Идеи Байеса для менеджеров

Кто такой Байес? и какое отношение он имеет к менеджменту? – может последовать вполне справедливый вопрос. Пока поверьте мне на слово: это очень важно!.. и интересно (по крайней мере, мне).

Томас Байес – английский пресвитерианский священник и по совместительству математик, живший в XVIII веке (1702–1761). Он развивал теорию вероятностей. Основная (и как я понимаю, единственная) его работа была опубликована уже после смерти в 1764 г. Байес предложил теорему, в последствии названную в его честь.

На самом деле, Байес не просто предложил формулу, позволяющую вычислять вероятность одного события на основании знания вероятностей других событий. Он изменил парадигму (к слову, и мою тоже). Люди (практически все, даже математически подкованные) довольно плохо обращаются с вероятностями. Все мы *преувеличиваем* значение конкретного наблюдения, *преуменьшая* априорное знание (то есть, знание имевшееся у нас до наблюдения).

Хорошей иллюстрацией такого подхода служит детская загадка: «Допустим ты капитан. Тебе нужно… (затем следует рассказ на несколько минут, и наконец вопрос…) Сколько лет капитану?» Всё наше внимание сосредоточено на рассказе, и мы пытаемся найти в нем подсказку, но, увы, тщетно. Мы совершенно забыли об априорном знании – установке, что капитан то я сам. Так и Байес учит нас не забывать о том, что было известно с самого начала. Последующие наблюдения (эксперименты) могут уточнить наше априорное знание, могут поколебать его, но не могут полностью затмить.

В какой парадигме действуют большинство менеджеров: если я наблюдаю нечто, какие выводы могу из этого сделать? Чему учит Байес: что должно быть на самом деле, чтобы мне довелось наблюдать это нечто? Именно так развиваются все науки, и об этом пишет Томас Кун (цитирую по памяти): человек, у которого нет в голове теории, будет шарахаться от одной идеи к другой под воздействием различных событий (наблюдений). Не даром говорят: нет ничего более практичного, чем хорошая теория.

Пример из практики. Мой подчиненный совершает ошибку, и мой коллега (руководитель другого отдела) говорит, что надо бы оказать управленческое воздействие на нерадивого сотрудника (проще говоря, наказать/обругать). А я знаю, что этот сотрудник делает 4–5 тысяч однотипных операций в месяц, и совершает за это время не более 10 ошибок. Чувствуете различие в парадигме? Мой коллега реагирует на наблюдение, а я обладаю априорным знанием, что сотрудник допускает некоторое количество ошибок, так что еще одна не повлияла на это знание… Вот если по итогам месяца окажется, что таких ошибок, например, 15!.. Это уже станет поводом для изучения причин несоответствия стандартам.

Убедил в важности Байесовского подхода? Заинтриговал? Надеюсь, что «да». А теперь ложка дегтя. К сожалению, идеи Байеса редко даются с первого захода. Мне откровенно не повезло, так как я знакомился с этими идеями по популярной литературе, после прочтения которой оставалось много вопросов. Планируя написать заметку, я собрал всё, что ранее конспектировал по Байесу, а также изучил, что пишут в Интернете. Предлагаю вашему вниманию мое лучшее предположение[1] на тему *Введение в Байесовскую вероятность*.[2]

Вывод теоремы Байеса

Рассмотрим следующий эксперимент:[3] мы называем любое число лежащее на отрезке [0; 1] и фиксируем, когда это число будет, например, между 0,1 и 0,4 (рис. 1а). Вероятность этого события равна отношению длины отрезка [0,1; 0,4] к общей длине отрезка [0; 1], при условии, что появления чисел на отрезке [0, 1] *равновероятны*. Математически это можно записать $p(0{,}1 <= x <= 0{,}4) = 0{,}3$, или кратко $p(X) = 0{,}3$, где p – вероятность, x – случайная величина в диапазоне [0; 1], X – случайная величина в диапазоне [0,1; 0,4]. То есть, вероятность попадания в отрезок [0,1; 0,4] равна 30%.

