Интуитивное объяснение теоремы Байеса

Ричард Фейнман однажды сказал, что если бы ядерная война привела человечество к потере всех знаний и людям пришлось бы начинать с нуля, но он как-то мог бы передать им один-единственный кусок нынешних научных знаний, он бы сказал им следующее: Все вещи сделаны из атомов – маленьких частиц, что перемещаются вокруг в постоянном движении, притягиваясь на малом расстоянии друг к другу, но отталкиваясь на слишком малом. Для Фейнмана, это был единственный наиболее полезный и важный кусок информации, который мы могли бы передать будущему человечеству, которое потеряло все свои знания. Это превосходный выбор, особенно в рамках редукционизма.[1]

После некоторых размышлений, я решил[2], что вторым куском информации, который можно было бы передать будущему обществу, была бы теорема Байеса. Когда вы смотрите на мир через призму байесовской теоремы это похоже на то, будто вы видите сквозь Матрицу. Ничто не будет тем же самым, после того как вы поймете теорему Байеса. Однако я не буду просто давать уравнение и объяснять его части, поскольку если вы не поймете логику, лежащую за данным уравнением, то не поймете, как его правильно применять. Цель данного туториала – это не научить вас, как подобрать учительский пароль и дать верные ответы на экзамене. Нет, цель – дать вам настоящее понимание теоремы Байеса, так, чтобы суметь правильно применять ее в условиях обыденной жизни, которая проходит за пределами экзаменационного класса. К концу туториала вы не просто сможете рассказать теорему Байеса наизусть; вы пропитаетесь ею до мозга костей.

Наиболее популярный онлайн-туториал по теореме Байеса, "Интуитивное объяснение теоремы Байеса" Элизера Юдковски, начинается так: 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. Просто одно уравнение. Найденная страница дает определение теоремы, но не рассказывает, что это, чем полезно и почему ваши друзья этим интересуются. Это похоже на какую-то случайную штуку из статистики.

Итак, вы заходите на эту страницу. Может быть, вы не понимаете о чём говорит это уравнение. Может быть, понимаете в теории, но каждый раз, когда вы пытаетесь применить формулу на практике, путаетесь, пытаясь вспомнить разницу между $p(a|x)$ и $p(x|a)$, и должно ли $p(a)*p(x|a)$ находиться в числителе или знаменателе. Может быть вы видите теорему, понимаете теорему, можете использовать теорему, но вы не можете понять почему вашим друзьям и/или коллегам кажется, что в ней заложен секрет вселенной. Может быть, все ваши друзья носят футболки с теоремой Байеса, и вы чувствуете себя одиноким. Может быть вы девушка, которая ищет парня, но парень, который вам нравится, отказывается встречаться с «не байесианцами». Единственное, что важно — это то, что Байес крут, а если вы не знаете Байеса, то вы не круты.

Почему математическая концепция порождает в студентах 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?

Скоро вы узнаете. Скоро вы будете одним из нас.

---

[1] Редукционизм (от лат. reductio — возвращение, приведение обратно) — методологический принцип, согласно которому сложные явления могут быть полностью объяснены с помощью законов, свойственных явлениям более простым (например, социологические явления объясняются биологическими или экономическими законами).

[2] Настоящий текст является переводом © Remlin. Оригинал написал Luke Muehlhauser. Любопытно, что оригинал является перепевом другого оригинала, написанного Eliezer Yudkowsky. Ну, а над переводом творчески поработал я ☺

Объяснение Элизером этого весьма важного закона вероятности, возможно, лучшее в интернете, однако я боюсь, что оно может быть все еще слишком сложным для тех, кто не помнит алгебру со школы. Элизер зовет свое объяснение "мучительно плавным", однако скорее всего измеряет "плавность" с точки зрения тех людей, кто читал Фейнмана в девять лет и проводил вычисления в 13, то есть для таких же, как сам Элизер. Так что я решил написать еще более плавное введение в теорему Байеса. Достаточно плавное для обычных людей.

Есть моменты, когда Юдковски вводит новые термины без объяснения их смысла (например "пересмотренная вероятность"). В других моментах он оставляет вас с трудной проблемой без ресурсов, нужных для ее решения (например, проблема, поставленная прямо перед фразой "означает пересмотренная вероятность"). В этих случаях, я полагаю, многие люди, не особо дружащие с математикой, просто сдаются и перестают читать дальше. Если у вас не вышло прочитать объяснение Элизера, я надеюсь, что вы попробуете прочитать мое. Оно плавнее.

И если вы застряли где-то в данной статье, пожалуйста, отметьте это где-то в комментариях или отпишитесь мне на электронную почту и объясните где точно вы запнулись. Это будет для меня знаком, что в определенном месте статью следует сделать еще плавнее.

Из-за того, что эта статья плавнее, чем у Юдковски, она также длиннее. Так что я советую вам изучать по одному параграфу в день. Перечень параграфов приведен ниже. В данном введении интерактивные элементы из объяснения Юдковски были заменены на ряды изображений, так что вы можете читать это объяснение и на мобильных устройствах, таких как Kindle. Здесь можно скачать PDF (ссылка).

Надеюсь, что вы найдете это полезным!

Содержание:

Скорее всего вы уже использовали байесовское мышление, хоть и не знали об этом. Обсудим пример, который я взял у Нэйла Мэнсона: Вы солдат в битве, который прячется в окопе. Вы точно знаете, что на поле боя остался только один вражеский солдат, на расстоянии примерно в 400 ярдов. Вы также знаете, что если это обычный солдат, то он не сможет попасть в вас с такого расстояния. Однако если этот солдат - снайпер, то вполне возможно, что он сможет в вас попасть. Но снайперов в армии врага мало, так что скорее всего это обычный солдат. Вы приподнимаете голову из окопа, стараясь оглядеться получше. Бам! Пуля чиркает по вашей каске и вы падаете обратно в окоп.

Хорошо, думаете вы. Я знаю, что снайперы редки, однако этот парень попал в меня с четыреста ярдов. Все еще велик шанс, что это обычный солдат, однако шанс того, что это снайпер, уже выше, ведь он попал в меня с такого большого расстояния. Спустя несколько минут вы осмеливаетесь выглянуть еще раз и приподнимаете голову над окопом. Бам! Вторая пуля чиркает по вашей каске! Вы падаете обратно.

Вот черт, думаете вы. Это определенно снайпер. Не имеет значения, насколько они редки, однако обычный солдат не может два раза подряд попасть с такого расстояния. Это определенно снайпер. Лучше мне вызвать подмогу. Если это грубое приближение того, как бы вы думали в подобной ситуации, тогда мои поздравления! Вы уже размышляете как байесианец, по крайней мере иногда.

Но конечно было бы полезнее быть поточнее, чем описано выше, и было бы полезнее знать когда и как наше мышление может зависеть от того, насколько правильно мы выполняем байесовские вычисления. На деле, в научном изучении искажений в человеческом мышлении, искажение определяется как систематическое отклонение от идеального байесовского мышления. Так что хватит мыслить Байесовски случайным образом. Давайте учиться применять данный навык сознательно.

Мы начнем с проблемы, похожей на те, с которыми вы имели дело в школе. Однако я обещаю вам, изучение теоремы Байеса будет намного полезнее почти всего изученного вами в школе. Итак, вот проблема. 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. 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 нет, я не преувеличиваю. Можете посмотреть: Casscells et. al. (1978), Eddy (1982), Gigerenzer & Hoffrage (1995). Но возьмите калькулятор и посмотрите, можете ли вы получить верный ответ. Математика здесь проста, однако есть хитрость.

Ладно. Вы же попытались, да? У вас не получится выучить теорему Байеса только читая. Теорему Байеса можно выучить только применяя ее. Так что вам действительно надо попытаться сделать все предлагаемые упражнения. Давайте еще раз. Попытайтесь. Я все еще буду тут когда вы закончите. Какой ответ вы получили? Большинство врачей дают оценку между 70% и 80%, но это очень далеко от правды.

Давайте рассмотрим более легкую версию той же проблемы. При ней, около половины врачей дает правильный ответ. В ходе проведения плановых медицинских осмотров установлено, что в сорокалетнем возрасте 100 женщин из 10 000 болеет раком молочной железы. 80 женщин из 100 больных раком получают положительные результаты маммографии. 950 женщин из 9900 здоровых также получают положительные результаты. Если проведён осмотр 10 000 женщин данной возрастной группы, какая доля женщин с положительным результатом маммографии действительно болеет раком молочной железы?

Попробуйте. Какой ответ?

7,8%. Только 7,8% женщин с положительной маммограммой будут иметь рак груди! Итак, следите за логикой... Всегда начинайте с определения того, что вам известно. В этом случае мы хотим знать долю (или процент) женщин с положительной маммограммой, которые имеют на самом деле рак груди. Сначала, давайте определим, как много женщин имеют положительные маммограммы. Это будет знаменатель нашей дроби. В исходных данных говорится, что 950 из 9900 женщин не имеющих рака груди будут иметь положительные маммограммы. Так что эти 950 женщин с положительным тестом приплюсовываем сюда. Также говорится, что 80 из 100 женщин, у которых есть рак груди, также получат положительный результат теста. Вместе с этими 80 у нас получается 80 + 950 = 1030 женщин, имеющих положительный результат теста.

Хорошо. У нас есть половина нашей дроби. Теперь как найти числитель? Как много из этих 1030 женщин с положительным результатом на самом деле имеют рак груди? Так, в задаче сказано что 80 из 100 женщин с раком груди получают положительные результаты, так что наш числитель равен 80. Отношение женщин, у кого действительно рак груди, к тем, кто имеет положительный результат маммограммы, составляет 80/1030, то есть 7,8%.

Так что, если одна из этих сорокалетних женщин получает положительную маммограмму, и доктору известно о вышеприведенной статистике, то он должен сказать женщине, что у нее только 7,8% вероятности того, что она имеет рак груди, несмотря на положительную маммограмму. Это куда менее пугающе, для женщины, нежели доктор скажет, что шанс равен 70–80%, как делает большинство докторов!

Итак, мы могли видеть, как осторожные рассуждения данного вида могли приводить к совершенно реальным последствиям. Это не "просто математика". Почему даже врачи, которым нужно решить эту задачу, ошибаются так часто? Вычисления здесь не такие сложные, так в чем же дело? Наиболее общая ошибка в концентрации только на тех женщинах с раком груди, которые получают положительный результат, одновременно забывая о другой важной информации, такой как исходная доля женщин с раком груди и доле женщин без рака груди, которые также получили положительные результаты.

Но вам всегда требуются все три куска информации, чтобы дать верный ответ. Чтобы ощутить, почему нужны всегда три куска информации, представьте альтернативную вселенную, в которой только одна женщина на миллион имеет рак груди. И давайте скажем, что тест маммографии определяет рак груди в восьми из десяти случая, давая фальшивый положительный результат только в десяти процентах случаев.

Теперь, я думаю вы можете видеть, что в этой вселенной начальная вероятность что женщина имеет рак груди настолько исчезающе мала, что даже если женщина получает положительный результат теста, все еще скорее всего можно сказать, что у нее нет рака груди.

Почему? Потому что куда больше женщин с фальшивым положительным результатом (10% из 99,9999% женщин), чем тех, для кого положительный результат верен (80% из 0,0001% женщин). Так что, если женщина получает положительный результат, это скорее всего фальшивый результат, а не настоящий.

Пример с преувеличениями, такой как этот, показывает, что новые данные, полученные с маммографии, не заменяют данные, которые были у вас с самого начала о том, насколько вероятно то, что женщина имеет рак груди. Вместо этого, представляйте, что вы начали с исходной вероятностью, что у женщины рак груди, и затем получаете новое свидетельство в виде теста маммографии, которое сдвигает вероятность в том или ином направлении с начального значения, в зависимости от того, положителен результат или отрицателен. Таким образом, тест маммографии сдвигает вероятность того, что женщина имеет рак груди в сторону результата теста.

