

UNITED WISDOM

Вероятность и частота: где интуиция теряет масштаб

Мир и сознание

Andrey Zarinsh & Maria Evilkut

W-03-N04 · RU

@andrey.zarinsh @ilkutki

Вероятность и частота: где интуиция теряет масштаб

Эпиграф

Но это предполагает равновозможность различных случаев.

Лаплас, «Философский опыт о вероятностях», главаIII, начало второго принципа. Собственный перевод авторов статьи с английского перевода Траскотта и Эмори; речь об условии предшествующего счёта случаев.

О чём эта статья

Вероятность легко услышать как обещание, а частоту — принять за точное описание будущего. Между тем фраза «так было в четырёх случаях из двадцати» сообщает о выбранном наборе наблюдений; фраза «шанс следующего случая равен одной пятой» требует дополнительного основания. В этой статье мы различим наблюдаемую долю, вероятность внутри заданной модели и решение при неопределённости. Небольшие вымышленные примеры с карточками и библиотечными книгами покажут роль знаменателя, условий отбора, независимости и направления вопроса. Фрагменты Лапласа и Юма помогут увидеть, что расчёт возможных исходов и перенос опыта в ожидание — разные задачи. Екклесиаст будет рассмотрен отдельно как религиозное размышление о пределах гарантированного результата, без превращения его в статистическую теорию. Мы не будем приписывать читателю неизбежную ошибку или универсальный психологический механизм. Практическая цель — научиться точнее читать числовое утверждение и понимать, какое решение оно действительно поддерживает, а какая уверенность добавлена сверх имеющихся сведений.

@andrey.zarinsh @ilkutki

I. Одна пятая чего именно

Представим вымышленную библиотеку, где за двадцать проверенных дней четыре раза пришлось переносить небольшую встречу. Доля перенесённых встреч среди этих двадцати составляет одну пятую. Пока это описание выбранных случаев, а не свойство всех библиотечных встреч. Нужно знать, что включено в двадцать: все назначенные мероприятия, только завершившиеся записи или дни, о которых сохранилась переписка. Если несколько несостоявшихся встреч вовсе отсутствуют в журнале, изменится и набор наблюдений. Число может быть вычислено правильно и всё же отвечать не тому вопросу, ради которого его используют. Арифметика не исправляет незаметно изменённый предмет подсчёта.

Уточним ещё слово «перенести». Одни участники считают переносом новую дату, другие включают начало на час позже. До сравнения нужно выбрать значение. В противном случае четыре случая одного списка могут оказаться шестью случаями другого, хотя оба составителя добросовестно читали записи. Общий вопрос требует общего признака и доступного основания его применения. Можно сохранить две разные задачи, если различие обозначено: перенос даты и задержка начала. Но нельзя поставить их числа рядом без пояснения и назвать разницу противоречием данных. Отдельное число не содержит автоматически собственного определения.

Полезная фраза выглядит длиннее привычного процента: «Среди этих двадцати назначенных встреч, учтённых по такому правилу за такой период, дату перенесли четыре раза». Она сообщает и долю, и границу. Сокращение до двадцати процентов возможно после того, как граница уже понятна. Когда число переходит в другой разговор, исходное пояснение тоже нужно сохранить. Иначе слушатель может услышать утверждение обо всех будущих встречах. Точность состоит не в количестве знаков после запятой, а в возможности восстановить, что было сосчитано. Иногда дробь четыре из двадцати лучше сохраняет масштаб материала, чем отдельно стоящий процент.

II. Вероятность требует устройства задачи

Для другого вопроса зададим искусственную модель: шесть одинаково доступных карточек с числами от одного до шести. Одну выбирают случайно; после каждого выбора возвращают и вновь перемешивают. В этой модели заранее указанное число имеет вероятность одну шестую в одном выборе. Речь идёт о заданных возможностях, а не о доле уже наблюдавшихся результатов. Даже до первого выбора можно вычислить дробь из условий модели. Однако утверждение об одинаковой доступности не следует из количества надписей. Оно должно входить в условия или быть отдельно обосновано для реального устройства. Перечень исходов ещё не устанавливает их равновозможность.