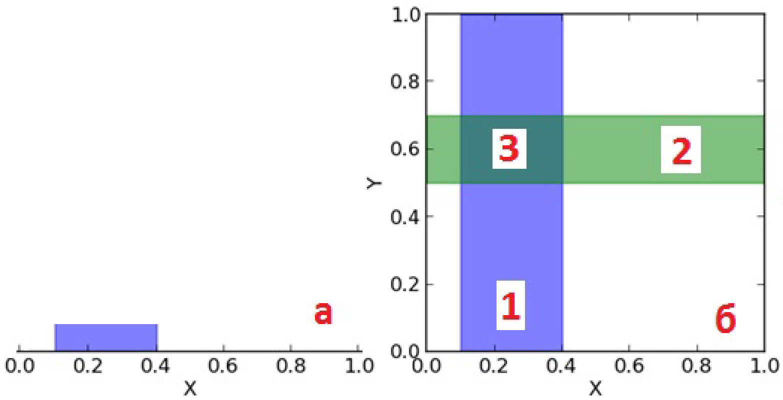

Рис. 1. Графическая интерпретация вероятностей

Теперь рассмотрим квадрат [0; 1] x [0; 1] (рис. 1б). Допустим, мы должны называть пары чисел (x, y), каждое из которых больше нуля и меньше единицы. Вероятность того, что x (первое число) будет в пределах отрезка [0,1; 0,4] (синяя область 1), равна отношению площади синей области к площади всего квадрата, то есть (0,4 – 0,1) * (1 – 0) / (1 * 1) = 0,3, то есть те же 30%. Вероятность того, что y находится внутри отрезка [0,5; 0,7] (зеленая область 2) равна отношению площади зеленой области к площади всего квадрата $p(0{,}5 <= y <= 0{,}7) = 0{,}2$, или кратко $p(Y) = 0{,}2$.

Что можно узнать о значениях одновременно x и y. Например, какова вероятность того, что одновременно x и y находятся в соответствующих заданных отрезках? Для этого надо посчитать отношение площади области 3 (пересечения зеленой и синей полос) к площади всего квадрата: $p(X, Y) = (0{,}4 – 0{,}1) * (0{,}7 – 0{,}5) / (1 * 1) = 0{,}06$.

А теперь допустим мы хотим знать какова вероятность того, что y находится в интервале [0,5; 0,7], если x уже находится в интервале [0,1; 0,4]. То есть фактически у нас есть фильтр и когда мы называем пары (x, y), то мы сразу отбрасывает те пары, которые не удовлетворяют условию нахождения x в заданном интервале, а потом из отфильтрованных пар мы считаем те, для которых y удовлетворяет нашему условию и считаем вероятность как отношение количества пар, для которых y лежит в вышеупомянутом отрезке к общему количеству отфильтрованных пар (то есть для которых x лежит в отрезке [0,1; 0,4]). Мы можем записать эту вероятность как $p(Y|X)$. Читается: «вероятность попадания y в диапазон [0,5; 0,7], при условии, что x попал в диапазоне [0,1; 0,4]». Очевидно, что эта вероятность равна отношению площади области 3 к площади синей области 1. Площадь области 3 равна (0,4 – 0,1) * (0,7 – 0,5) = 0,06, а площадь синей области 1 (0,4 – 0,1) * (1 – 0) = 0,3, тогда их отношение равно 0,06 / 0,3 = 0,2. Другими словами, вероятность нахождения y на отрезке [0,5; 0,7] при условии, что x принадлежит отрезку [0,1; 0,4] $p(Y|X) = 0{,}2$.

---

[1] Физик Нобелевский лауреат Ричарда Фейнмана, отзываясь об одном философе с особо большим самомнением, как-то сказал: «Меня раздражает вовсе не философия как наука, а та помпезность, которая создана вокруг нее. Если бы только философы могли сами над собой посмеяться! Если бы только они могли сказать: «Я говорю, что это вот так, а Фон Лейпциг считал, что это по-другому, а ведь он тоже кое-что в этом смыслит». Если бы только они не забывали пояснить, что это всего лишь их ***лучшее предположение***»

[2] Дальнейшее изложение требует некоторого понимания вероятностей; если вы хотите освежить свои знания в этой области начните с Википедии и/или заметки Основные понятия теории вероятностей.

[3] Я адаптировал пример, подсмотренный у Maxim_I.

В предыдущем абзаце мы фактически сформулировали тождество: $p(Y|X) = p(X, Y) / p(X)$. Читается: «вероятность попадания *y* в диапазон [0,5; 0,7], при условии, что *x* попал в диапазон [0,1; 0,4], равна отношению вероятности одновременного попадания *x* в диапазон [0,1; 0,4] и *y* в диапазон [0,5; 0,7], к вероятности попадания *x* в диапазон [0,1; 0,4]».

По аналогии рассмотрим вероятность $p(X|Y)$. Мы называем пары (x, y) и фильтруем те, для которых *y* лежит между 0,5 и 0,7, тогда вероятность того, что *x* находится в отрезке [0,1; 0,4] при условии, что *y* принадлежит отрезку [0,5; 0,7] равна отношению площади области 3 к площади зеленой области 2: $p(X|Y) = p(X, Y) / p(Y)$.