Чтобы проиллюстрировать это, рассмотрим снова исходную задачу. В ее условии, 1% сорокалетних женщин (100 из 10 000) имеют рак груди. 80% женщин с раком (80 из 100) получают положительный тест, и 9,6% женщин без рака (950 из 9900) также получают положительный результат. Когда мы проводим вычисления, мы обнаруживаем, что положительный результат теста сдвигает шансы женщины на наличие рака груди от 1% до 7,6%.

Вы не можете заменить исходную вероятность новой информацией. Вы можете только дополнить ее новой информацией, путем сдвига ее в том или ином направлении от исходной вероятности. Исходная вероятность все еще имеет значение. Помните, что нам всегда нужны все три куска информации. Нам нужно знать исходную долю женщин с раком, долю женщин с раком и положительным результатом теста и долю женщин без рака, но с положительным результатом.

Чтобы понять, почему нужен последний кусок информации, доля женщин без рака, но с положительным результатом – рассмотрим туже задачку, но с новыми условиями. Маммография – имеет тот же процент фальшивых отрицательных результатов, что и ранее, но также она имеет невероятно высокий уровень фальшивых положительных результатов: 80%!

Итак, условие задачи: 1% женщин больны раком груди. 80% женщин с раком груди получают положительный результат теста. 80% женщин без рака груди также получают положительный результат теста. Женщина имеет положительный результат маммографии. Какова вероятность, что она на самом деле имеет рак груди?

Теперь прекратите чтение и посчитайте ответ. Получилось?

Хорошо, начнем вычислять, какой процент женщин получит положительный результат теста. 80% от 1% женщин с раком получат положительный результат, то есть это 0,8%. Также, 80% из 99% женщин без рака груди также получат положительный результат, это 79,2%. Складываем 0,8% + 79,2% = 80%, то есть 80% женщин получат положительный результат теста.

Хотя больных женщин на самом деле только 1%! Так что уже можно сказать, что третий кусок информации может дать значительную разницу. Но все же давайте закончим вычисления. Какова доля женщин с положительным результатом маммографии на самом деле больны?

Первое – сколько женщин всего получит положительный результат? Это наш знаменатель. Так, есть две группы женщин с положительным результатом: те, у кого на самом деле рак груди (0,8%) и те, у кого нет рака груди (79,2%). Складываем вместе, получаем, что наш знаменатель равен 80%.

Теперь пора определиться с числителем. Сколько из этих 80% женщин с положительным результатом на самом деле больны раком? Мы уже знаем ответ, так как знаем процент женщин с раком груди и положительным результатом: 0,8%. Так что доля женщин, которые имеют положительный результат и на самом деле больны раком, составляет 0,8%/80%, то есть 1%

Женщина начала с вероятности иметь рак в 1% и после теста она все еще имеет 1% шанс на наличие рака. Как так получилось? То есть тест нам ничего не сказал вообще? Неа. Почему он нам ничего не сказал? Вспомните, тест маммографии имеет такой высокий уровень фальшивых положительных результатов, что женщина может получить положительный результат в независимости от того, есть у нее рак груди или нет! Если у нее есть рак, у нее 80% вероятность получения положительного результата теста. И если она здорова, у нее так же 80% вероятности получить тот же результат.

Вот почему тест не сказал нам ничего. Поскольку мы изменили вероятности. Скорее всего женщина получит один и тот же результат в любом случае, так что тест не скажет нам какая возможность правильна. В таком случае, маммография никак не коррелирует с наличием рака груди, поскольку дает одинаковые результаты во всех случаях. На деле, в этом случае нет причин называть один результат "положительным", а другой "отрицательным", поскольку ни один из них не сдвигает вашу вероятность в каком-либо направлении.

Что означает, что с тем же успехом в качестве "теста" можно использовать подкидывание монетки. Оно будет в той же степени не коррелировать с частотой случаев заболевания раком груди. Если у женщины есть рак груди, у монетки 50% шансов выпасть орлом, и если у нее нет рака груди, у монетки также 50% шансов выпасть орлом.

Или вы можете просто использовать тест, всегда дающий одинаковый результат. Пусть вашим тестом будет 2+2. Если у женщины есть рак груди, то значение теста равно четырем. И если у нее нет рака груди, значение теста равно четырем.

Все эти тесты одинаково плохи, поскольку они дают одинаковый результат с одним и тем же распределением все время, в независимости от того, больна ли женщина или нет. Чтобы сделать тест, который даст нам информацию, которую мы можем использовать для дополнения вероятности что у женщины есть рак груди, тест должен коррелировать с наличием рака груди хоть каким-либо способом. Тест должен выдавать разные результаты, в зависимости от того, больна ли женщина, или нет. Тогда он будет являться настоящим тестом на рак груди.

Но помните: вероятность существует в вашем сознании, не в реальности. Даже полезный тест маммографии на самом деле не меняет тот факт, больна ли женщина раком или нет на самом деле. Рак у нее либо есть, либо нет. Это не реальность не уверена о наличии рака у женщины. Это мы не уверены, есть ли у нее рак. Это наша информация, наше суждение не определено, а не реальность.

Исходная доля женщин с раком груди известна нам в качестве *априорной* вероятности. Это вероятность что у женщины есть рак груди, априорна по отношению к каким-либо новым свидетельствам, которые мы получаем. А что насчет доли женщин с раком груди и положительным результатом, и долей женщин без рака груди и положительным результатом? Это было в условиях нашей задачи, так что эти вероятности называют *условными*. Вместе, априорная и условные вероятности составляют наши исходные данные. Они – кусочки информации которые нам нужно знать для вычисления результата, который называется пересмотренной или *апостериорной* вероятностью.

Как показано выше, если две условные вероятности равны друг другу - если положительный результат получается в 80% случаев в независимости от того больна ли женщина – апостериорная вероятность равна исходной (априорной) вероятности.

Откуда мы берем исходные данные? Как нам узнать, какова исходная и условные вероятности? Их мы берем, тестируя реальность, как и все остальное. Для примера, если вы думаете, что 100 из 10 000 женщин имеют рак груди, а на самом деле только 50 из них, тогда ваши исходные данные неправильны и вам нужно провести больше исследований. Также есть несколько легких обозначений которые вы должны знать, поскольку они нужны для работы с подобными задачами.

Для наглядности, Элизер приводит задачу о пластиковых яйцах: Предположим, что бочка содержит множество маленьких пластиковых яиц. Некоторые из них окрашены в красный, а некоторые в синий. 40% яиц содержат жемчужины, а 60% пусты. 30% яиц, содержащих жемчуг, окрашены в синий цвет, и 10% пустых яиц также синие. Какова вероятность того, что синее яйцо содержит жемчуг?

Перед тем как вы приступите к решению данной проблемы, давайте рассмотрим символы о которых я говорил. Чтобы сказать, что "вероятность того, что в яйце содержится жемчужина равна 0,4" мы пишем: p(жемчуг) = 0,4. Теперь, следующая запись: p(синее|жемчуг) = 0,3. Что значит линия между "синее" и "жемчуг"? Это означает "учитывая, что". И здесь, слово "синий" ставится в начале. Так что мы можем прочитать приведенное выше обозначение как: "Вероятность что яйцо синее, если в нем есть жемчуг, равна 0,3". Еще одним символом является тильда:~ Это значит отрицание, то есть: p(синий|~жемчуг) = 0,1. Читается: "Вероятность что яйцо синее, учитывая то, что оно не содержит жемчуг, равна 0,1". Теперь мы готовы выделить три кусочка информации из задачи выше, но в символической форме:

p(жемчуг) = 0,4

p(синее|жемчуг) = 0,3

p(синее|~жемчуг) = 0,1

И, конечно, то что мы ищем:

p(жемчуг|синее) = ?

Вам следует прочитать эти обозначения вслух:

Вероятность что данное яйцо содержит жемчуг равна 0.4.

Вероятность что данное яйцо синее, учитывая то, что оно содержит жемчуг, равна 0.3.

Вероятность что данное яйцо синее, учитывая, что в нем нет жемчуга, равна 0.1.

Вероятность что данное яйцо содержит жемчуг, учитывая, что оно синее, равна...чему?

Это наша задача. Остановитесь и попробуйте решить ее, не подглядывая далее.

Каково же решение? Мы ищем вероятность, что яйцо содержит жемчуг, учитывая то, что оно синее. (Похоже на попытку определить вероятность что женщина имеет рак груди, учитывая, что у нее положительный результат маммографии)

40% яиц содержат жемчуг, и 30% из них синие, так что 12% яиц одновременно и содержат жемчуг и синие.

60% яиц не содержат жемчуг, и 10% из них синие, так что 6% яиц синие и не содержат жемчуг.

12% + 6% = 18%, так что всего доля синих яиц - 18%.

Мы уже знаем, что 12% яиц синие и содержат жемчуг, так что шанс того, что в синем яйце содержится жемчуг равен 12/18 или около 67%.

Один из самых известных случаев ошибки при применении Теоремы Байеса связан с британской женщиной Салли Кларк. После того, как ее двое детей умерли от синдрома внезапной детской смерти, ее арестовали и осудили за убийство ее детей. Педиатр Рой Мидоу оценил шансы что оба ребенка умрут одновременно от данного синдрома как один к 73 миллионам. Он получил такое число, возведя в квадрат шанс одного ребенка умереть от данного синдрома, который равен 1 к 8500

Основываясь на данном свидетельстве, Салли Кларк был вынесен приговор. Королевское сообщество статистиков публично осудило "это неправильное использование статистических данных в суде", но первая апелляция Салли Кларк была отклонена. Ее освободили после 4 лет в женской тюрьме, где все считали, что она убила своих детей. Она никогда не оправилась от этого, впала в зависимость от алкоголя и умерла от алкогольной интоксикации в 2007 году.

Статистическая ошибка, сделанная Роем Мидоу, была, помимо всего прочего, провальной при игнорировании исходной вероятности что Салли Кларк убила своих детей. В то время как смерть сразу двоих детей от синдрома внезапной смерти довольно редкое явление, однако мать, убивающая двух своих детей, значительно более редкое явление.

Может быть полезно представить происходящее. На странице Юдковского, где приведено объяснение Теоремы Байеса, есть интерактивная панель, которая позволяет вам изменять каждую из трех величин независимо от других и смотреть как это отображается на результате. Но это работает только если у вас есть Java и не на Маке (по крайней мере, не на моем) или на мобильных устройствах таких как Kindle, так что здесь я собираюсь использовать картинки. (На деле, это скриншоты интерактивной панели Юдковского).

Сначала, исходная проблема (рис. 1). Предположим, что бочка содержит множество маленьких пластиковых яиц. Некоторые из них окрашены в красный, а некоторые в синий. 40% из яиц содержат жемчужины, а 60% пусты. 30% яиц, содержащих жемчуг, окрашены в синий цвет, и 10% пустых яиц также синие. Какова вероятность того, что синее яйцо содержит жемчуг?

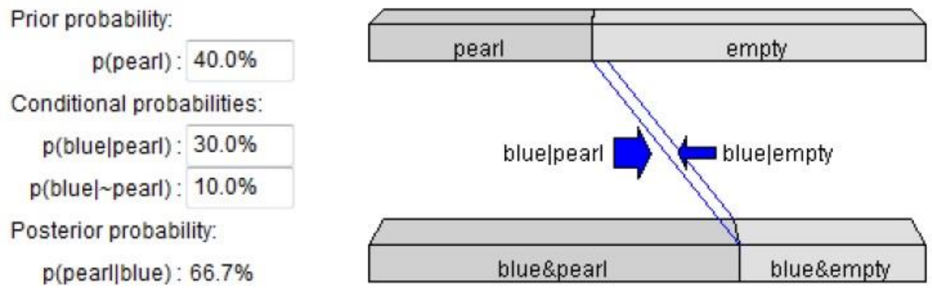

Рис. 1. Как априорная вероятность под воздействием условной вероятности переходит в апостериорную вероятность

Полоска вверху, разделенная между жемчугом и пустыми, показывает исходную (априорную) вероятность что яйцо содержит жемчуг. Вероятность найти жемчуг равна 40%, так что разделитель верхней полоски немного сдвинут влево от центра (центр означал бы 50%).