В начале главы III «Философского опыта о вероятностях» Лаплас прямо уточняет условие равновозможности случаев. Прочитанный фрагмент доступен в переводе Траскотта и Эмори с обозначенного шестого французского издания. Мы используем этот ограниченный математический вопрос, а не принимаем весь исторический взгляд автора на мир. Если две большие стороны монеты перечислены как два исхода, это само по себе не исследование её поведения. Аналогично шесть карточек в нашей истории задают модель только после нужных условий выбора. Нельзя перейти от «мы можем назвать шесть вариантов» к «каждый имеет одинаковый шанс» без дополнительного шага.

В библиотеке варианты тоже можно перечислить: встреча состоится, перенесётся или отменится. Из трёх названий не получается по одной трети на каждый вариант. Они могут зависеть от расписания, помещения и решений людей. Более того, категории могут перекрываться: сначала перенесённую встречу позже отменили. Тогда нужно определить момент оценки и взаимоисключающие исходы. Хорошая модель не обязана быть сложной, но обязана соответствовать вопросу. Иногда её пока нет. В таком случае доступна частота прошлого набора и остаётся открытой вероятность будущего случая; неизвестность не нужно закрывать красивой дробью, полученной из числа слов.

Даже группировка исходов влияет на впечатление. «Шестёрка» и «любое другое число» дают два названия, но в нашем наборе им соответствуют одна и пять карточек. Их вероятности поэтому различаются. Названий два, возможностей выбора шесть. Это небольшое различие достаточно, чтобы показать, почему словесное разделение вариантов должно быть сопоставлено с устройством испытания.

III. Опыт становится ожиданием не автоматически

Юм в разделе VI «Исследования о человеческом разумении» рассматривает вероятность и переход прошлых случаев в ожидание. Среди различных прежних результатов он предлагает не упускать остальные, даже когда один встречался чаще. Это философский разбор уверенности и опыта, не современное исследование библиотечных мероприятий. Он позволяет уточнить наш вопрос: на какой прежний набор мы опираемся и какие различия между нынешним случаем и прежними сохраняются? Четыре переноса могут быть важной частью опыта, но число ещё не показывает, что условия следующей встречи будут такими же. Сам переход требует объяснения, соответствующего предмету.

Предположим, все четыре переноса связаны с ремонтом одного помещения, а ремонт закончился. Тогда прежняя доля описывает период с определённым обстоятельством. Она не исчезает из истории, но её использование для следующего месяца нужно пересмотреть. Возможна и обратная ситуация: новых записей мало, зато доступное расписание прямо показывает конфликт двух мероприятий. Общая частота не должна заслонять сведения о конкретном случае. Чтобы пользоваться опытом, полезно

спросить, что в нём сопоставимо, какие условия изменились и не пропадает ли важная деталь при объединении наблюдений. Одной формулы для любого переноса здесь не предлагается.

Иногда организатор может сказать лишь: «В прошлой записи доля была такой; подходящего основания для точного шанса следующего случая у меня нет». Это содержательный ответ, а не отказ мыслить. Можно дополнить его доступным действием: проверить помещение, заранее уточнить дату или подготовить способ сообщить об изменении. Решение допускает неопределённость, если она названа. Исторические философские вопросы помогают не путать описание опыта с гарантией; они не заменяют необходимый сбор нынешних сведений. Уверенность может расти после нового основания, а не только после повторения прежнего числа более решительным голосом.

IV. Короткая серия не распределяет исходы равномерно

Вернёмся к карточкам. Если пять выборов не дали число шесть, в заданной независимой модели вероятность шестёрки при следующем выборе остаётся одной шестой. Карточка возвращена, условия выбора восстановлены, прошлые результаты не меняют состав набора. Из этого не следует, что в реальной системе условия действительно восстанавливаются: если карточки не возвращаются, вопрос другой. Именно поэтому слово «независимо» должно иметь содержание, а не быть украшением расчёта. Оно обозначает нужное отношение между выборами в нашей модели. История результатов не обязана управлять следующим исходом только потому, что мы ожидаем более ровную картину.