Заметим, что вероятности $p(X, Y)$ и $p(Y, X)$ равны, и обе равны отношению площади зоны 3 к площади всего квадрата, а вот вероятности $p(Y|X)$ и $p(X|Y)$ не равны; при этом вероятность $p(Y|X)$ равна отношению площади области 3 к области 1, а $p(X|Y)$ – области 3 к области 2. Заметим также, что $p(X, Y)$ часто обозначают как $p(X\&Y)$.

Итак, мы ввели два определения: $p(Y|X) = p(X, Y) / p(X)$ и $p(X|Y) = p(X, Y) / p(Y)$

Перепишем эти равенства виде: $p(X, Y) = p(Y|X) * p(X)$ и $p(X, Y) = p(X|Y) * p(Y)$

Поскольку левые части равны, равны и правые: $p(Y|X) * p(X) = p(X|Y) * p(Y)$

Или мы можем переписать последнее равенство в виде:

$$p(X|Y) = \frac{p(Y|X) \times p(X)}{p(Y)}$$

Это и есть теорема Байеса!

Неужели столь несложные (почти тавтологические) преобразования рождают великую теорему!? Не спешите с выводами. Давайте еще раз проговорим, что же мы получили. Имелась некая исходная (априорная) вероятность $p(X)$, того, что случайная величина *x* равномерно распределенная на отрезке [0; 1] попадает в диапазон *X* [0,1; 0,4]. Произошло некое событие *Y*, в результате которого мы получили апостериорную вероятность той же самой случайной величины *x*: $p(X|Y)$, и эта вероятность отличается от $p(X)$ на коэффициент $\frac{p(Y|X)}{p(Y)}$. Событие *Y* называется свидетельством, в большей или меньшей степени подтверждающим или опровергающим *X*. Указанный коэффициент иногда называют *мощностью свидетельства*. Чем мощнее свидетельство, тем больше факт наблюдения Y изменяет априорную вероятность, тем больше апостериорная вероятность отличается от априорной. Если свидетельство слабое, апостериорная вероятность почти равна априорной.

Формула Байеса для дискретных случайных величин

В предыдущем разделе мы вывели формулу Байеса для непрерывных случайных величин x и y, определенных на отрезке [0; 1]. Рассмотрим пример с дискретными случайными величинами, принимающими каждая по два возможных значения. В ходе проведения плановых медицинских осмотров установлено, что в сорокалетнем возрасте 1% женщин болеет раком молочной железы. 80% женщин больных раком получают положительные результаты маммографии. 9,6% здоровых женщин также получают положительные результаты маммографии. В ходе проведения осмотра женщина данной возрастной группы получила положительный результат маммографии. Какова вероятность того, что у неё на самом деле рак молочной железы?

Если вы испытываете трудности с ответом, вам следует знать, что только 15% врачей дают верный ответ. И тем не менее, не спишите читать дальше, попробуйте найти решение самостоятельно.

Ход рассуждений/вычислений следующий. Из 1% больных раком маммография даст 80% положительных результатов = 1%*80% = 0,8%. Из 99% здоровых женщин маммография даст 9,6% положительных результатов = 99%*9,6% = 9,504%. Итого из 10,304% (9,504% + 0,8%) с положительными результатами маммографии, только 0,8% больных, а остальные 9,504% здоровых. Таким образом, вероятность того, что при положительном результате маммографии женщина больна раком составляет 0,8%/10,304% = 7,764%. А вы думали, что 80% или около того?

В нашем примере формула Байеса принимает следующий вид:

$$p(X_1|Y_1) = \frac{p(Y_1|X_1) * p(X_1)}{p(Y_1|X_1) * p(X_1) + p(Y_1|X_2) * p(X_2)}$$

Давайте еще раз проговорим «физический» смысл этой формулы. X – случайная величина (диагноз), принимающая значения: X_1 – болен и X_2 – здоров; Y – случайная величина (результат измерения – маммографии), принимающая значения: Y_1 – положительный результат и Y_2 – отрицательный результат; $p(X_1)$ – вероятность болезни до проведения маммографии (априорная вероятность), равная 1%; $p(Y_1|X_1)$ – вероятность положительного результата в случае, если пациентка больна (условная вероятность, так как она должна быть задана в условиях задачи), равная 80%; $p(Y_1|X_2)$ – вероятность положительного результата в случае, если пациентка здорова (также условная вероятность), равная 9,6%; $p(X_2)$ – вероятность того, что пациентка здорова до проведения маммографии (априорная вероятность), равная 99%; $p(X_1|Y_1)$ – вероятность того, что пациентка больна, при условии положительного результата маммографии (апостериорная вероятность).

