Краткое введение в различия между частотным и байесовским подходами к вероятности в оценке стоимости

К. А. Мурашев

31 декабря 2021 г.

При словах «математическая статистика» и «статистические методы» большинство тех, кто не является математиком или специалистом по машинному обучению, вспоминают такие понятия как «нулевая гипотеза», «уровень значимости», «статистика критерия» и многое другое, что преподаётся на курсах статистики в вузах. Возможно для кого-то станет открытием: всё это лишь часть того, что есть современная наука о вероятности. Традиционно изучаемые методы относятся к одному направлению: статистике, основанной на частотном подходе к вероятности [8]. Это обстоятельство в первую очередь объясняется тем, что данное направление статистики существует уже не одно столетие и является исторически первым известным человечеству. При этом такой подход позволяет создавать описательные и предсказательные модели, обеспечивающие хорошую интерпретацию результатов, доступную для понимания даже со стороны не специалистов. Действительно, когда британские учёные совершают очередной научный прорыв, доказав, что количество пластинок с записями Вагнера находится в тесной связи с объёмом грудной клетки их владельца, любой читатель может сформировать общее представление о вопросе.

Однако наука не стоит на месте. Примерно с 1950-х годов XX века распространение начал получать и другой подход к понятию вероятности — т. н. байесовский подход к вероятности [4]. Данный подход стал особенно популярным в последние 20–30 лет, и на сегодняшний день можно говорить о том, что именно он определяет развитие такой области как машинное обучение, являющейся предметом интереса всего исследования применения методов искусственного интеллекта в оценочной деятельности. Забегая вперёд, можно сказать, что эти два подхода не являются взаимоисключающими, используют один математический аппарат и при определённых условиях один из них перерождается в другой.

При этом, следует признать, что даже частотный подход к вероятности, равно как и математические методы в целом, пока что не получили широкого распространения в практике российской оценки. Данный материал призван сформировать представление о сути каждого из подходов к вероятности, помочь оценщикам получить некоторое представление о возможных путях развития методов оценки, а также о дальнейших перспективах развития отрасли и своём месте в ней.

1. Историческая справка

Невозможно сказать, когда именно появились первые представления о вероятности и статистических свойствах явлений, объектов и процессов. Известно лишь то, что появились они вследствие любви человека к двум вещам:

- деньгам;
- играм.

Первые известные работы, в которых были рассмотрены вопросы вероятности, описывают различные ситуации, связанные с азартными играми. Первым серьёзным общетеоретическим научным выводом о вероятности можно считать высказанное Д. Кардано утверждение о том, что *«при небольшом числе игр реальное количество исследуемых событий может сильно отличаться от теоретического, но чем больше игр в серии, тем меньше доля этого различия»* [6]. Известный междисциплинарный исследователь Галилео Галилей в 1718 г. опубликовал свой трактат «О выходе очков при игре в кости», в котором детально описал многие вопросы вероятности. В своей главной работе «Диалог о двух главнейших системах мира, птолемеевой и коперниковой» Галилей также указал на возможность оценки погрешности астрономических и иных измерений, причём заявил, что малые ошибки измерения вероятнее, чем большие, отклонения в обе стороны равновероятны, а средний результат должен быть близок к истинному значению измеряемой величины. Эти качественные рассуждения стали первым в истории предсказанием нормального распределения [3] ошибок.

Дальнейшее развитие науки о вероятности произошло благодаря переписке Блеза Паскаля и Пьера Ферма, вызванной поставленной шевалье де Мере т. н. «задачей об очках», также посвящённой вопросам азартных игр. в ходе данной переписки зародился ещё один раздел математики — комбинаторика. Обсуждение подобных вопросов вдохновило Христиана Гюйгенса, предложившего собственное решение задачи, а также опубликовавшего работу «О расчётах в азартных играх». При этом Гюйгенса стал первым исследователем, обратившим внимание на серьёзную научную основу вероятностных задач. Предисловие к вышеуказанной работе гласило «Я полагаю, что при внимательном изучении предмета читатель заметит, что имеет дело не только с игрой, но что здесь закладываются основы очень интересной и глубокой теории». Гюйгенс впервые ввёл понятие математического ожидания, т. е.

теоретического среднего арифметического случайной величины, дал окончательное решение задачи о разделе ставок при досрочном прекращении игры.

История изучения вопросов вероятности в XVIII веке связана прежде всего с именами Бернулли, Муавра, Эйлера и Лапласа, разработавших первые законы распределения. Примерно в это же время фокус внимания переключился с игр на более практические области: демографию, экономику, страхование. В 1763 году Томас Байес опубликовал свою знаменитую формулу, ставшую спустя примерно 200 лет основной нового подхода к вероятности. Тем не менее, несмотря на его открытие, практическое распространение в тот период получал лишь частотный подход к вероятности.

XIX век дал существенное число работ по статистике, перечислить даже малую часть которых не представляется возможным в рамках данного фрагмента. Следует отметить, что итоговое формирование частотной математической статистики произошло в 1930-е годы вследствие работ К. Пирсона, Р. Фишера и А. Н. Колмогорова.