Первая условная вероятность – это p(синее|жемчуг) или "вероятность, что яйцо синее учитывая, что оно содержит жемчуг". Стрелка с левой стороны отражает размер этой вероятности. Вторая условная вероятность – это "вероятность что яйцо синее, учитывая, что оно не содержит жемчуг". Стрелка с правой стороны заметно меньше.

Нижняя строка показывает вероятности что яйцо содержит или нет жемчуг, учитывая, что оно синее. Может смущать тот факт, что полоски вверху и внизу не равны друг другу, хотя нарисованы они одного размера. Верхняя отражает все яйца, как синие, так и красные. А нижняя отражает только синие. Но не дайте этому смутить вас; подобный рисунок позволяет нам наиболее ясно проиллюстрировать эффекты каждой из трех вероятностей: исходной, первой условной и второй условной. Обе полоски, нижняя и верхняя, рисуются нами для отображения шансов, что яйцо содержит жемчуг. Это то же самое, что мы "запускаем тест", берем яйцо и оно оказывается синее, так что устраняет из рассмотрения остальные яйца.

Перекос линии посередине отражает то, как мы должны пересмотреть нашу вероятность что яйцо содержит жемчуг после нашего первого теста (то есть синее ли яйцо). С начала мы знаем, что если мы возьмем яйцо с корзины, есть 40% вероятность, что в нем есть жемчуг. Теперь же мы взяли яйцо и видим, что оно синее. Так как мы знаем, что шансы того, что яйцо будет синим, если в нем есть жемчуг, были выше, чем шансы на то, что оно было бы синим без жемчуга, мы таким образом можем заключить, что теперь вероятность того, что в яйце есть жемчуг, куда выше, нежели до этого. Так что мы сдвигаем нашу вероятность, что в яйце есть жемчуг, в сторону увеличения. Вот почему линия, соединяющая верхнюю и нижнюю полоски изображена наклоненной вправо. Знание насколько надо сдвинуть нашу вероятность, разумеется, потребует некоторых вычислений.

Теперь, давайте посмотрим на эффект изменения апостериорной вероятности если априорная вероятность меняется. Что если только 10% яиц содержат жемчуг (рис. 2)? Наши условные вероятности не изменили, так что относительный наклон линии посередине не изменился. Таким образом, степень, в которой мы пересматриваем вероятность что яйцо содержит жемчуг, если мы знаем, что оно синее – эта степень не требует изменения. Однако исходная вероятность стала намного ниже, что означает, что и апостериорная вероятность соответственно станет ниже.

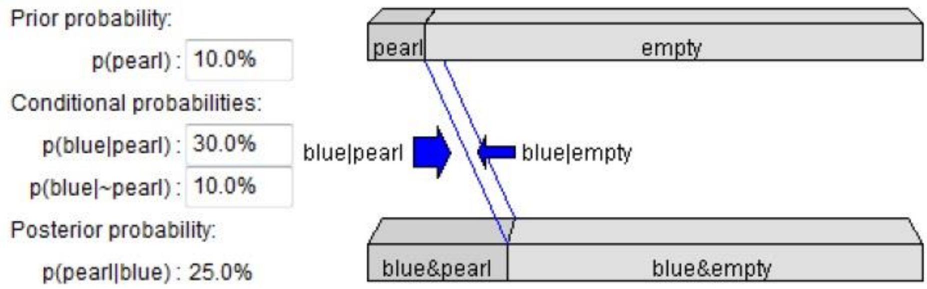

Рис. 2. Три вероятности: априорная, условная, апостериорная; по сравнению с рис. 1 изменена априорная вероятность

Помните, как изменилась вышеприведенная задача о женщинах с раком груди, когда в условии изменили априорную вероятность на то, что только одна женщина из миллиона больна раком? Если мы нарисуем для этой задачи диаграмму, похожую на ту, что приведена выше, наклонная линия будет завалена напротив левой границы рисунка и апостериорная вероятность что у женщины рак будет очень малой, вследствие чего результат теста почти не будет иметь значения, каковы бы не были условные вероятности. (Вам нужно иметь действительно сильно наклоненную линию условных вероятностей, чтобы пересмотреть что-либо далекое от левой границы диаграммы.

И что же происходит если мы сохраняем условные вероятности такими же, но увеличиваем априорную вероятность до 80% (рис. 3)? Разумеется, раз априорная вероятность стала много больше, то апостериорная вероятность также станет больше. И опять, степень в которой мы пересматриваем остается той же самой, так что линия все еще наклонена вправо. Но заметим, что точное количество величины, на которую пересматривается вероятность, не та же самая. Наклон линии не такой, какой был при вероятности в 40%, и не такой, какой был на вероятности в 10%. Почему так? Это потому, что величина, на которую нам нужно пересмотреть нашу вероятность после обнаружения что яйцо синее, зависит не только от разницы между двумя условными вероятностями (в этом случае, 30% и 10%), но также и от априорной вероятности. Данное свойство вытекает просто из вычислений. И если вы как следует это обдумаете, это станет для вас очевидно. Что если априорная вероятность что яйцо содержит жемчуг равна 99.999%, а условные вероятности такие же? Если это бы случай, когда мы бы пересматривали на ту же величину что и ранее, тогда вероятность того что в синем яйце есть жемчуг была бы больше ста процентов! Наклонная линия вышла бы за правую границу рисунка!

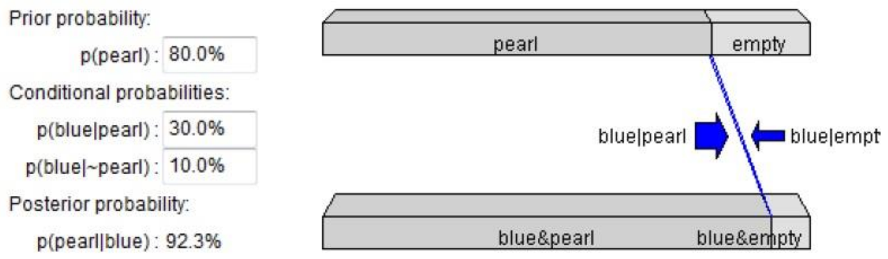

Рис. 3. Новое значение априорной вероятности

Если это происходит, это значит, что вы где-то ошиблись в вычислениях. Сдвиг априорной вероятности на 99% просто делает величину пересмотра, которую нам надо сделать очень маленькой в абсолютных величинах, так как нам все еще нужно отрегулировать изменение, но вероятность что яйцо содержит жемчуг не может стать много больше, чем она уже есть (рис. 4).

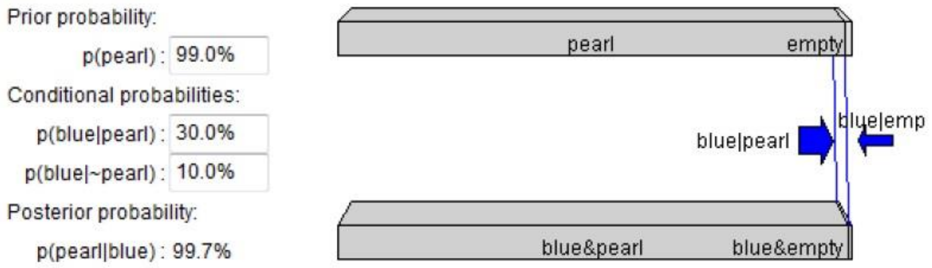

Рис. 4. При экстремально большой априорной вероятности апостериорная вероятность должна оставаться ниже границы в 100%

Что если мы вернем априорную вероятность к ее исходному значению в 40%, но изменим первую условную вероятность? Сейчас первая условная вероятность оказывает значительно большее влияние на апостериорную вероятность, и наша линия в большей степени наклоняется вправо. Это значит, что мы должны пересмотреть нашу вероятность в сторону того, что яйцо скорее всего содержит жемчуг, если мы знаем, что оно синее.

Почему так? Ну, первая условная вероятность – это p(синее|жемчуг) или "вероятность что яйцо синее, учитывая, что оно содержит жемчуг". Что будет, если данная вероятность будет больше чем вторая условная вероятность, "вероятность что яйцо синее, учитывая то, что в нем нет жемчуга" (рис. 5)? Если так, тогда там будет куда больше синих яиц, содержащих жемчуг, нежели синих яиц без жемчуга. Так что если вы обнаружили, что вытащенное яйцо синее, вы знаете что куда вероятнее то, что в нем есть жемчуг… поскольку синих яиц с жемчугом куда больше, нежели пустых синих яиц.

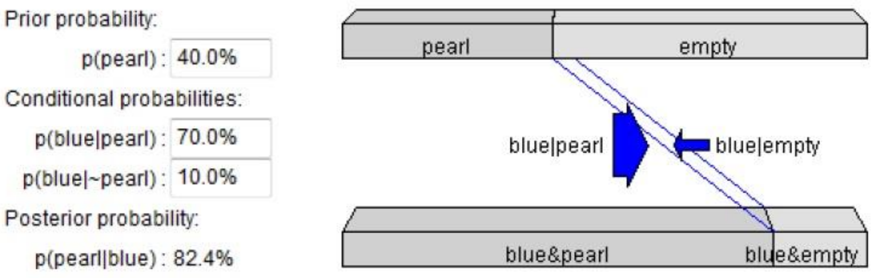

Рис. 5. Условная вероятность изменилась, и теперь 70% яиц, содержащих жемчуг, окрашены в синий цвет

И опять, именно разница между двумя условными вероятностями определяет насколько нам нужно пересмотреть нашу вероятность в зависимости от результата теста. Если разница между двумя условными вероятностями мала, нам не придется слишком сильно менять априорную вероятность. Если же она велика, то нам надо будет пересмотреть априорную вероятность в значительной степени.

Давайте обратим побольше внимания на разницу между двумя условными вероятностями, которая определяет степень, на которую мы меняем нашу вероятность, а не на их абсолютные величины, давайте посмотрим, что случится если обе условные вероятности очень высоки, но не слишком различаются (рис. 6).

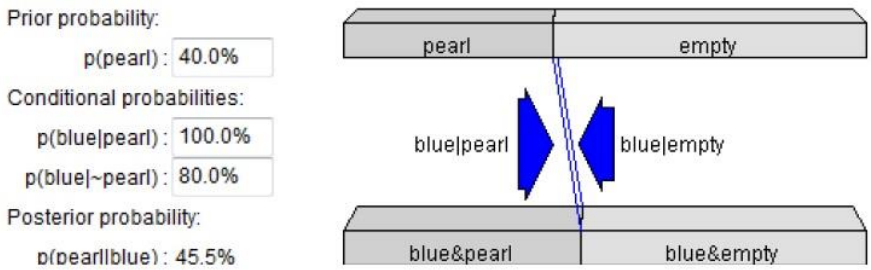

Рис. 6. Обе условные вероятности высоки, но близки по значению

Таким образом сейчас можно сказать, что если яйцо содержит жемчуг, то оно определенно синее. Нет красных яиц с жемчугом. Однако, есть также большая степень "фальшивых положительных результатов" в нашем "тесте", так как очень может быть, что яйцо без жемчуга будет синим. Так как яйцо может быть синим в независимости от наличия в нем жемчуга, факт того, что оно синее не скажет вам о том, содержит ли оно жемчуг, то есть тот факт, что яйцо синее не позволит нам в какой-либо мере пересмотреть нашу вероятность. Это разница между двумя условными вероятностями, которая говорит нам насколько большая величина пересмотра вероятности должна быть реализована.

Что если вторая условная вероятность больше чем первая? Линия наклонена влево и нам надо пересмотреть нашу вероятность в другом направлении (рис. 7). Опять же, это очевидно. Теперь мы видим, что "вероятность что яйцо синее, учитывая, что оно не содержит жемчуг" больше чем "вероятность, что яйцо синее, учитывая, что оно содержит жемчуг", что значит, что в данной задаче больше яиц, которые синие и не содержат жемчуг, чем синих яиц с жемчугом. Так что если мы вытаскиваем синее яйцо, куда вероятнее, что оно пустое, нежели в том случае, когда мы еще не знали, какого оно цвета (априорная вероятность). Так что нам надо пересмотреть нашу вероятность что яйцо содержит жемчуг в сторону уменьшения в этот раз, в независимости от значения априорной вероятности.