Видно, что апостериорная вероятность (то, что мы ищем) пропорциональна априорной вероятности (исходной) с несколько более сложным коэффициентом $\frac{p(Y_1|X_1) * p(X_1)}{p(Y_1|X_1) * p(X_1) + p(Y_1|X_2) * p(X_2)}$. Подчеркну еще раз. На мой взгляд, это фундаментальный аспект Байесовского подхода. Измерение (Y) добавило некоторое количество информации к первоначально имевшейся (априорной), что уточнило наше знание об объекте.

Примеры

Для закрепления пройденного материала попробуйте решить несколько задач.[4]

Пример 1. Имеется 3 урны; в первой 3 белых шара и 1 черный; во второй — 2 белых шара и 3 черных; в третьей — 3 белых шара. Некто подходит наугад к одной из урн и вынимает из нее 1 шар. Этот шар оказался белым. Найдите апостериорные вероятности того, что шар вынут из 1-й, 2-й, 3-й урны.

Решение. У нас есть три гипотезы: H_1 = {выбрана первая урна), H_2 = {выбрана вторая урна}, H_3 = {выбрана третья урна}. Так как урна выбирается наугад, то априорные вероятности гипотез равны: $P(H_1) = P(H_2) = P(H_3) = 1/3$.

В результате опыта появилось событие A = {из выбранной урны вынут белый шар}. Условные вероятности события A при гипотезах H_1, H_2, H_3: $P(A|H_1) = 3/4$, $P(A|H_2) = 2/5$, $P(A|H_3) = 1$. Например, первое равенство читается так: «вероятность вынуть белый шар, если выбрана первая урна равна 3/4 (так как всего шаров в первой урне 4, а белых из них – 3)».

Применяя формулу Бейеса, находим апостериорные вероятности гипотез:

$$\mathbf{P}(H_1|A) = \frac{(1/3)\,(3/4)}{(1/3)\,(3/4) + (1/3)\,(2/5) + (1/3)\cdot 1} = \frac{15}{43}$$

$$\mathbf{P}(H_2|A) = \frac{8}{43} \qquad \mathbf{P}(H_3|A) = \frac{20}{43}$$

Таким образом, в свете информации о появлении события A вероятности гипотез изменились: наиболее вероятной стала гипотеза H_3, наименее вероятной — гипотеза H_2.

Пример 2. Два стрелка независимо друг от друга стреляют по одной и той же мишени, делая каждый по одному выстрелу. Вероятность попадания в мишень для первого стрелка равна 0,8, для второго — 0,4. После стрельбы в мишени обнаружена одна пробоина. Найти вероятность того, что эта пробоина принадлежит первому стрелку (Исход {обе пробоины совпали} отбрасываем, как ничтожно маловероятный).

Решение. До опыта возможны следующие гипотезы: H_1 = {ни первый, ни второй стрелки не попадут}, H_2 = {оба стрелка попадут}, H_3 — {первый стрелок попадет, а второй — нет}, H_4 = {первый стрелок не попадет, а второй попадет). Априорные вероятности гипотез:

$P(H_1) = 0{,}2*0{,}6 = 0{,}12$; $P(H_2) = 0{,}8*0{,}4 = 0{,}32$; $P(H_3) = 0{,}8*0{,}6 = 0{,}48$; $P(H_4) = 0{,}2*0{,}4 = 0{,}08$.

[4] Задачи взяты из учебника Е.С.Вентцель, Л.А.Овчаров. Теория вероятностей и ее инженерные приложения.

Условные вероятности наблюденного события A = {в мишени одна пробоина} при этих гипотезах равны: $P(A|H_1) = P(A|H_2) = 0$; $P(A|H_3) = P(A|H_4) = 1$

После опыта гипотезы H_1 и H_2 становятся невозможными, а апостериорные вероятности гипотез H_3, и H_4 по формуле Бейеса будут:

$$P(H_3|A) = \frac{0{,}48\cdot 1}{0{,}48\cdot 1 + 0{,}08\cdot 1} = \frac{6}{7}$$
$$P(H_4) = 1 - P(H_3|A) = 1 - 6/7 = 1/7$$

Байес против спама

Формула Байеса нашла широкое применение в разработке спам-фильтров. Предположим, вы хотите обучить компьютер определять, какие из писем являются спамом. Будем исходить из словаря и словосочетаний, используя байесовские оценки. Создадим вначале пространство гипотез. Пусть относительно любого письма у нас есть 2 гипотезы: H_A – это спам, H_B – это не спам, а нормальное, нужное, письмо.[5]

Вначале «обучим» нашу будущую систему борьбы со спамом. Возьмем все имеющиеся у нас письма и разделим их на две «кучи» по 10 писем. В одну отложим спам-письма и назовем ее кучей H_A, в другую – нужную корреспонденцию и назовем ее кучей H_B. Теперь посмотрим: какие слова и словосочетания встречаются в спам- и нужных письмах и с какой частотой? Эти слова и словосочетания назовем свидетельствами и обозначим E_1, E_2... Выясняется, что общеупотребительные слова (например, слова «как», «твой») в кучах H_A и H_B встречаются примерно с одинаковой частотой. Таким образом, наличие этих слов в письме ничего не говорит нам о том, к какой куче его отнести (слабое свидетельство). Присвоим этим словам нейтральное значение оценки вероятности «спамности», скажем, 0,5.