В шести независимых выборах не обещана одна шестёрка. Возможны отсутствие шестёрок, одна и несколько. Среднее количество в такой модели равно одному, но это характеристика распределения, а не обязательный результат одной серии. Для вероятности хотя бы одной шестёрки можно рассмотреть противоположный случай: шесть раз выбрать другое число. Его вероятность равна произведению шести долей пять шестых. Вероятность хотя бы одной — единица минус этот результат. Она не равна единице. Вычисление объясняет, почему «шесть попыток при шансе одна шестая» не следует переводить как гарантированное получение нужной карточки.

Этот пример не устанавливает, что библиотечные переносы независимы. Один перенос может изменить расписание и вызвать другой. Одна общая причина может затронуть несколько встреч. Поэтому формула карточек остаётся в своей области. Чтобы применить похожий расчёт к мероприятиям, пришлось бы определить процесс, зависимости и подходящие сведения. Если этого нет, лучше не умножать проценты только потому, что действие умножения знакомо. Числовая уверенность должна сохранять условия, из которых получена. Чем проще пример, тем яснее можно увидеть

ограничение; простота не является разрешением переносить его на любые реальные события.

V. Направление вопроса меняет знаменатель

Вымышленный пример с книгами покажет другое различие. Представим заранее составленный учебный набор из тысячи книг. Двадцать из них имеют повреждённый переплёт. Условная система пометила шестнадцать из этих двадцати и ещё восемьдесят книг без повреждения. Это придуманные числа, не измерение настоящей системы и не обещание качества проверки. Среди повреждённых книг помечены шестнадцать из двадцати, то есть восемьдесят процентов. Но среди всех помеченных книг повреждены шестнадцать из девяноста шести, то есть одна шестая. Обе доли вычислены из одного набора. Они отвечают на разные вопросы.

Первая доля спрашивает: «Если книга повреждена, насколько часто она оказалась помеченной в этом наборе?» Вторая: «Если книга помечена, насколько часто она оказалась повреждённой?» Поменялись условие и знаменатель. Из восьмидесяти процентов нельзя получить вторую долю простым изменением порядка слов. Чтобы понять пометку, нужно учитывать и повреждённые, и неповреждённые книги, которые её получили. Большая группа неповреждённых книг способна дать много пометок, даже если каждая отдельная доля пометок в ней сравнительно невелика. Масштаб группы составляет часть вопроса, а не случайную подробность, которую можно убрать.

В настоящем каталоге сначала пришлось бы проверить определение повреждения и способ составления набора. Но уже искусственная таблица показывает логическую ошибку перестановки условия. Она не сообщает, какие показатели у реальной библиотеки и что делать с конкретным переплётом. Для действия можно проверить книгу непосредственно или уточнить, как устроена пометка. Учебный расчёт полезен как раз своей ограниченностью: мы знаем все числа и поэтому можем проследить переход. Если число перенесено в совет без соответствующего вопроса, даже верная арифметика перестаёт поддерживать тот смысл, который слушатель извлёк из неё.

VI. Большое число случаев не исправляет отбор

Количество записей имеет значение, но прежде него нужно понять, какие записи включены. Если о переносах отвечают только организаторы, которым хотелось поделиться удачной работой, тысяча ответов описывает этих ответивших. Она не становится автоматически всеми библиотеками. Если список содержит только завершённые встречи, отменённые могли исчезнуть ещё до подсчёта. Нельзя устранить такой пробел простым расширением того же списка. Нужен другой способ составления материала или более узкий вывод. Статья не оценивает неизвестное исследование; она показывает, какое условие надо спросить, прежде чем принять масштаб числа за масштаб применимости.

Также нужно знать, что считается единицей наблюдения. Один посетитель мог оставить несколько отзывов, одна встреча фигурировать в двух журналах, одно сообщение перепечататься в разных местах. Число строк тогда не равно числу независимых случаев. Для подсчёта доли может понадобиться сопоставить записи и обозначить повтор. Это не обвинение составителя: материал мог создаваться для другой задачи. Журнал сообщений подходит для истории общения и не обязательно для подсчёта мероприятий. Назначение исходного документа влияет на то, какой вопрос к нему можно задать. Переход к новой задаче требует проверки устройства данных.