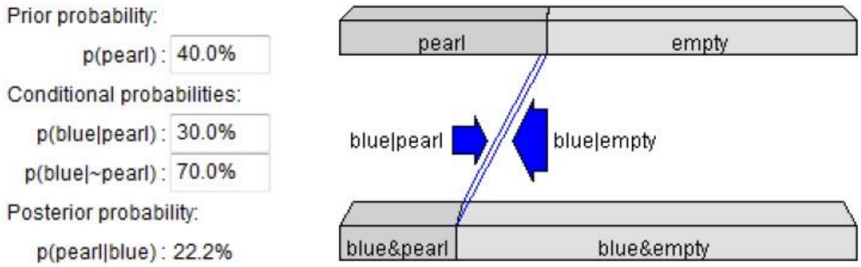

Рис. 7. Условная вероятность синего при условии пустого выше, чем синего, при условии, что оно с жемчугом

И что если две условных вероятности одинаковы? Если наши две условных вероятности одинаковы, тогда они с одинаковой силой влияют на нашу величину изменения вероятности, что значит, что она вообще не меняется. Если яйцо может быть синим, учитывая то что оно содержит жемчуг, а может быть синим, учитывая, что оно не содержит жемчуг, тогда не имеет значения, сколько синих яиц содержат жемчуг, так же как сколько яиц синие и не содержат жемчуг, так как тот факт, что вытащенное яйцо синее не даст нам никакой новой информации о том, есть ли в нем жемчуг. Таким образом мы не получаем новой информации и не можем пересмотреть априорную вероятность. Это тот случай, когда не имеет значения каковы условные вероятности, пока они равны друг другу.

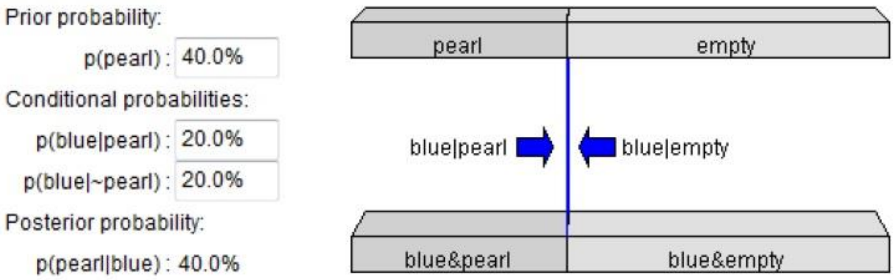

Рис. 8. Условная вероятность синего при условии пустого такая же, как и синего, при условии, что оно с жемчугом; в этом случае априорная вероятность не изменяется

Обычная ошибка при рассуждении о таких проблемах это игнорирование априорной вероятности и концентрация на условных вероятностях. Но, как мы можете видеть, все три куска информации требуются для корректного вычисления апостериорной вероятности.

Юдковски объясняет. Исследования показывают, что множество врачей мысленно заменяют исходную 1%-ную вероятность на 80%-ную вероятность что женщина с раком груди получит положительный результат маммограммы. Как и в случае задачи с яйцами, большинство респондентов, незнакомые с байесовским мышлением скорее всего ответят что синее яйцо содержит жемчуг с вероятностью 30% или возможно 20% (30% шанс настоящего положительного результата минус 10% шанс фальшивого). Даже если эта ментальная операция кажется хорошей идеей, это не имеет смысла с точки зрения вопроса. Это похоже на эксперимент, в котором вы спрашиваете: "Если 18 людей село в автобус, а потом зашло еще семь, то сколько лет водителю автобуса?" Некоторые ответят: "Двадцать пять". Они понимают когда их подталкивают к определенной ментальной процедуре, но не могут связать ее с реальностью. Точно так же, чтобы найти вероятность что женщина, имеющая положительные результат маммограммы, имеет рак груди не имеет смысла заменять исходную вероятность что женщина может иметь рак вероятностью что женщина с раком может получить положительный результата. И вы не можете просто вычесть вероятность фальшивого положительного результат из вероятности истинного положительного результата. Эти операции в значительной степени не связаны, так же как сумма людей в автобусе с возрастом водителя.

Человек, который знает, что у него есть генетическая предрасположенность к алкоголизму, может отреагировать на данный факт избеганием алкоголя более целенаправленно, нежели в любом другом случае. Похожим образом, если мы понимаем почему наши мозги обрабатывают вероятности плохо, мы можем быть способны спланировать наперед и предотвратить ошибки мышления, которые для нас обычны.

Так почему мозг человека, даже мозг дипломированного врача, обычно решает такие задачи неправильно? К счастью, недавние исследования пролили немного света на данную проблему. Согласно данным исследований, мы решаем задачи более или менее верно часто в зависимости от того, как сформулирована проблема. Тип задач, в которых мы чаще всего ошибаемся, формулируется в терминах процентов или вероятностей: "1% женщин..." и так далее. Немного лучше когда проблема сформулирована в терминах частоты:"1 из ста женщин имеет рак груди..." и "80 из 100 женщин с раком груди получают положительный результат теста..." Очевидно, эта формулировка помогает нам представить единичную женщину в пустом пространстве в противовес ста женщинам или 80 женщин в пространстве со 100.

И лучше всего мы решаем задачи, сформулированные в абсолютных числах, которые называются собственные частоты: "400 из 1000 яиц содержат жемчуг..." и "50 из 400 яиц с жемчугом синие..." Это ближе всего к настоящему эксперименту, который мы мысленно проводим и можем представить, как часто мы вытаскиваем синее яйцо и как часто в синем яйце мы находим жемчуг. Следующий очевидный ход – это представлять всю медицинскую статистику в виде собственных частот. К сожалению, хотя собственные частоты являются шагом в верном направлении, этого скорее всего недостаточно. Когда задачи представляются в виде собственных частот, доля людей использующих байесовское мышление, возрастает примерно до 50%.

Существенное улучшение, но все же этого недостаточно если мы говорим о реальных врачах и реальных пациентах.

Представление задачи о яйцах и жемчужинах в виде собственных частот может выглядеть так (рис. 9).

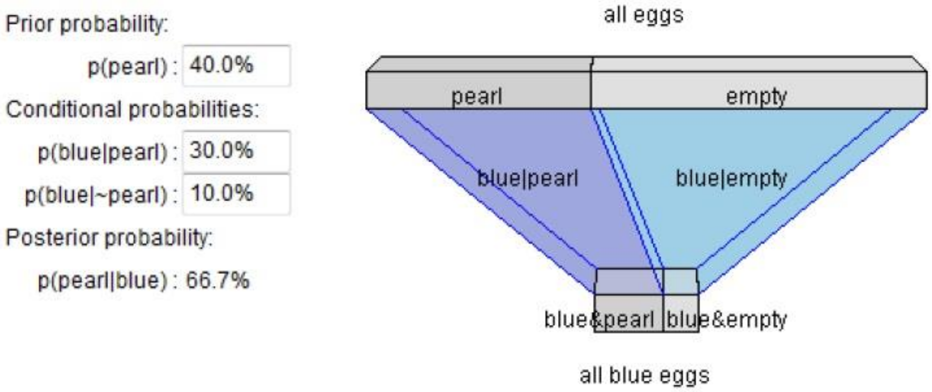

Рис. 9. Представление задачи о яйцах и жемчужинах в виде собственных частот

Так как мы рассматриваем абсолютные значения вместо процентов, полоска вверху больше, чем внизу, потому что сумма всех яиц больше чем сумма только синих. Верхняя полоска выглядит так же, как и ранее, и средняя линия имеет тот же наклон, но теперь нижняя полоска намного меньше, так как на ней отображены только синие яйца. В случае такой визуализации, мы не видим насколько мы пересматриваем вероятность по наклону линии, но можем заметить разницу в пропорциях между верхней и нижней полосках. В примере выше, вы можете видеть, что наличие жемчуга занимает большую часть нижней полоски, нежели на верхней, что значит что мы пересмотрели нашу вероятность в сторону увеличения.

Как такой вид представления выглядит если условные вероятности равны? В этом случае (рис. 10) мы видим, что пропорции на нижней и на верхней полосках одинаковые, хотя линия посередине наклонена.

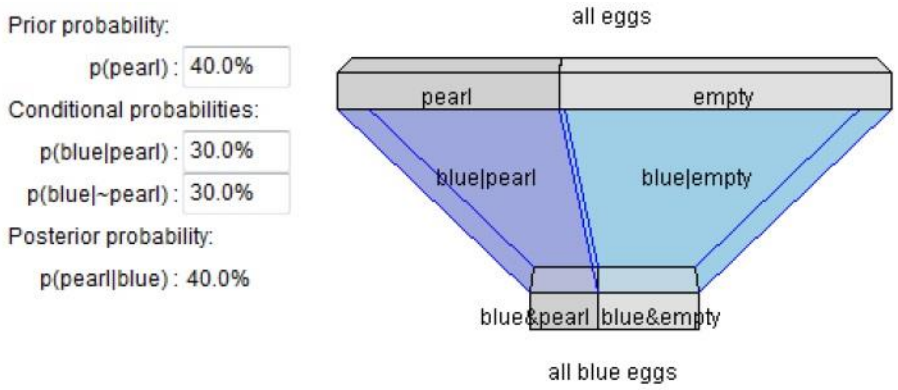

Рис. 10. Представление в виде собственных частот, когда условные вероятности равны

Представление в виде собственных частот показывает кое-что, чего не видно на представлении с помощь вероятностей. Представление в виде собственных частот позволяет нам увидеть, что когда мы уменьшаем две доли на одну величину, результирующие доли остаются тем же самыми.

Теперь давайте взглянем на представление в виде собственных частот для исходной задачи о раке груди. 1% женщин имеют рак груди, 80% из них получат положительный результат теста, и 9,6% женщин без рака груди также получат положительный результат (рис. 11).

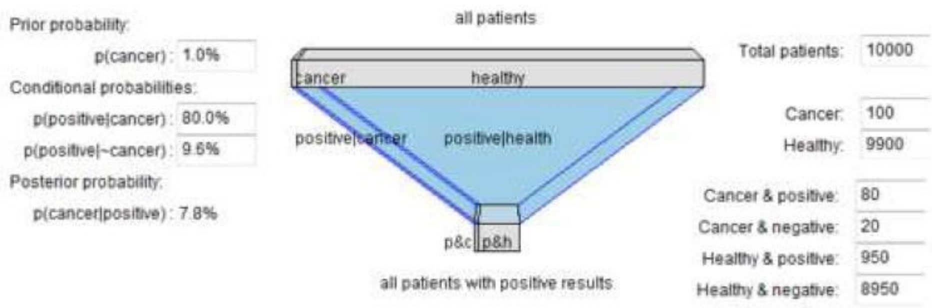

Рис. 11. Представление в виде собственных частот для задачи о раке груди

Несмотря на то, что маммография достаточно точна (только 20% ложных отрицательных и 9,6% ложных положительных), из-за очень маленькой априорной вероятности рака груди (всего 1%), у нас нет веских причин думать, что женщина имеет рак груди после положительного результата теста (апостериорная информация остается довольно низкой – 7,8%).

Дает ли тест полезную информацию? Похоже, что нет. ☹ Для женщин, имеющих положительные результаты теста, он лишь изменяет априорную вероятность в 1%, повышая ее до апостериорных 7,8%. Эта вероятность всё еще остается слишком низкой, чтобы говорить что-то определенное о болезни.