Пусть словосочетание «разговорный английский» встречается всего в 10 письмах, причем чаще в спам-письмах (например, в 7 спам-письмах из всех 10), чем в нужных (в 3 из 10). Поставим этому словосочетанию для спама более высокую оценку 7/10, а для нормальных писем более низкую: 3/10. И наоборот, выяснилось, что слово «дружище» чаще встречалось в нормальных письмах (6 из 10). И вот мы получили коротенькое письмо: *«Дружище! Как твой разговорный английский?»*. Попробуем оценить его «спамность». Общие оценки $P(H_A)$, $P(H_B)$ принадлежности письма к каждой куче поставим, воспользовавшись несколько упрощенной формулой Байеса и нашими приблизительными оценками:

$P(H_A) = A/(A+B)$, *где* $A = p_{a1}*p_{a2}*...*p_{an}$, $B = p_{b1}*p_{b2}*…*p_{b}n = (1 – p_{a1})*(1 – p_{a2})*... *(1 – p_{an})$.

Таблица 1. Упрощенная (и неполная) Байес-оценка письма

	Куча H_A: спам-письма		Куча H_B: нужные письма	
Свидетельства:	Писем	Приблизительная оценка p_{ai}	Писем	Приблизительная оценка $p_{bi} = 1 – p_{ai}$
как (E1)		0,5		0,5
твой (E2)		0,5		0,5
дружище (E3)	4	0,4	6	0,6
разговорный английский(E4)	7	0,7	3	0,3
. . .				
A (произведение p_{ai})		0,07		
B (произведение p_{bi})				0,045
$P(H_i)$: вероятность принадлежности к куче	$P(H_A)$ = A/(A+B)=	0,609	$P(H_B)$ = B/(A+B)=	0,391

[5] По материалам заметки Сергея Козловского Что и как может Байес?

Таким образом, наше гипотетическое письмо получило оценку вероятности принадлежности с акцентом в сторону «спамности». Можем ли мы принять решение о том, чтобы бросить письмо в одну из куч? Выставим пороги принятия решений:

- Будем считать, что письмо принадлежит куче H_i, если $P(H_i) \geq T$.
- Письмо не принадлежит куче, если $P(H_i) \leq L$.
- Если же $L \leq P(H_i) \leq T$, то нельзя принять никакого решения.

Можно принять T = 0,95 и L = 0,05. Поскольку для рассматриваемого письма и $0,05 < P(H_A) < 0,95$, и $0,05 < P(H_B) < 0,95$, то мы не сможем принять решение, куда отнести данное письмо: к спаму (H_A) или к нужным письмам (H_B). Можно ли улучшить оценку, используя больше информации?

Да. Давайте вычислим оценку для каждого свидетельства другим способом, так, как это, собственно, и предложил Байес. Пусть:

F_a – это общее количество писем спама;
F_{ai} – это количество писем со свидетельством *i* в куче спама;
F_b – это общее количество нужных писем;
F_{bi} – это количество писем со свидетельством *i* в куче нужных (релевантных) писем.

Тогда: $p_{ai} = F_{ai}/F_a$, $p_{bi} = F_{bi}/F_b$. $P(H_A) = A/(A+B)$, $P(H_B) = B/(A+B)$, *где* $A = p_{a1}*p_{a2}*...*p_{an}$, $B = p_{b1}*p_{b2}*...*p_b n$

Обратите внимание – оценки слов-свидетельств p_{ai} и p_{bi} стали объективными и их можно вычислять без участия человека.

