктор также сможет прикинуть процент людей среди населения, у которых есть корь, — допустим, их 20%. Эта вероятность записывается как $P(K) = 0{,}2$.

Нам еще нужна лишь одна цифра — $P(C)$, т. е. процент людей среди населения, у которых сыпь. Итак, вероятность того, что у кого-то сыпь, есть вероятность того, что у кого-то корь и сыпь плюс вероятность того, что у кого-то кори нет, а сыпь есть. Исходя из нашего ключевого выражения, $P(C) = 0{,}9 \times 0{,}2 + 0{,}15 \times 0{,}8 = 0{,}3$. Подставляем все значения в формулу Байеса:

$$P(K \mid C) = \frac{0{,}2}{0{,}3} \times 0{,}9 = 0{,}6.$$

Таким образом, из всех своих пациентов с сыпью врач правильно определяет корь в 60% случаев. Допустим, врач получает больше информации о штамме кори и вероятность диагностирования растет, т. е. $P(C \mid K)$ — вероятность появления сыпи из-за кори — возрастает с 0,9 до 0,95, а $P(C \mid \tilde{K})$ — вероятность появления сыпи по иным причинам — упадет с 0,15 до 0,1. Как это повлияет на точность диагностики? Каково новое значение $P(K \mid C)$? С учетом новых данных $P(C) = 0{,}95 \times 0{,}2 + 0{,}1 \times 0{,}8 = 0{,}27$, а $P(K \mid C)$ по формуле Байеса есть 0,2, деленные на $P(C) = 0{,}27$ и умноженные на 0,95, что в итоге даст нам 0,704. Следовательно, доктор Почему теперь сможет определять корь в 70% случаев, т. е. точнее. Если вероятности станут 0,99 и 0,01 соответственно, вероятность определения кори, $P(K \mid C)$, станет 0,961, т. е. вероятность правильного диагноза возрастет до 96%.

Современные байесовцы Традиционные статистики совершенно не против формулы Байеса — когда вероятность можно измерить. Камень преткновения — интерпретация вероятности как степени уверенности или, как ее еще иногда называют, субъективной вероятности.

В суде вопрос вины или невиновности подчас определяется «балансом вероятностей». Говоря строго, этот критерий применим лишь к гражданским делам, но можно вообразить и сценарий, в котором его применяют к делам уголовным. Статистики, придерживающиеся частотного подхода, теряются перед задачей определения значения вероятности того, что подсудимый виновен. Байесовцы же не стесняются учитывать свои чувства. Как же это устроено? Если применять подход баланса вероятностей к оценке вины и невиновности, необходимо разобраться, как можно жонглировать вероятностями. Рассмотрим возможное развитие сюжета.

Член жюри присяжных выслушал дело в суде и решил, что вероятность вины осужденного — примерно 1 к 100. Во время совещания жюри этого присяжного вызывают в зал суда, заслушать еще одно свидетельство от обвиняющей стороны. У подсудимого в доме найдено оружие и глава обвинения считает, что вероятность нахождения оружия дома у подсудимого — целых 0,95, если подсудимый виновен, а если он невиновен, то вероятность нахождения у него оружия — всего 0,1.

Следовательно, вероятность нахождения оружия дома у подсудимого гораздо выше, если подсудимый виновен, чем в ситуации его невиновности. Вопрос к присяжному таков: как скорректировать отношение к подсудимому в свете этой новой информации? Применим нашу форму записи и обозначим *В* ситуацию, в которой подсудимый виновен, и *П* — нахождение новых показаний. Присяжный произвел начальную оценку вероятности $\mathrm{P}(B) = 1/100$ или 0,01. Эта вероятность называется априорной. Заново оцененная вероятность $\mathrm{P}(B \mid \Pi)$ — пересмотренная вероятность вины с учетом показаний *П*, ее называют апостериорная вероятность. Формула Байеса в таком случае приобретает вид:

$$\mathrm{P}(B \mid \Pi) = \frac{\mathrm{P}(\Pi \mid B)}{\mathrm{P}(\Pi)} \times \mathrm{P}(B)$$

и иллюстрирует мысль о том, что априорная вероятность корректируется новой, апостериорной вероятностью $\mathrm{P}(B \mid \Pi)$. Аналогично вычислению $\mathrm{P}(C)$ в медицинском примере найдем $\mathrm{P}(\Pi)$ и получим:

$$\mathrm{P}(B \mid \Pi) = \frac{0{,}95}{0{,}95 \times 0{,}01 + 0{,}1 \times 0{,}99} \times 0{,}01 = 0{,}088.$$

Такой результат ставит присяжного в трудное положение: его исходная оценка виновности, равная 1%, возросла почти до 9%. Если бы сторона обвинения предположила, что вероятность нахождения оружия — аж целых 0,99, если подсудимый виновен, а если невиновен — 0,01, тогда, повторив вычисления по формуле Байеса, получим, что в этой ситуации присяжному пришлось бы пересмотреть свою оценку с 1% до 50%.

Применение формулы Байеса к таким случаям критикуют. Более всего оказывается под ударом способ оценки априорной вероятности. В свою защиту байесовский анализ предлагает метод обращения с субъективными вероятностями и корректирования их при получении дополнительных данных. Байесовский метод применим в широком диапазоне областей знания — от науки и предсказания погоды до уголовного права. Его сторонники доказывают разумность и практичность метода при обращении с неопределенностью. В его пользу говорит многое.

В сухом остатке: Убеждения с поправкой на сведения