Давайте подробнее рассмотрим все различные сущности в тесте:

p(рак): 0,01. Группа один: 100 женщин с раком груди.

p(~рак): 0,99. Группа два: 9900 женщин без рака груди.

p(положительный результат|рак): 0,8, то есть 80% женщин с раком груди получат положительные результаты теста.

p(~положительный результат|рак): 0,2, то есть 20% женщин с раком груди получат отрицательные результаты.

p(положительный результат|~рак): 0,096. 9,6% здоровых женщин получат положительный результат.

p(~положительный результат|~рак): 0,904. 90,4% здоровых женщин получат отрицательный результат.

p(рак&положительный результат): 0,008 (0,01*0,8). Группа А: 80 женщин с раком груди и положительным результатом.

p(рак&~положительный результат): 0,002 (0,01*0,2). Группа Б: 20 больных женщин с отрицательным результатом теста.

p(~рак&положительный результат): 0,095 (0,99*0,096). Группа В: 950 здоровых женщин, получивших положительный результат.

p(~рак&~положительный результат): 0,895 (0,99*0,904). Группа Г: 8950 здоровых женщин, получивших отрицательный результат.

p(положительный результат): 0,103 (0,008+0,095). 1030 женщин с положительным результатом.

p(~положительный результат): 0,897 (0,002+0,895). 8970 женщин с отрицательным результатом.

p(рак|положительный результат): 0,078 (0,008/0,103). 7,8%-ный шанс, что у вас рак груди, если результат теста положителен.

p(~рак|положительный результат): 0,922 (0,095/0,103) 92,2%-ный шанс, что вы здоровы, если результат теста положительный.

p(рак|~положительный результат): 0,002 (0,002/0,897). 0,2%-ный шанс, что у вас рак, если результат отрицателен.

p(~рак|~положительный результат): 0,998 (0,895/0,897). 99,8%-ный шанс, что вы здоровы, если результат отрицателен.

Как вы можете представить, можно легко перепутать эти множества одно с другим. p(рак&положительный результат) это *то же самое*, что и p(положительный результат&рак), однако p(рак|положительный результат) определенно *не* то же самое, что p(положительный результат|рак). Вероятность, что женщина больна, учитывая, что у нее положительный результат теста, не эквивалентна вероятности что женщина получит положительный результат, если она больна раком. Если вы запутаетесь в этом, то дадите неправильный результат! И, конечно же, p(рак&положительный результат) совершенно отличается от p(рак|положительный результат). Вероятность того, что женщина больна и имеет положительный результат абсолютно не то же самое, что вероятность больной женщины получить положительный результат.

Позже, я покажу теорему Байеса и если вы разберетесь в формуле, вы не перепутаете эти множества. Но и сейчас хорошо бы знать, что каждое из них значит и как они соотносятся друг с другом. Чтобы увидеть, как они соотносятся друг с другом, рассмотрим "степени свободы" между ними. Что значит "степень свободы"? Википедия говорит следующее: в статистике, количество степеней свободы — это количество значений в итоговом вычислении статистики, способных варьироваться.

Что это значит? Рассмотрим это на примере. Между p(рак) и p(~рак) только одна степень свободы, поскольку если вы знаете одну из этих величин, вы знаете и другую. Если одна величина имеет определенное значение, то вторая тоже может иметь только одно значение; это и есть одна степень свободы. Если вероятность рака равна 90%, то вероятность его отсутствия – 10%. Если рака нет с вероятностью 45%, то с 55% вероятностью он наличествует. Тут им просто некуда больше деться, поскольку: p(рак) + p(~рак) = 100%. И, конечно же, то же самое верно для положительного и отрицательного результатов теста: p(положительный результат) + p(~положительный результат) = 100%

Другая пара величин с одной степенью свободы – p(положительный результат|рак) и p(~положительный результат|рак). Если у женщины есть рак (что истинно при любой из двух этих вероятностей), тогда она может получить либо положительный, либо отрицательный результат. Так что если вы знаете одну из этих величин, то вы знаете и другую. Если вероятность получить положительный результат при наличии рака равна 20%, то шанс на получение отрицательного должен быть равен 80%. Почему? Вот поэтому: p(положительный результат|рак) + p(~положительный результат|рак) = 100%

Запомните, это поможет вам в будущем читать подобные обозначения. Приведенное выше обозначение читается как: "Вероятность, что женщина, больная раком, получит положительный результат, плюс вероятность, что она получит отрицательный результат, равны 100%".

И, разумеется, то же самое верно для рассмотрения наличия или отсутствия рака при положительном результате теста: p(рак|положительный результат) + p(~рак|положительный результат) = 100%. Это читается так: "Вероятность, что у женщины рак, учитывая, что у нее положительный результат теста, плюс вероятность, что у нее нет рака, учитывая, что у нее положительный результат теста, равны 100%". Если прочитать это таким образом, это становится очевидно.

И точно так же: p(положительный результат|рак) + p(~положительный результат|рак) = 100%; p(рак|~положительный результат) + p(~рак|~положительный результат) = 100%

Теперь рассмотрим отношение между p(положительный результат|рак) и p(положительный результат|~рак). Если вспомнить то, как поставлено условие в исходной задаче, то первая вероятность равна 80%, а вторая 9,6%. То есть, возможно такое, что вероятность получения положительного результата теста для женщины, при условии, что у нее рак, составляет 80%, а если рака у нее нет – то 9,6%. Эти две величины независимы друг от друга. Шанс ложного отрицательного результата при этом может быть не только 20% вместо 2%, в то время как шанс ложного положительного результата все еще 9,6%. Так что эти две величины – вероятность положительного результата при раке и положительного результата при отсутствии рака – имеют, как принято говорить, две степени свободы. Оба числа могут меняться независимо друг от друга.

Теперь возьмем триплет величин и посмотрим степень свободы между ними. Наши три величины это (1) вероятность рака и положительного результата, (2) вероятность положительного результата при раке и (3) вероятность рака. Как много степеней свободы между ними? Поскольку у нас три величины, то степеней свободы должно быть три. Однако давайте сначала проверим.

В данном случае, мы можем найти одну из этих величин, если у нас есть две другие. В частности: p(положительный результат&рак) = p(положительный результат|рак) × p(рак). Почему именно так? Если мы знаем вероятность наличия рака у женщины и знаем долю больных женщин, у которых будет положительный результат, то вычислить вероятность что больная женщина имеет положительный результат, несложно. Мы умножаем вероятность наличия рака на вероятность того, что больная женщина получит положительный результат, и получаем вероятность p(положительный результат&рак).

Так как мы можем использовать две величины для вычисления третьей, между этими тремя величинами только две степени свободы.

То же самое верно и для такой схемы: p(~положительный результат&рак) = p(~положительный результат|рак) × p(рак). Если мы знаем вероятность что женщина больна, и знаем вероятность что больная женщина получит отрицательный результат, тогда у нас есть и вероятность что женщина больна и не имеет положительного результата.

Давайте рассмотрим другой триплет величин: (1) вероятность получения положительного результата, (2) вероятность рака и положительного результата и (3) вероятность наличия положительного результата и отсутствия рака. Как много степеней свободы между ними? Должно быть достаточно очевидно следующее: p(положительный результат&рак) + p(положительный результат&~рак) = p(положительный результат). Каждая женщина, которая имеет положительный результат, либо больна, либо нет. Эти две возможности в сумме дают 100% женщин с положительным результатом. Так что вы можете использовать их для вычисления общей доли женщин с положительным результатом. То есть, только две степени свободы в данном триплете.

Теперь рассмотрим набор из четырех величин: p(положительный результат&рак), p(положительный результат&~рак), p(~положительный результат&рак), и p(~положительный результат&~рак). На первый взгляд может показаться что здесь только две степени свободы, потому что вы можете вычислить все остальные величины, зная только две из них: p(положительный результат) и p(рак). Например: p(положительный результат&~рак) = p(положительный результат) × p(~рак).

Но это не так! Заметим, что вышепривдененное уравнение верно только если p(положительный результат) и p(~рак) статистически независимы. Уравнение выше истинно только если больная женщина не влияет на шансы получить положительный результат. Но это не наш случай! В соответствии с нашей историей, она скорее получит положительный результат, если она больна, нежели тогда, когда здорова.

Но есть более простой путь увидеть почему это неправильно – заметить, что все эти четыре величины относятся к разным группам женщин и конечно же в каждой группе будет разное количество женщин. У нас может быть 500 женщин больных раком и с положительным результатом (пусть это будет группа А), 150 женщин с раком груди и отрицательным результатом теста (группа Б), 50 женщин без рака и положительным результатом (группа В), и 900 женщин без рака и отрицательным результата (группа Г). И каждая из этих величин может иметь свое значение, независимое от остальных.

Теперь вы думаете, что этот набор из четырех величин имеет четыре степени свободы и вы были бы правы, если бы не тот факт, что все четыре в сумме дают 100% женщин. То есть, все эти четыре вероятности в сумме должны дать 100%. Для примера, в предыдущем параграфе я поместил 500 женщин в группу А, 150 в группу Б, 50 в группу В и 900 в группу Г. Всего получается 1600 женщин. Таким образом, вероятность, что женщина принадлежит к группе А равна 31,25%. То есть, p(A)=31,25%. Соответственно, p(Б)=9,375%, p(В)=3,125% и p(Г)=56,25%. И, что неудивительно, эти проценты в сумме дают 100%.

Исходя из этого, мы можем использовать три величины, чтобы найти четвертую, поскольку мы знаем, что в сумме они должны давать 100%, что стоит им одной степени свободы: p(положительный результат&рак) + p(положительный результат&~рак) + p(~положительный результат&рак) + p(~положительный результат&~рак) = 100%

На деле, раз у вас есть все четыре группы (А = рак и положительный результат, Б = рак и отрицательный результат, В = нет рака и положительный результат, Г = нет рака и отрицательный результат), вы можете легко использовать их для вычисления всех остальных величин. Например: p(рак|положительный результат) = А / (А + В). Вероятность что у женщины рак, при условии, что она получила положительный результат, равна вероятности что у нее рак и положительный результат (группа А), деленная на вероятность, что у нее положительный результат (А+В): p(положительный результат) = А+В.

Второй пример: вероятность, что женщина имеет рак при том, что ее результат теста отрицательный это вероятность, что у нее рак и что тест отрицательный, деленная на вероятность, что у нее отрицательный результат (Б+Г): p(рак|~положительный результат) = Б / (Б + Г). Третий пример. Вероятность что женщина больна равна вероятности что у нее рак и положительный результат (группа А) плюс вероятность что у нее рак и отрицательный результата (группа Б): p(рак) = А + Б.

Таким образом мы вычисляем все величины, которые хотим, если у нас есть А, Б, В, и Г, и таким образом четыре этих величины имеют три степени свободы. Но это не должно удивлять вас, ведь вы уже знаете, что вы можете решать подобные задачи, имея только три куска информации: исходную вероятность и две условных.

Вернемся к задаче с яйцами и жемчугом. Напомню, у вас есть большая бочка, наполненная пластиковыми яйцами. В некоторых из них находится жемчуг, в других ничего нет. Некоторые синего цвета, остальные красные. Пусть 40% яиц – синие, 5\13 яиц, содержащих жемчуг – синие, и 20% яиц пустые и окрашены в красный цвет. Какова вероятность, что синее яйцо содержит жемчуг? Не читайте дальше. Попробуйте решить, используя отношения, которые мы открыли ранее.

Какие куски информации у нас есть? Мы знаем, что 40% яиц – синие: p(синее) = 40%. Мы также знаем, что 5/13 яиц, которые содержат жемчуг, синие: p(синее|жемчуг) = 5/13. И мы знаем, что 20% яиц одновременно красные и пустые: p(~синее&~жемчуг) = 20%. Что хотим узнать: p(жемчуг|синее) = ?