Таблица 2. Более точная (но неполная) Байес-оценка по наличным признакам из письма

	Куча H_A: спам-письма			Куча H_B: нужные письма		
Всего писем	F_a=	500		F_b=	30	
Свидетельства:		Писем	Условная вероятность $p_{ai}=F_{aei}/F_a$		Писем	Условная вероятность $p_{bi}=F_{bei}/F_b$
как (E1)	F_{ae1}=	250	0,5	F_{be1}=	15	0,5
твой (E2)	F_{ae2}=	250	0,5	F_{be2}=	15	0,5
дружище (E3)	F_{ae3}=	4	0,008	F_{be3}=	6	0,2
разговорный английский(E4)	F_{ae4}=	7	0,014	F_{be4}=	3	0,1
. . .						
A (произведение p_{ai})			0,000028			
B (произведение p_{bi})						0,005
$P(H_i)$: вероятность принадлежности к куче		$P(H_A)$ = A/(A+B)=	0,0056		$P(H_B)$ = B/(A+B)=	0,9944

Мы получили вполне определенный результат – с большим перевесом с вероятностью письмо можно отнести к нужным письмам, поскольку $P(H_B) = 0,997 > T = 0,95$. Почему результат изменился? Потому, что мы использовали больше информации – мы учли количество писем в каждой из куч и, кстати, гораздо более корректно определили оценки p_{ai} и p_{bi}. Определили их так, как это сделано у самого Байеса, вычислив условные вероятности. Другими словами, p_{a3} – это вероятность появления в письме слова «дружище» при условии того, что это письмо уже принадлежит спам-куче H_A. Результат не заставил себя ждать – кажется, мы можем принять решение с большей определенностью.

Байес против корпоративного мошенничества

Любопытное применение Байесовского подхода описал MAGNUS8.

В моем текущем проекте (ИС для выявления мошенничества на производственном предприятии) используется формула Байеса для определения вероятности фрода (мошенничества) при наличии/отсутствии нескольких фактов, косвенно свидетельствующих в пользу гипотезы о возможности совершения фрода. Алгоритм самообучаем (с обратной связью), т.е. пересчитывает свои коэффициенты (условные вероятности) при фактическом подтверждении или неподтверждении фрода при проверке службой экономической безопасности.

Стоит, наверное, сказать, что подобные методы при проектировании алгоритмов требуют достаточно высокой математической культуры разработчика, т.к. малейшая ошибка в выводе и/или реализации вычислительных формул сведет на нет и дискредитирует весь метод. Вероятностные методы особенно этим грешат, поскольку мышление человека не приспособлено для работы с вероятностными категориями и, соответственно, отсутствует «наглядность» и понимание «физического смысла» промежуточных и итоговых вероятностных параметров. Такое понимание есть лишь для базовых понятий теории вероятностей, а дальше нужно лишь очень аккуратно комбинировать и выводить сложные вещи по законам теории вероятностей — здравый смысл для композитных объектов уже не поможет. С этим, в частности, связаны достаточно серьезные методологические баталии, проходящие на страницах современных книг по философии вероятности, а также большое количество софизмов, парадоксов и задачек-курьезов по этой теме.

Еще один нюанс, с которым пришлось столкнуться — к сожалению, практически все мало-мальски ПОЛЕЗНОЕ НА ПРАКТИКЕ на эту тему написано на английском языке. В русскоязычных источниках в основном только общеизвестная теория с демонстрационными примерами лишь для самых примитивных случаев.

* * *

Полностью соглашусь с последним замечанием. Например, Google при попытке найти что-то типа «книги Байесовская вероятность», ничего внятного не выдал. Правда, сообщил, что книгу с байесовской статистикой запретили в Китае. (Профессор статистики Эндрю Гельман сообщил в блоге Колумбийского университета, что его книгу «Анализ данных с помощью регрессии и многоуровневых/иерархических моделей» запретили публиковать в Китае. Тамошнее издательство сообщило, что «книга не получила одобрения властей из-за различных политически чувствительных материалов в тексте».) Интересно, не аналогичная ли причина привела к отсутствию книг по Байесовской вероятности в России?

Консерватизм в процессе обработки информации человеком

Вероятности определяют степень неопределенности.[6] Вероятность, как согласно Байесу, так и нашей интуиции, составляет просто число между нулем и тем, что представляет степень, для которой несколько идеализированный человек считает, что утверждение верно. Причина, по которой человек несколько идеализирован, состоит в том, что сумма его вероятностей для двух взаимно исключающих событий должна равняться его вероятности того, что произойдет любое из этих событий. Свойство аддитивности имеет такие последствия, что мало реальных людей могут соответствовать им всем.

Теорема Байеса – это тривиальное последствие свойства аддитивности, бесспорное и согласованное для всех сторонников вероятностей, как Байеса, так и других. Один их способов написать это следующий. Если $P(H_A|D)$ – последующая вероятность того, что гипотеза А была после того, как данная величина D наблюдалась, $P(H_A)$ – его априорная вероятность до того, как наблюдалась данная величина D, $P(D|H_A)$ – вероятность того, что данная величина D будет наблюдаться, если верно H_A, а P(D) – безусловная вероятность данной величины D, то

(1) $P(H_A|D) = P(D|H_A) * P(H_A) / P(D)$

P(D) лучше всего рассматривать как нормализующую константу, заставляющую апостериорные вероятности составить в целом единицу по исчерпывающему набору взаимно исключающих гипотез, которые рассматриваются. Если ее необходимо подсчитать, она может быть такой:

$$P(D) = \sum_i P(D|H_i) \cdot P(H_i)$$

[6] Цитируется по Канеман, Словик, Тверски. Принятие решений в неопределенности: Правила и предубеждения, глава 25

Но чаще P(D) устраняется, а не подсчитывается. Удобный способ устранять ее состоит в том, чтобы преобразовать теорему Байеса в форму отношения вероятность–шансы.