Два периода тоже нельзя объединять бездумно. Допустим, первая небольшая группа встреч проводилась во время ремонта, а вторая — после него. Общая доля описывает смесь, но может плохо отвечать на вопрос о следующей встрече после ремонта. При этом разделение на всё более мелкие группы имеет предел: в каждой может остаться слишком мало материала для сильного вывода. Мы не предлагаем универсальный размер достаточной выборки. Такой вопрос зависит от задачи и метода. Наш общий шаг скромнее: назвать группу, время и способ включения, а затем показать, к какой части нынешнего решения относится полученная доля.

При объединении групп нужна и правильная арифметика масштаба. В ещё одном искусственном наборе первая группа содержит десять встреч и пять переносов, вторая — девяносто встреч и девять переносов. Их доли равны пятидесяти и десяти процентам. Общая доля — четырнадцать переносов из ста, то есть четырнадцать процентов, а не среднее двух процентов, равное тридцати. Группы имеют разное количество встреч и поэтому входят в общую долю с разным весом. Это чистый подсчёт выбранного набора, не рекомендация объединять любые периоды. Перед объединением всё равно остаётся вопрос об общей задаче и сопоставимости определений. Верный расчёт смеси не доказывает, что смесь подходит для нашего будущего случая.

VII. Частота не содержит причины

Если встречи вечером переносятся чаще, чем утром, наблюдаемая разница ещё не показывает, что причиной является вечер. Вечером могли назначать более крупные мероприятия, использовать другое помещение или чаще зависеть от доступности гостей. Даже добросовестная запись всех случаев не выбирает автоматически между такими объяснениями. Можно искать дополнительные сведения и сопоставимые условия. Можно сохранить найденную разницу без утверждения причины. Разделение не лишает числа пользы: оно сообщает, где возник вопрос. Но из описания распределения нельзя сразу вывести, что перенос встречи на утро устранит обстоятельство, из-за которого менялась дата.

Обратное тоже возможно: причина известна из конкретного распоряжения, хотя общей таблицы переносов нет. Если помещение закрыто в определённый день, запись о закрытии может прямо относиться к выбранной встрече. Тогда не требуется ждать большой частоты, чтобы уточнить место. Разные основания отвечают на разные вопросы. Число наблюдений не является единственным видом знания; непосредственное сообщение об условии может иметь самостоятельное значение. Нужно проверить, что оно относится к нужной дате и месту. Мы сравниваем доступные основания, не назначая статистическую долю обязательной заменой любого другого сведения.

В предложенной читательской работе полезно написать два предложения отдельно: «Я вижу такую долю случаев» и «Я предполагаю такое объяснение». Затем назовите переход между ними. Если переход не найден, объяснение остаётся предположением. Если нужно принять решение, можно выбрать ограниченное действие, которое не требует считать причину установленной: уточнить расписание или предусмотреть запас времени. Удобство решения не доказывает причинную версию. И наоборот, неопределённость причины не запрещает действовать при достаточных сведениях о конкретном условии. Эти различия удерживают число в той роли, которую оно действительно может выполнить.

VIII. Ожидаемое число и ценность исхода

Вероятность события не сообщает, насколько оно желательно или приемлемо. Два организатора могут понимать одну и ту же неопределённость и выбирать по-разному. Один допускает изменение даты, другой приезжает только на этот вечер. Их расхождение не обязательно означает ошибку расчёта. Оно может касаться последствий для каждого. Сначала нужно понять, одинаково ли они прочли число; затем обсудить условия участия. Перевод предпочтения в заявление о «правильной вероятности» способен скрыть настоящий вопрос. Человек вправе отказаться от неудобной неопределённости без доказательства, что событие почти наверняка не состоится.

В карточной модели среднее количество нужных результатов равно одному на шесть выборов, но это не выбирает за нас, стоит ли делать серию. Для решения нужны цена действия, значение результата и допустимые варианты. Они не содержатся в дроби одна шестая. В этой статье мы не предлагаем денежной стратегии или оптимального порога для реального риска. Нам достаточно увидеть границу: описание возможных результатов и оценка их значения — разные части решения. Даже точный расчёт не превращает предпочтение одного участника в обязанность всех остальных. Число помогает разговору, если его роль ясна.