Ладно, как нам получить эту апостериорную вероятность? Поскольку вы новички в теореме Байеса, вы возможно не уверены, каков наибыстрейший путь к ответу, так что давайте начнем с нахождения наиболее очевидных величин из тех, что мы можем узнать, среди тех 16, что присутствуют в задаче:

p(синее) = 40% …дано в условии
p(~синее) =
p(жемчуг) =
p(~жемчуг) =
p(жемчуг&синее) =
p(жемчуг&~синее) =
p(~жемчуг&синее) =
p(~жемчуг&~синее) = 20%... дано в условии
p(синее|жемчуг) = 5/13 …дано в условии
p(~синее|жемчуг) =
p(синее|~жемчуг) =

p(~синее|~жемчуг) =
p(жемчуг|синее) = ???
p(~жемчуг|синее) =
p(жемчуг|~синее) =
p(~жемчуг|~синее) =

Как нам узнать значения этих величин? Ну, давайте проверим отношения между величинами, что мы открыли и посмотрим, какие можем найти. Ниже приведены правила, которые мы нашли, когда обсуждали задачу про больных раком:

p(рак) + p(~рак) = 100%
p(положительный результат) + p(~положительный результат) = 100%
p(положительный результат|рак) + p(~положительный результат|рак) = 100%
p(рак|положительный результат) + p(~рак|положительный результат) = 100%
p(положительный результат|рак) + p(~положительный результат|рак) = 100%
p(рак|~положительный результат) + p(~рак|~положительный результат) = 100%
p(положительный результат&рак) = p(положительный результат|рак) × p(рак)
p(~положительный результат&рак) = p(~положительный результат|рак) × p(рак)
p(положительный результат&рак) + p(положительный результат&~рак) = p(положительный результат)
p(положительный результат&рак) + p(положительный результат&~рак) + p(~положительный результат&рак) + p(~положительный результат&~рак) = 100%
p(рак|положительный результат) = p(рак&положительный результат) / [p(рак&положительный результат + p(~рак&положительный результат)]
p(положительный результат) = p(рак&положительный результат) + p(~рак&положительный результат)
p(рак|~положительный результат) = p(рак&~положительный результат) / [p(рак&~положительный результат) + p(~рак&~положительный результат)]
p(~положительный результат) = p(рак&~положительный результат) + p(~рак&~положительный результат)
p(рак) = p(рак&положительный результат) + p(рак&~положительный результат)
p(~рак) = p(~рак&положительный результат) + p(~рак&~положительный результат)

А теперь давайте переведем эти правила для задачи с синим и жемчугом, меняя каждое упоминание о раке (который мы пытались обнаружить) на жемчуг (который мы пытаемся найти), и все вхождения о положительном результате (нашем предыдущем тесте) на синее (нашем текущем тесте):

p(жемчуг) + p(~жемчуг) = 100%
p(синее) + p(~синее) = 100%
p(синее|жемчуг) + p(~синее|жемчуг) = 100%
p(жемчуг|синее) + p(~жемчуг|синее) = 100%
p(синее|жемчуг) + p(~синее|жемчуг) = 100%
p(жемчуг|~синее) + p(~жемчуг|~синее) = 100%
p(синее&жемчуг) = p(синее|жемчуг) × p(жемчуг)
p(~синее&жемчуг) = p(~синее|жемчуг) × p(жемчуг)
p(синее&жемчуг) + p(синее&~жемчуг) = p(синее)
p(синее&жемчуг) + p(синее&~жемчуг) + p(~синее&жемчуг) + p(~синее&~жемчуг) = 100%
p(жемчуг|синее) = p(жемчуг&синее) / [p(жемчуг&синее + p(~жемчуг&синее)]
p(синее) = p(жемчуг&синее) + p(~жемчуг&синее)
p(жемчуг|~синее) = p(жемчуг&~синее) / [p(жемчуг&~синее) + p(~жемчуг&~синее)]
p(~синее) = p(жемчуг&~синее) + p(~жемчуг&~синее)
p(жемчуг) = p(жемчуг&синее) + p(жемчуг&~синее)
p(~жемчуг) = p(~жемчуг&синее) + p(~жемчуг&~синее)

Ладно, какие из этих уравнений мы можем использовать, учитывая то, что мы знаем? Ну, наиболее очевидными выглядят следующие: p(синее) + p(~синее) = 100% и p(синее|жемчуг) +

p(~синее|жемчуг) = 100%. Это дает нам 60% для p(~синее) и 8/13 для p(~синее|жемчуг). И еще вспомним следующее: p(~синее) = p(жемчуг&~синее) + p(~жемчуг&~синее). Так, мы знаем две величины из этого уравнения, так что: 60% = 20% + p(жемчуг&~синее). Что значит: p(жемчуг&~синее) = 60% – 20% = 40%. Так что мы можем добавить и его в нашу таблицу. Теперь она выглядит так:

p(синее) = 40%
p(~синее) = 60% …поскольку p(синее) + p(~синее) = 100%
p(жемчуг) =
p(~жемчуг) =
p(жемчуг&синее) =
p(жемчуг&~синее) = 40% …поскольку p(~синее) = p(жемчуг&~синее) + p(~жемчуг&~синее)
p(~жемчуг&синее) =
p(~жемчуг&~синее) = 20%
p(синее|жемчуг) = 5/13
p(~синее|жемчуг) = 8/13 …поскольку p(синее|жемчуг) + p(~синее|жемчуг) = 100%
p(синее|~жемчуг) =
p(~синее|~жемчуг) =
p(жемчуг|синее) = ???
p(~жемчуг|синее) =
p(жемчуг|~синее) =
p(~жемчуг|~синее) =

Теперь, что еще мы можем сделать? Просмотрим все варианты отношений между величинами выше и увидим, что есть одно соотношение, которое позволяет продвинуться дальше. Вот оно: p(жемчуг|~синее) = p(жемчуг&~синее) / [p(жемчуг&~синее) + p(~жемчуг&~синее)]. Подставляя уже найденные нами величины получаем: p(жемчуг|~синее) = 40% / (40% + 20%). Что значит: p(жемчуг|~синее) = 2/3. Это дает нам возможность определить величину p(~жемчуг|~синее) вследствие: p(жемчуг|~синее) + p(~жемчуг|~синее) = 100%. Таким образом: p(~жемчуг|~синее) = 1/3.

И теперь наша таблица величин выглядит следующим образом:

p(синее) = 40%
p(~синее) = 60%
p(жемчуг) =
p(~жемчуг) =
p(жемчуг&синее) =
p(жемчуг&~синее) = 40%
p(~жемчуг&синее) =
p(~жемчуг&~синее) = 20%
p(синее|жемчуг) = 5/13
p(~синее|жемчуг) = 8/13
p(синее|~жемчуг) =
p(~синее|~жемчуг) =
p(жемчуг|синее) = ???
p(~жемчуг|синее) =
p(жемчуг|~синее) = 2/3 …поскольку p(жемчуг|~синее) = p(жемчуг&~синее) / [p(жемчуг&~синее) + p(~жемчуг&~синее)]
p(~жемчуг|~синее) = 1/3 …поскольку p(жемчуг|~синее) + p(~жемчуг|~синее) = 100%

Мы значительно продвинулись! Еще одно уравнение может быть решено: p(~синее&жемчуг) = p(~синее|жемчуг) × p(жемчуг). Итак: 40% = (8/13) × p(жемчуг). И таким образом: p(жемчуг) = 40% / (8/13) = (2/5) / (8/13) = 13/20 = 65%. Теперь, когда мы знаем p(жемчуг), мы также можем решить следующее из наших уравнений: p(синее&жемчуг) = p(синее|жемчуг) × p(жемчуг). Так: p(синее&жемчуг) = (5/13) × (13/20) = 1/4 = 25%. Дополняя нашу таблицу величин, мы теперь имеем:

p(синее) = 40%
p(~синее) = 60%
p(жемчуг) = 65% …поскольку p(~синее&жемчуг) = p(~синее|жемчуг) × p(жемчуг)
p(~жемчуг) =
p(жемчуг&синее) = 25% …поскольку p(синее&жемчуг) = p(синее|жемчуг) × p(жемчуг)
p(жемчуг&~синее) = 40%
p(~жемчуг&синее) =
p(~жемчуг&~синее) = 20%
p(синее|жемчуг) = 5/13
p(~синее|жемчуг) = 8/13
p(синее|~жемчуг) =
p(~синее|~жемчуг) =
p(жемчуг|синее) = ???
p(~жемчуг|синее) =
p(жемчуг|~синее) = 2/3
p(~жемчуг|~синее) = 1/3

Последний кусок найти легко: p(синее) = p(жемчуг&синее) + p(~жемчуг&синее). Что дает нам: 40% = 25% + p(~жемчуг&синее). То есть: p(~жемчуг&синее) = 40% – 25% = 15%. И теперь мы можем наконец найти p(жемчуг|синее), поскольку: p(жемчуг|синее) = p(жемчуг&синее) / [p(жемчуг&синее + p(~жемчуг&синее)]. Из чего получается: p(жемчуг|синее) = 25% / (25% + 15%). То есть: p(жемчуг|синее) = 25% / 40%. И в результате: p(жемчуг|синее) = 62,5%.

На деле, нам не нужно вычислять все эти величины. Но было полезно попрактиковаться. :)

Теперь проверим наши вычисления. Имеют ли они смысл? Вернемся к исходному условию: Предположим, что у вас есть большая бочка, наполненная пластиковыми яйцами. В некоторых из них находится жемчуг, в других ничего нет. Некоторые синего цвета, остальные красные. Пусть 40% яиц - синие, 5\13 яиц, содержащих жемчуг – синие, и 20% яиц пустые и окрашены в красный цвет. Какова вероятность, что синее яйцо содержит жемчуг?

Вспомним, что мы нашли, что p(жемчуг) равен 13/20. Это наша априорная вероятность: 65%. То есть существует шанс в 65% что в яйце содержится жемчуг до того, как мы посмотрели, какого оно цвета. Каковы наши условные вероятности? Одна из них дается в условии. Вероятность, что яйцо, которое содержит жемчуг, синее, p(синее|жемчуг), равна 5/13 – что нельзя нормально округлить. Другая условная вероятность, вероятность что яйцо с жемчугом будет окрашено в красный цвет – p(~синее|жемчуг) = 8/13. Так что если в яйце есть жемчуг, то вероятней, что оно красное, нежели, что оно синее (хотя совсем чуть-чуть). Так что когда мы обнаруживаем, что яйцо, которое было извлечено из бочки, синее, это снижает вероятность того, что в нем есть жемчуг, по сравнению с вероятностью, которую мы знали до того, как извлекли яйцо. Так что после теста, в котором яйцо оказалось синим, вероятность наличия в нем жемчуга немного снижается.

И вуаля! Вот что мы видим – наша априорная вероятность что наше выбранное яйцо содержит жемчуг составляет 65% и в соответствии с нашими расчетами апостериорная вероятность что наше выбранное яйцо содержит жемчуг немного ниже – 62,5%. Так что да, похоже, что наши вычисления, которые мы провели с данными задачи, верны и работают корректно.

***

Примечание Багузина. *Я решил изобразить решение задачи графически (рис. 15). Напомню условия: (1) пусть 40% яиц – синие, (2) 5\13 яиц, содержащих жемчуг – синие, и (3) 20% яиц пустые и окрашены в красный цвет. Вертикальная линия на отметке 0,4 по шкале Цвет отражает условие (1). Любая горизонтальная линия, отсекающая яйца с жемчугом от пустых даст то же соотношение 4/6, следовательно, линия должна идти под углом с небольшим ростом слева направо, чтобы из всех яиц с жемчугом на долю синих пришлось менее 4/6, а именно 5/13. Эта наклонная линия должна проходить на такой высоте, чтобы площадь Г составляла 0,2 от площади квадрата 1*1.*

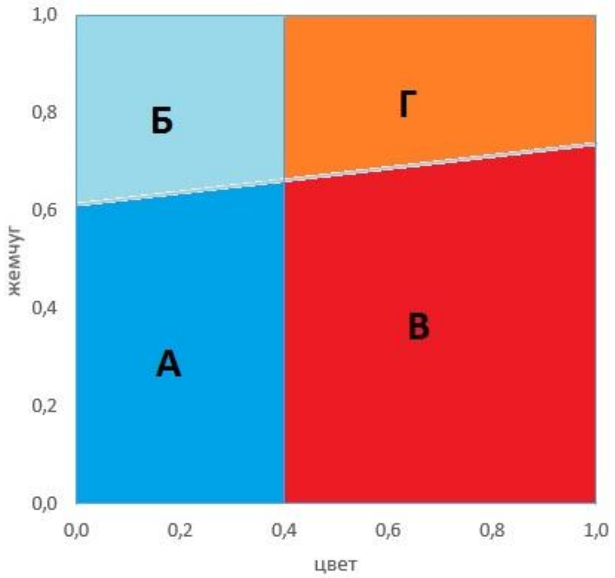

Рис. 15. Задача про яйца и жемчуг: А – синее с жемчугом, Б – синее пустое, В – красное с жемчугом, Г – красное пустое