Рассмотрим другую гипотезу, H_B, взаимно исключающую H_A, и изменим мнение о ней на основе той же самой данной величины, которая изменила ваше мнение о H_A. Теорема Байеса говорит, что

(2) $P(H_B|D) = P(D|H_B) * P(H_B) / P(D)$

Теперь разделим Уравнение 1 на Уравнение 2; результат будет таким:

$$\frac{P(H_A|D)}{P(H_B|D)} = \frac{P(D|H_A)}{P(D|H_B)} \bullet \frac{P(H_A)}{P(H_B)} \text{ или}$$

(3) $\Omega_1 = L \cdot \Omega_0$

где Ω_1 – апостериорные шансы в пользу H_A через H_B, Ω_0 – априорные шансы, а L – количество, знакомое статистикам как отношение вероятности. Уравнение 3 – это такая же соответствующая версия теоремы Байеса как и Уравнение 1, и часто значительно более полезная особенно для экспериментов, с участием гипотез. Сторонники Байеса утверждают, что теорема Байеса – формально оптимальное правило о том, как пересматривать мнения в свете новых данных.

Мы интересуемся сравнением идеального поведения, определенного теоремой Байеса, с фактическим поведением людей. Чтобы дать вам некоторое представление о том, что это означает, давайте попробуем провести эксперимент с вами как с испытуемым. Эта сумка содержит 1000 покерных фишек. У меня две такие сумки, причем в одной 700 красных и 300 синих фишек, а в другой 300 красных и 700 синих. Я подбросил монету, чтобы определить, какую использовать. Таким образом, если наши мнения совпадают, ваша вероятность в настоящее время, что выпадет сумка, в которой больше красных фишек – 0,5. Теперь, Вы наугад составляете выборку с возвращением после каждой фишки. В 12 фишках вы получаете 8 красных и 4 синих. Теперь, на основе всего, что вы знаете, какова вероятность того, что выпала сумка, где больше красных? Ясно, что она выше, чем 0,5. Пожалуйста, не продолжайте читать, пока вы не записали вашу оценку.

Если вы похожи на типичного испытуемого, ваша оценка попала в диапазон от 0,7 до 0,8. Если бы мы проделали соответствующее вычисление, тем не менее, ответ был бы 0,97. Действительно очень редко человек, которому предварительно не продемонстрировали влияние консерватизма, приходит к такой высокой оценке, даже если он был знаком с теоремой Байеса.

Если доля красных фишек в сумке – p, то вероятность получения r красных фишек и ($n – r$) синих в n выборках с возвращением – $p^r(1–p)^{n–r}$. Так, в типичном эксперименте с сумкой и покерными фишками, если H_A означает, что доля красных фишек составляет p_A и H_B – означает, что доля составляет p_B, тогда отношение вероятности:

(4) $$L = \frac{p_A^r(1-p_A)^{n-r}}{p_B^r(1-p_B)^{n-r}}$$

При применении формулы Байеса необходимо учитывать только вероятность фактического наблюдения, а, не вероятности других наблюдений, которые он, возможно, сделал бы, но не сделал. Этот принцип имеет широкое воздействие на все статистические и нестатистические применения теоремы Байеса; это самый важный технический инструмент размышления Байеса.

Байесовская революция

Ваши друзья и коллеги разговаривают о чем-то, под названием «Теорема Байеса» или «Байесовское правило», или о чем-то под названием байесовское мышление. Они действительно заинтересованы в этом, так что вы лезете в интернет и находите страницу о теореме Байеса и... Это уравнение. И все… Почему математическая концепция порождает в умах такой энтузиазм? Что за «байесианская революция» происходит в среде учёных, причем утверждается, что даже сам экспериментальный подход может быть описан, как её частный случай? В чём секрет, который знают последователи Байеса? Что за свет они видят?[7]

---

[7] Цитируется по Интуитивное объяснения теоремы Байеса.