В Екклесиасте 9:7–12 способность и труд соседствуют с признанием того, что желаемый исход не гарантирован. Фрагмент прочитан в King James Version и сохраняет религиозную область. Он не устанавливает независимость карточных выборов, частоты переносов или допустимый порог участия. Наше сопоставление касается ограниченности гарантии, а не общей математической теории. Из него здесь не выводится бесполезность подготовки. Подготовка, неопределённость и ценность деятельности могут обсуждаться рядом, не сводясь к одному проценту. Религиозный вопрос о жизни и математический вопрос о заданных исходах остаются различными вопросами.

IX. Как сообщить число без лишнего обещания

Организатор может дать участникам полезное сообщение: в таком периоде из двадцати назначенных встреч четыре сменили дату; для нынешней встречи помещение подтверждено, а участие гостей ещё уточняется. Здесь прошлая доля не скрыта и не превращена в точный шанс будущего. Названы сведения о нынешнем случае. Слушатель понимает, что уже известно и какая часть решения остаётся открытой. Сообщение можно сократить, но важно не потерять переходы. Отдельное «надёжность восемьдесят процентов» оставило бы неизвестными и определение надёжности, и группу наблюдений, и основание переноса результата на нынешнюю дату.

Если используется оценка вероятности, полезно назвать её происхождение: расчёт в заданной модели, вывод из определённого материала или мнение при недостатке данных. Эти основания нельзя автоматически уравнять. Они могут быть подходящими для разных задач, но требуют разных вопросов. Важно также не выдавать отсутствие сведений за нулевой шанс, а отсутствие гарантии — за равенство всех исходов. Иногда точная цифра пока не оценена. Это можно сказать прямо, дополнив ответ доступными фактами. Честное описание неизвестности помогает выбрать следующий шаг и не обязывает украшать её псевдоточностью.

Проверка сообщения возможна с обеих сторон. Составитель спрашивает, поддерживают ли материалы его слова. Читатель спрашивает, какой вывод он сам добавляет. «Четыре переноса» может превратиться в воображении в «четыре случая из всех возможных», если граница забыта. Мы не объявляем этот переход неизбежным свойством человека; в условной ситуации он просто возможен. Разговор позволяет его уточнить. При обнаружении ошибки лучше исправить конкретное число, группу или условие и объяснить изменение. Нет необходимости объявлять бесполезной всю прошлую запись, если исправление относится к одному определённому переходу.

Особенно важно различать ноль записанных случаев и невозможность события. Если в двадцати наблюдениях переносов не было, наблюдаемая доля равна нулю. Из этого без иных оснований не следует нулевая вероятность следующего переноса. В карточной модели определённое число может не появиться в ограниченной серии, хотя его вероятность задана как ненулевая. Реальная библиотека не становится карточной моделью от этого сравнения; пример лишь показывает различие утверждений. Чтобы назвать будущее событие невозможным, нужны соответствующие условия, а не только отсутствие его в небольшом списке. Точно так же один наблюдавшийся перенос не устанавливает, что он обязательно повторится. Ноль и единица в записи сохраняют свой предмет: сколько подходящих случаев оказалось именно в выбранном материале.

X. Масштаб удерживается условиями

Частота сохраняет связь с набором наблюдений, вероятность — с устройством вопроса и основанием оценки, решение — с последствиями и предпочтениями. Эти отношения легко потерять при сокращении фразы до одного процента. Возвращение условий не делает числовое знание слабее. Оно показывает его действительное применение. Одна пятая может точно описывать двадцать случаев, одна шестая — заданную карточную модель, а одна шестая среди помеченных книг — другую долю в искусственном наборе. Одинаковая дробь не означает одинаковое основание. Именно понимание основания позволяет пользоваться числом вне первоначальной красивой формулировки.

Лаплас в выбранном фрагменте помогает спросить о равновозможности и соединении событий; Юм — о переносе разных прежних результатов в ожидание; Екклесиаст — о жизни без гарантированного желаемого исхода. Сравнение не создаёт одного общего учения и не заменяет современные методы исследования. Мы использовали исторические вопросы для оригинального разбора небольших примеров. Если потребуется оценка реального процесса с существенными последствиями, её нужно строить на подходящих данных и методе. Наша практика готовит более ясный вопрос, а не выдаёт готовое число для всех ситуаций.