Наши 16 условий:
p(синее) = (А + Б)/(А + Б+ В + Г); доля синего во всем
p(~синее) = (В + Г)/(А + Б+ В + Г); доля красного во всем
p(жемчуг) = (А + В)/(А + Б+ В + Г); доля жемчуга во всем
p(~жемчуг) = (Б + Г)/(А + Б+ В + Г); доля пустых во всем
p(жемчуг&синее) = А/(А + Б+ В + Г); доля синего жемчуга во всем
p(жемчуг&~синее) = В/(А + Б+ В + Г); доля красного жемчуга во всем
p(~жемчуг&синее) = Б/(А + Б+ В + Г); доля синего пустого во всем
p(~жемчуг&~синее) = Г/(А + Б+ В + Г); доля красного пустого во всем
p(синее|жемчуг) = А/(А + В); доля синего в жемчуге
p(~синее|жемчуг) = В/(А + В); доля красного в жемчуге
p(синее|~жемчуг) = Б/(Б + Г); доля синего в пустом
p(~синее|~жемчуг) = Г/(Б + Г); доля красного в пустом
p(жемчуг|синее) = А/(А+Б); доля жемчуга в синем
p(~жемчуг|синее) = Б/(А+Б); доля пустого в синем
p(жемчуг|~синее) = В/(В+Г); доля жемчуга в красном
p(~жемчуг|~синее) = Г/(В+Г); доля пустого в красном

***

Работая над этими задачами вы могли заметить, что сильное, но редкое свидетельство, которое направлено в одну сторону, должно быть уравновешено слабым, но частым свидетельством, направленным в другую сторону. Поясним это на примере задачи про больных раком: p(рак) = p(рак|положительный результат) * p(положительный результат) + p(рак|~положительный результат) * p(~положительный результат). Это читается так: "Вероятность что женщина больна раком равна [вероятности что у нее рак, при условии что ее тест положителен умноженной на вероятность что у нее положительный результат теста] плюс [вероятность что у женщины рак, при условии что у нее отрицательный результат теста умноженной на вероятность что у нее отрицательный результат теста]."

Таким образом, редкое, но сильное свидетельство от одной из условных вероятностей должно быть уравновешено частым но слабым свидетельством от другой условной вероятности, поскольку две из них должны в сумме давать, в этом примере, p(рак). Юдковски называет это принципом сохранения вероятности.

Теперь еще один термин: *отношение правдоподобия* – это правдоподобие получения настоящего положительного результата против правдоподобия получения ложного положительного результата. В частности, отношение правдоподобия представляет собой вероятность, что тест даст настоящий положительный результат, деленную на вероятность получения ложного положительного результата. Однако отношение правдоподобия не особо много говорит нам, что делать, если мы получили отрицательный результат.

Например, p(жемчуг|синее) не зависит от p(жемчуг|~синее). Даже если мы знаем отношение правдоподобия, и таким образом знаем, что делать если получаем положительный результат в нашем тесте, это не говорит нам что делать, если тест дает отрицательный результат. Чтобы в дальнейшем это проиллюстрировать, рассмотрим следующую проблему, снова предложенную Элизером.

Предположим, что есть две бочки, каждая из которых содержит пластиковые яйца. В обоих бочках только по 40% яиц содержат жемчуг. В обоих бочках некоторые яйца окрашены в синий цвет, а остальные в красный. В первой бочке 30% яиц с жемчугом – синие, и 10% пустых тоже синие. Во второй бочке 90% яиц с жемчугом синие, как и 30% пустых. [Предположим, что вам нужен жемчуг, а не наоборот] вы бы предпочли синее яйцо из первой или из второй бочки? А может вы бы взяли красное яйцо из второй или первой бочки?

В этом случае нам нужно вычислить вероятность что синее яйцо из первой бочки содержит жемчуг и сравнить с вероятностью что синее яйцо из второй бочки содержит жемчуг. Для второго вопроса мы вычисляем вероятность что красное яйца из первой бочки содержит жемчуг и сравниваем с вероятностью, что красное яйцо из второй бочки содержит жемчуг.

В обоих бочках 40% яиц содержат жемчуг. Так что наша исходная вероятность, p(жемчуг) равна 40% в обеих бочках. И если вы задержитесь здесь, вместо того чтобы продолжать, тут вам интуитивно может стать ясно, что не имеет значения, из какой бочки доставать синее яйцо, поскольку, в первой бочке, p(синее|жемчуг) / p(синее|~жемчуг) = 30/10; во второй бочке, p(синее|жемчуг) / p(синее|~жемчуг) = 90/30… что представляет собой одно и тоже отношение. Они оба равны трем. И так как исходная вероятность – в этом случае, p(жемчуг) – одинакова для обоих бочек, и отношение между условными вероятностями одинаково для обоих бочек, это дает нам, что p(жемчуг|синее) будет одинаково для обоих бочек. Предпочесть ли нам яйцо из первой или второй бочки? Нет никакой разницы, p(жемчуг|синее) одинакова в любом случае.

А как с красными яйцами? Нам стоит предпочесть красное яйцо из первой бочки или из второй? То есть: для какой бочки p(жемчуг|~синее) выше? В первой бочке 70% яиц с жемчугом красные и 90% пустых яиц тоже. А во второй бочке в красный цвет окрашены 10% яиц с жемчугом и 70% яиц без него. Здесь отношение между условными вероятностями для каждой бочки свое, а именно: в первой бочке p(~синее|жемчуг) / p(~синее|~жемчуг) = 70/90; во второй бочке p(~синее|жемчуг) / p(~синее|~жемчуг) = 10/70.

Поскольку отношение условных вероятностей для первой бочки отличается от отношения условных вероятностей для второй бочки, мы можем сказать, что для одной из бочек условия лучше, нежели для другой. И даже без вычислений можно сказать что вероятность p(жемчуг|~синее) будет выше для первой бочки, так что лучше будет взять красное яйцо из нее, нежели из второй. Отношение p(~синее|жемчуг) к p(~синее|~жемчуг) выше для первой бочки, чем для второй.

Снова, вам нужно прочесть это вслух: "Отношение вероятности, что яйцо красное учитывая, что в нем есть жемчуг к вероятности что яйцо красное при условии, что оно пустое, выше для первой бочки, чем для второй, так что мы скорее предпочтем достать красное яйцо из первой бочки, а не из второй, если мы, конечно, хотим получить жемчужину." Эта проблема иллюстрирует тот факт, что p(жемчуг|синее) и p(жемчуг|~синее) имеют две степени свободы, даже если p(жемчуг) фиксирована. Для любой из бочек p(жемчуг) одна и та же, однако это не значит, что отношение p(жемчуг|синее) к p(жемчуг|~синее) одно и то же для обеих бочек, поскольку p(синее) для каждой из бочек свое. Как говорит Юдковски: Во второй бочке пропорция синих яиц с жемчугом та же, что и в первой бочке, однако синих яиц там намного больше! Это изменяет набор красных яиц таким образом, что пропорции [между условными вероятностями] меняются.

Возвращаясь к задаче про рак. Отношение правдоподобия медицинского теста – число настоящих положительных результатов, деленное на число ложных положительных результатов – говорит нам все, что нужно знать о значении положительного результата. Но не говорит нам о значении отрицательного результата и не говорит нам как часто тест полезен. Для примера, маммография с числом попаданий 80% для пациентов с раком и числом ложных положительных результатов 9,6% для здоровых пациентов имеет то же отношение правдоподобия, как тест с 8% верных результатов и уровнем ложных положительных результатов в 0,96%. Хотя эти два теста имеют одно и то же отношение правдоподобия, первый тест более полезен в любом случае – он определяет болезнь значительно чаще и у него отрицательный результат теста является более сильным свидетельством здоровья.

Отношение правдоподобия для положительного результата суммирует разное давление двух условных вероятностей для положительного результата, и таким образом суммирует насколько положительный результат сдвигает априорную вероятность... Конечно, отношение правдоподобия не может сказать все; отношение правдоподобия и априорная вероятность вместе – это два числа, в то время как задача имеет три степени свободы.

Покойный мастер Байесианства Э. Т. Джейнс полагал, что свидетельство должно измеряться в децибелах. Почему децибелы? В децибелах измеряют экспоненциальную разницу в энергии звука, точно так же как по шкале Рихтера измеряют экспоненциальную разницу в сейсмической энергии, выражающейся в землетрясении. По шкале Рихтера, землетрясение с магнитудой в 7 баллов сильнее землетрясения в 6 баллов не на один балл, а в десять раз. А с магнитудой в 8 баллов оно в сто раз сильнее, нежели землетрясение в 6 баллов. Подобным же образом, полная тишина представляет собой 0 децибел, шепот равен 20 децибелам, обычный разговор – 60 децибелам. При этом обычный разговор высвобождает энергии не в три раза больше, чем шепот, а в 10 000 раз, при этом вырастая на 40 децибелов.

Чтобы получить звук в децибелах используется следующее уравнение: децибелы = 10 × log10(интенсивность). Понимание логарифмов поможет нам понять, что значит думать о свидетельстве в виде децибелов (в виде степеней).

Вернемся к задаче про рак. Предположим, что мы начали с 1%-ной априорной вероятности, что у женщины рак груди. Тогда мы назначаем три разных теста на рак, у каждого из которых свое отношение правдоподобия. Пусть они будут равны соответственно 25:3, 18:1 и 7:2.

Если мы последуем совету Джейнса буквально и измерим нашу априорную вероятность в децибелах, мы получим: 10 × log10(1/99) = –20 децибелов свидетельства, что у женщины рак груди. Теперь, предположим, что мы назначили первый тест, с отношением правдоподобия 25/3, и женщина получила положительный результат. Это дает нам 9 положительных децибелов свидетельства, что у нее рак груди, поскольку: 10 × log10(25/3) = +9 децибелов свидетельства, что у женщины рак груди. Далее мы назначаем второй тест и она снова получает положительный результат: 10 × log10(18/1) = +13 децибелов свидетельства, что у женщины рак груди. И третий тест для нее дает положительный результат: 10 × log10(7/2) = +5 децибелов свидетельства, что у женщины рак груди.

Бедняжка начала с очень низкой вероятности, что у нее рак груди, однако сейчас три теста, обладающие неплохой эффективностью, дали для нее положительный результат одновременно. Это не хорошо! Она начала с –20 децибелов свидетельства, что у нее рак груди, однако эти три теста дали в сумме 27 децибелов свидетельства (9+13+5) в пользу того, что она больна, так что теперь у нас +7 децибелов свидетельства, что у нее рак груди. На линейной шкале +7 смотрятся как нечто маленькое, однако на экспоненциальной +7 децибелов свидетельства означают, что вероятность наличия рака 83%!

Заметим, что +7 децибелов свидетельства не так много, равно как и –20 децибелов свидетельства не мало. Исходные –20 децибелов свидетельства значат, что с вероятностью 99% женщина здорова, однако +7 децибелов свидетельства значат 83%, что она больна. Тогда, конечно же, +20 децибелов свидетельства значили бы, что она больна с вероятностью в 99%. Теперь, когда вы понимаете экспоненциальную силу, которую свидетельство имеет в вероятностном мышлении, попробуйте примерно, не делая вычислений, дать ответ на следующую задачу, которую я взял у Юдковски и слегка перефразировал.

Перед вами лежит сумка, в которой находится 1000 фишек для покера. Я начал с двух таких сумок, в одной из которых лежит 700 красных фишек и 300 синих, в другой 300 красных и 700 синих. Я подбрасываю монетку, чтобы определить какую из сумок вам показать, так что ваша априорная вероятность, что в сумке перед вами больше красных фишек, равна 50%. Теперь вы вслепую достаете из сумки случайную фишку. Вы смотрите на ее цвет, записываете его и кладете фишку обратно в сумку, после чего фишки в ней перемешиваются. Вы делаете так 12 раз и у вас получается, что из этих 12 "примеров" 8 красных и 4 синих. Какова вероятность, что это сумка, в которой больше красных фишек, чем синих?