Байесовская революция в науке произошла не потому, что все больше и больше когнитивных ученых внезапно начали замечать, что ментальные явления имеют байесовскую структуру; не потому, что ученые в каждой области начали использовать байесовский метод; но потому, что наука сама по себе является частным случаем теоремы Байеса; экспериментальное свидетельство есть байесовское свидетельство. Байесовские революционеры утверждают, что когда вы выполняете эксперимент и получаете свидетельство, которое «подтверждает» или «опровергает» вашу теорию, это подтверждение или опровержение происходит по байесовским правилам. Для примера, вы должны принимать во внимание не только то, что ваша теория может объяснить явление, но и то, что есть другие возможные объяснения, которые также могут предсказать это явление.

Ранее, наиболее популярной философией науки была [фальсификация Карла Поппера](#) – старая философия, которая была смещена байесовской революцией. Идея Карла Поппера, что теории могут быть полностью фальсифицированы, однако никогда не могут быть полностью подтверждены, это еще один частный случай байесовских правил; если $p(X|A) \approx 1$ – если теория делает верные предсказания, тогда наблюдение ~X очень сильно фальсифицирует A. С другой стороны, если $p(X|A) \approx 1$ и мы наблюдаем X, это не очень сильно подтверждает теорию; возможно какое-то другое условие B, такое что $p(X|B) \approx 1$, и при котором наблюдение X не свидетельствует в пользу A но свидетельствует в пользу B. Для наблюдения X определенно подтверждающего A, мы должны были бы знать не то, что $p(X|A) \approx 1$, а что $p(X|\sim A) \approx 0$, что мы не можем знать, поскольку мы не можем рассматривать все возможные альтернативные объяснения. Например, когда эйнштейновская теория общей относительности превзошла ньютоновскую хорошо подтверждаемую теорию гравитации, это сделало все предсказания ньютоновской теории частным случаем предсказаний эйнштейновской.

Похожим образом, попперовское заявление, что идея должна быть фальсифицируема может быть интерпретировано как манифестация байесовского правила о сохранении вероятности; если результат X является положительным свидетельством для теории, тогда результат ~X должен опровергать теорию в каком-то объеме. Если вы пытаетесь интерпретировать оба X и ~X как «подтверждающие» теорию, байесовские правила говорят, что это невозможно! Чтобы увеличить вероятность теории вы должны подвергнуть ее тестам, которые потенциально могут снизить ее вероятность; это не просто правило, чтобы выявлять шарлатанов в науке, но следствие из теоремы байесовской вероятности. С другой стороны, идея Поппера, что нужна только фальсификация и не нужно подтверждение является неверной. Теорема Байеса показывает, что фальсификация это очень сильное свидетельство, по сравнению с подтверждением, но фальсификация все еще вероятностна по своей природе; она не управляется фундаментально другими правилами и не отличается в этом от подтверждения, как утверждает Поппер.

Таким образом, мы обнаруживаем, что многие явления в когнитивных науках, плюс статистические методы, используемые учеными, плюс научный метод сам по себе – все они являются частными случаями теоремы Байеса. В этом и состоит Байесовская революция.

Добро пожаловать в Байесовский Заговор!

Литература по Байесовской вероятности

1. Рекомендую начать с относительно большого, но более-менее понятного и доступного Интуитивного объяснения теоремы Байеса.

2. Очень много различных применений Байеса описывает нобелевский лауреат по экономике Канеман (со товарищи) в замечательной книге Принятие решений в неопределенности: Правила и предубеждения. Только в моем кратком конспекте этой очень большой книги я насчитал 27 упоминаний имени пресвитерианского священника. Минимум формул.

3. Глава 6. Ложная положительность и положительная ложность из книги Леонард Млодинов. (Не)совершенная случайность. Как случай управляет нашей жизнью. Ни одной формулы. Кстати, примеры на тему Байесовской вероятности встречаются не только в упомянутой главе.

4. Глава 10. Кое-что о Байесе, статистике и подходах из книги Дуглас Хаббард. Как измерить всё, что угодно. Оценка стоимости нематериального в бизнесе.

5. Раздел Условная вероятность. Теорема Байеса из книги Левин (со товарищи). Статистика для менеджеров с использованием Microsoft Excel.

6. Раздел Теорема гипотез (формула Байеса) из книги Вентцель, Овчаров. Теория вероятностей и ее инженерные приложения. Кратко и по делу. Формулы и несколько примеров с решениями.

7. Глава 5. Наше восприятие мира — это фантазия, совпадающая с реальностью из книги Криса Фрита. Мозг и душа.

8. Байесовский вывод. Коротенькая заметка в Викинауке – свободной научной энциклопедии на русском языке.

8. Единственная найденная мною книга по теме (издана еще в СССР): Дж. Хей. Введение в методы байесовского статистического вывода. Не читал (планирую). В электронном виде можно найти в инете.