После такого разбора человек может пересчитать долю, уточнить, о каком случае идёт речь, или оставить будущую вероятность неоценённой. Каждый из этих результатов может быть полезным. Необязательно заканчивать полной уверенностью. Достаточно знать, что вычислено, какие условия приняты и какое действие ими поддерживается. Масштаб становится видимым, когда возвращены числитель, знаменатель, правило включения и предел вывода. Тогда знакомое число перестаёт быть общим обещанием и становится определённым знанием, которое можно обсудить, проверить и при необходимости изменить.

Что это значит на практике

Выберите одно небольшое числовое утверждение, которое уже влияет на ваш выбор. Запишите его полностью и спросите: это доля прошлых случаев, вероятность в заданной модели или оценка будущего? Для доли назовите числитель, знаменатель, период и способ включения. Для модели — возможные исходы и основания их равновозможности либо различной вероятности. Если этих сведений нет, не подставляйте произвольное число.

Проверьте направление вопроса. «Какая часть повреждённых книг помечена?» и «Какая часть помеченных повреждена?» требуют разных знаменателей. Затем отделите расчёт от предположения о причине и от оценки желательности результата. Одно число не отвечает на все три вопроса. Если утверждается независимость нескольких исходов, уточните, что обеспечивает её в выбранной задаче.

Выберите один следующий шаг: восстановить группу наблюдений, пересчитать простую долю, уточнить нынешнее условие или сохранить вероятность как пока неоценённую. Запишите, что изменилось после этого шага. Для серьёзного решения может потребоваться подходящая специализированная работа; маленькое упражнение её не заменяет. Это авторская практика чтения числовых утверждений, без обещания улучшить каждое решение. Её можно остановить после одного ясного вопроса, не превращая всю повседневность в постоянный подсчёт.

Ключевая цитата

> *мы не должны упускать другие результаты*

Часть предложения Юма, «Исследование о человеческом разумении», VI.4, обсуждение разных прежних исходов и их веса в ожидании. Собственный перевод авторов статьи с английского текста; явно ограниченный фрагмент, не самостоятельное правило расчёта.

Связи между традициями

1. **Лаплас / Юм:** расчёт по заданным случаям и перенос опыта в ожидание касаются неопределённости, но имеют разные основания. Философская уверенность не заменяет математического условия.

2. **Лаплас / наши карточки:** равновозможность и независимость входят в авторскую учебную модель. Они не устанавливаются одним перечнем названий или историческим авторитетом.

3. **Юм / библиотечная запись:** прежние разные исходы помогают задать вопрос об ожидании. Наша доля относится к вымышленному набору; философский текст не оценивает реальную библиотеку.

4. **Екклесиаст / два философско-математических фрагмента:** негарантированный исход — вопрос для сопоставления. Религиозная жизнь, философское основание ожидания и вычисление вероятности сохраняют разные предметы.

5. **Три источника / решение участника:** различение основания и гарантии помогает нашему разговору. Условия участия и добровольная практика предложены авторами; общей доказанной техники источники не устанавливают.

Первоисточники

1. Лаплас. «Философский опыт о вероятностях», выбранные фрагменты глав II–III; перевод Ф. У. Траскотта и Ф. Л. Эмори, английское издание1902. Прочитан ограниченный фрагмент условий расчёта.

2. Юм. «Исследование о человеческом разумении», VI, §§1–4 и примечание10; английский текст Hume Texts Online. Философское обсуждение вероятности и ожидания.

3. Екклесиаст9:7–12, King James Version. Прочитанный религиозный фрагмент о жизни и негарантированном исходе; не статистическая теория.

Теги

вероятность, частота, масштаб, доля, числитель, знаменатель, модель, равновозможность, независимость, условие, условная вероятность, ожидание, опыт, Юм, Лаплас, Екклесиаст, карточки, библиотека, отбор, повтор записи, объединение групп, причина, неопределённость, ноль наблюдений, среднее, гарантия, предпочтение, решение, проверка, читательская практика

Про авторов

Andrey Zarinsh — автор и основатель проекта United Wisdom. Исследует темы внутренней трансформации, единства традиций, межкультурного диалога и современной духовной дисциплины. Instagram: @andrey.zarinsh.

Maria Evilkut — соавтор United Wisdom. Помимо исследовательской работы, она стремится проверять эти идеи на практике и наблюдать, как принципы единства и взаимопонимания работают в реальной жизни между людьми и культурами. Instagram: @ilkutki.