Остановитесь здесь и покрутите задачу в голове и попробуйте грубо прикинуть ответ. Согласно Варду Эдвардсу и Лоренсу Филипсу, большинство людей, которым давалась эта задача, дают ответ между 70% и 80%. Ваша оценка получилась выше? Тогда поздравляю! Правильный ответ составляет порядка 97%.

Даже без вычислений вы можете здесь прийти к верному ответу. Как указано в задаче, отношение правдоподобия для результата теста вытаскивания красной фишки равно 7/3, а для результата вытаскивания синей – 3/7. Таким образом, положительный результат для любого теста имеет одинаковую степень воздействия на нашу финальную вероятность в одном направлении или в другом, однако, разумеется, красная фишка сдвигает p(сумка, где красных фишек больше) в противоположном направлении сдвига синей фишкой.

Если вы вытаскиваете красную фишку, кладете обратно и вытаскиваете синюю, эти два куска свидетельства аннулируют друг друга, но только потому, что отношение правдоподобия для другого теста абсолютно противоположны. Если вы вытащили красную фишку, а после нее синюю и больше ничего не делали, тогда ваша вероятность, что сумка перед вами содержит больше красных фишек, чем синих, по-прежнему равна 50%.

Вы вытащили 12 фишек и получили на четыре красных больше, чем синих. Это несколько "децибел" свидетельства в пользу того, что в сумке больше красных фишек, что значительно увеличивает вероятность. Когда вы уберете те фишки, что аннулируют друг друга, у вас останется несколько, которые сдвигают вашу вероятность, что это сумка с красными фишками, с силой отношения правдоподобия в 7/3! Так что даже без выполнения вычислений вы можете понять, что вероятность того, что в сумке больше красных фишек, похоже, весьма высока.

Если отношение правдоподобия вашего положительного теста равно 7/3 и вы получили на четыре положительных тестов больше, нежели отрицательных, то ваши шансы будут вычисляться следующим образом: $7^4:3^4 = 2401:81 \approx 30:1$ или 97%.

Ладно. Я думаю, что вы начали понимать, как работают эти типы вероятностей. Давайте посмотрим на последнюю задачу. Вы механик. Когда механизм перестает работать, есть 30% шанса, что это связано с засорившимся насосом. Если насос засорился, то с вероятностью 45% он будет искрить. Если же он в порядке, то искрить он может с вероятностью в 5%. Клиент приносит вам механизм, который не работает. Вы запускаете его и видите искры. Какова вероятность, что насос в механизме засорен?

Итак, мы хотим найти p(засорен|искрит), при этом мы уже знаем: p(засорен) = 30%, p(~засорен) = 70%, p(искрит|засорен) = 45%, p(искрит|~засорен) = 5%. Вспомним, что: p(искрит|засорен) × p(засорен) = p(искрит&засорен). Так что: p(искры&засорен) = 45% × 30% = 13,5%. Также вспомним, что: p(искрит|~засорен) × p(~засорен) = p(искрит&~засорен), так что: p(искрит&~засорен) = 5% × 70% = 3,5%. И наконец вспомним, что: p(засорен|искрит) = p(искрит&засорен) / [p(искрит&засорен + p(искрит&~засорен)], таким образом получая: p(засорен|искрит) = 13.5% / (13.5% + 3,5%), и ответ: p(засорен|искрит) = 79,4%. Теперь, если выразить все наши вычисления для этой задачи в одной формуле, мы получим:

$$p(\text{засорен}|\text{искрит}) = \frac{p(\text{искрит}|\text{засорен}) * p(\text{засорен})}{p(\text{искрит}|\text{засорен}) * p(\text{засорен}) + p(\text{искрит}|\sim\text{засорен}) * p(\sim\text{засорен})}$$

Общая форма для этого:

$$p(H|E) = \frac{p(E|H) \times p(H)}{p(E|H) \times p(H) + p(E|{\sim}H) \times p({\sim}H)}$$

Это теорема Байеса.

И поскольку: $p(E|H) \times p(H) + p(E|{\sim}H) \times p({\sim}H) = p(E)$

Мы можем свести теорему Байеса к следующему:

$$p(H|E) = \frac{p(E|H) \times p(H)}{p(E)}$$

Эта формулировка проще, так что вы можете встретить ее чаще, хотя она не дает такой ясной картины, как предыдущая. Хотя правильны обе.

Если дана гипотеза H, которую мы хотим исследовать и наблюдение E, которое является свидетельством для H, теорема Байеса говорит нам, как мы должны обновить нашу вероятность, что H верно, учитывая свидетельство E.

В задаче про рак, H – это "женщина больна раком", E – положительный результат теста маммографии. Теорема Байеса сообщает нам какова апостериорная вероятность, что женщина имеет рак груди (H), учитывая положительный результат теста маммографии (E).

Юдковски делает вывод. С данной точки зрения теорема Байеса может показаться вопиюще очевидной или даже тавтологичной, нежели возбуждающей и новой. Если так, то данное объяснение в полной мере достигло своей цели. Вот оно. Теперь вы поняли знаменитую теорему Реверенда Томаса Байеса. Он горд за вас.

Так зачем это вам? В чем сущность теоремы? Юдковски приводит пример, как кто-то думает, что человечество избегнет ядерной войны в следующие сто лет. Когда его спрашивают почему, он говорит, "все игроки, вовлеченные в решения касательно ядерной войны, в ней не заинтересованы". А почему срок определен в 100 лет? "Потому что я оптимист" – отвечает он. Что делает такой вид мышления иррациональным? Может быть то, что фраза "Потому что я оптимист" не дает нам уверенности, что утверждение верно? (Возможно, что оно верно, однако мы не поверим этому только потому, что кто-то говорит, что он оптимист.)

Юдковски объясняет. Другие интуитивные аргументы включают идею, что "в независимости от того, оптимист ты или нет, это никак не меняет возможности уничтожения человечества в ядерной войне", или "чистая надежда не является свидетельством для ядерной войны, поскольку это не наблюдение для нее." Также есть математический ответ, ясный, точный и содержащий все интуитивные знания как отдельные случаи. Он известен как теорема Байеса. Для примера, ответ "в независимости от того, оптимист ты или нет, это никак не меняет возможности уничтожения человечества в ядерной войне" может быть преобразован как: p(вы оптимист|человечество избежит ядерной войны в следующие сто лет) = p(вы оптимист | человечество не избежит ядерной войны в следующие сто лет).

Юдковски продолжает (он использует "A" для гипотезы и "X" для свидетельства, вместо H и E, как это выше делал я). Поскольку две вероятности для $p(X|A)$ и $p(X|{\sim}A)$ равны, теорема Байеса говорит, что $p(A|X) = p(A)$; как мы видела ранее, когда две условных вероятности равны, пересмотренная вероятность равна априорной вероятности. Если X и A не связаны – то есть статистически независимы – тогда нахождение, что X истинно, не может быть свидетельством, что A истинно; наблюдение X не меняет нашу вероятность для A; утверждение, что "X" не является аргументов в пользу A. В этом случае, свидетельство X (что вы оптимист) не заставляет нас обновлять нашу вероятность для A (что ядерная война уничтожит человечество в следующие сто лет).

Но предположим, что оптимист говорит: "Эй, но, поскольку я оптимист, то я надеюсь на завтра, я буду работать усерднее, немного подтолкну глобальную экономику и, наконец, мои усилия помогут выделить немного лишних денег на исследования ученому, который найдет способ предотвратить ядерную войну – как вы можете видеть, эти два события все же связаны, так что я могу использовать одно из них как свидетельство для другого."

Не так быстро… В каком-то смысле это верно – любая корреляция, не имеет значения насколько слабая, должна учитываться теоремой Байеса; однако теорема различает слабые и сильные свидетельства. То есть теорема Байеса не просто говорит нам что является свидетельством, а что нет, она также описывает силу свидетельства. Теорема Байеса не просто говорит нам пересматривать наши вероятности, но и говорит насколько мы должны их пересматривать. Корреляция между надеждой и предотвращением войны, возможно, и существует, однако она гораздо слабее, чем говорящий хотел бы; он пересматривает свою вероятность на слишком большую величину.

Статистические модели рассматриваются через призму байесовского метода потому что байесовские статистики хороши настолько, насколько возможно: "Байесовский метод определяет максимальный запас расстояния, который вы можете получить из данного куска свидетельства так же как термодинамика определяет максимальный запас работы, который может быть произведен из разницы температур." Вы также услышите, что когнитивные ученые сравнивают субъектов, принимающих решения с идеальными байесовскими агентами, поскольку когнитивные искажения рассматриваются как отклонения от идеального байесовского мышления.

Юдковски делает вывод. 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 $\sim X$ 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.

Вы даже можете формализовать попперовскую философию математически. Отношение правдоподобия для X, $p(X|A)/p(X|\sim A)$, определяет насколько наблюдение X сдвигает вероятность для A; отношение правдоподобия говорит, насколько сильно X в качестве свидетельства. Итак, в вашей теории A, вы можете предсказать X с вероятностью 1, если вам нравится; но вы не можете контролировать знаменатель отношения правдоподобия, $p(X|\sim A)$ – там всегда будут какие-то альтернативные теории которые тоже предсказывают X, и в то время как мы работаем с простейшей теорией, которая объясняет текущее свидетельство, вы когда-нибудь можете найти определенное свидетельство, которое объясняется альтернативной теорией, а вашей нет. Вот тут вы и попадаетесь. Здесь и есть предел пользы, которую можно извлечь из успешных предсказаний; здесь предел того, насколько большим может быть отношение правдоподобия для подтверждающего свидетельства.

С другой стороны, если вы считаете, что определенный кусок свидетельства Y определенно не предсказывался вашей теорией, это невероятно сильное свидетельство против вашей теории. Если p(Y|A) бесконечно мало, то отношение правдоподобия также будет бесконечно малым. Для примера, если p(Y|A) равно 0.00001% и p(Y|~A) равно 1%, то отношение правдоподобия p(Y|A)/p(Y|~A) будет 1:10000. –40 децибелов свидетельства! Или, если поменять все на противоположное, если p(Y|A) очень маленькое, тогда p(Y|~A)/p(Y|A) будет очень большим, что значит что наблюдение Y в значительной степени свидетельствует о превосходстве ~A над A. Фальсификация сильнее чем подтверждение. Это последствие ранней точки зрения, что очень сильное свидетельство это не продукт высокой вероятности, что A ведет к X, но продукт очень низкой вероятности, что не-A ведет к X. Это точное байесовское правило, которое лежит в основе эвристической ценности фальсификации Поппера.

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!

Бонусные ссылки:

- Yudkowsky's original: An Intuitive Explanation of Bayes' Theorem
- BetterExplained: An Intuitive (and Short) Explanation of Bayes' Theorem
- komponisto: Bayes' Theorem Illustrated (My Way)

Бонусная задача. В левой руке Морфеуса лежит 7 синих и 3 красных таблетки, а в правой 5 синих и 8 красных. Вы закрываете глаза и берете таблетку – она оказывается красной, однако вы не знаете из какой руки ее взяли. Какова вероятность, что вы взяли ее из правой руки?

Решение. По условию задачи: $\mathrm{p}(\text{красная}|\text{правая}) = 8/13$, $\mathrm{p}(\text{красная}|\sim\text{правая}) = 3/10$, $\mathrm{p}(\text{правая}) = 13/23$, $\mathrm{p}(\sim\text{правая}) = 10/23$. Применим формулу Байеса:

$$p(\text{правая}|\text{красная}) = \frac{p(\text{красная}|\text{правая}) * p(\text{правая})}{p(\text{красная}|\text{правая}) * p(\text{правая}) + p(\text{красная}|\sim\text{правая}) * p(\sim\text{правая})}$$

Получаем:

$$p(\text{правая}|\text{красная}) = \frac{\frac{8}{13} * \frac{13}{23}}{\frac{8}{13} * \frac{13}{23} + \frac{3}{10} * \frac{10}{23}} = \frac{8}{11}$$

Похоже, что задачу можно решить проще: в обеих руках 11 красных таблеток. Если я вытащил красную, то с вероятностью 8/11 она из правой руки, и с вероятностью 3/11 – из левой…