Байесовский подход к вероятности начал получать распространение лишь после публикации работ Ф. Рэмси «Истина и вероятность» (1926), Б. де Финетти «Предвидение: его логические законы, его субъективные источники» (1937). Однако его широкое признание произошло только после выхода в свет работы Л. Сэвиджа «Основания статистики» (1954). Суть субъективной интерпретации вероятности можно выразить словами Рэмси: «Степень уверенности (belief) — это её каузальное свойство (causal property of it), которое мы можем приблизительно сформулировать как степень, в какой мы готовы действовать в соответствии с нашей уверенностью».

2. Современное состояние

Как уже было сказано ранее, на сегодняшний день существуют два практических подхода к понятию вероятности: классический частотный и байесовский. Можно сказать, что ключевое различие между ними заключается в интерпретации понятия случайности. Частотный подход гласит о том, что случайность является следствием объективной неопределённости, которая может быть уменьшена только путём проведения серии экспериментов. В байесовском подходе утверждается, что неопределённость является лишь следствием субъективного незнания. Таким образом можно утверждать о том, что разные субъекты могут обладать разным незнанием. Современная наука придерживается позиции о том, что именно байесовская трактовка даёт более корректное понимание окружающего мира, соответствующее текущим представлениям о фундаментальных процессах. Дело в том, что по-настоящему неопределёнными являются всего два известных человечеству процесса:

- квантово-механические процессы;
- радиоактивный распад ядер атомов.[1]

Все остальные явления и процессы на самом деле являются детерминированными, т. е. происходящими в строгом соответствии с некоторыми законами. И лишь

[1] По крайней мере по современным представлениям в области физики.

субъективное незнание законов, в соответствии с которыми они протекают, либо отдельных параметров, являющихся необходимыми для вычисления состояния объекта либо процесса вызывает неопределённость. Классическим примером применения вероятности является эксперимент с подбрасыванием монеты. Несложно догадаться, что незнание исхода подбрасывания не является следствием склонности самой монетки либо неопределённости тех законов механики, в соответствии с которыми происходит её движение. На самом деле, нет никакой сложности в том, чтобы точно предсказать исход броска. Достаточно знать массу монеты, её геометрические свойства, упругость, расположение центра массы, параметры силы, приложенной к монете в момент броска, свойства среды (например кинематическая вязкость), в которой будет происходить её движение, высоту между точкой броска и поверхностью, свойства поверхности (например та же упругость) и ряд иных параметров. Знание этих начальных совершенно конкретных значений, а также умение применять полностью детерминированные законы классической механики позволят абсолютно точно предсказать исход каждого броска. И лишь отсутствие значений этих параметров, а также определённая вычислительная сложность позволяют говорить о непредсказуемости поведения монеты. Т. е. налицо лишь субъективная неопределённость, вызванная личным незнанием наблюдателя, но отнюдь не объективная, вызванная склонностью монеты либо недетерминированностью законов её движения.

Применительно к оценке стоимости можно сказать, что её неопределённость также является лишь субъективной. С точки зрения оценщика, определяющего стоимость, например запасов строительного материала — песка, действительно имеет место неопределённость, которую он стремится уменьшить путём применения ряда методов. Однако с точки зрения многих других субъектов, например собственника предприятия-правообладателя запасов, части его сотрудников, осуществляющих закупочную деятельность, поставщиков, производителей и перевозчиков строительных материалов в данном вопросе нет никакой неопределённости: им хорошо известна стоимость строительного песка на любой стадии его существования от залежей недр до строительной площадки. Кроме того, стоимость строительного песка на каждой стадии его перемещения между этими состояниями прирастает вполне определённым образом — путём привнесения в неё добавленной стоимости, образуемой субъектами того или иного этапа. Таким образом, можно говорить о том, что стоимость конкретного песка на конкретной строительной площадке вполне детерминирована вкладами всех участников производственно-логистической цепочки. И лишь субъективное незнание стороннего наблюдателя, вызванное в том числе множественностью возможных цепочек вызывает у него наличие субъективной неопределённости относительно стоимости песка. При этом вполне вероятна ситуация, когда один оценщик, может обладать знаниями, касающимися, например стоимости песка на всех карьерах территории, к которой относится исследуемая им строительная площадка, но при этом он также обладает незнанием всех остальных слагаемых стоимости. В этом случае можно говорить о том, что он является калиброванным экспертом в части стоимости «франко завод». Второй оценщик может ничего не знать о стоимости песка на карьерах, но при этом отлично знать стоимость логистики сыпучих грузов, являясь в таком случае калиброванным экспертом по вопросам затрат на перевозку. Третий оценщик не обладает никакими знаниями о стоимости самого песка и его перевозки, однако при этом в отличие от двух предыдущих обладает знаниями в вопросах налогообложения, вследствие чего он может делать правильные выводы о факте включения либо невключения НДС в стоимость. Таким образом, каждый из них обладает субъективным незнанием части слагаемых стоимости. Таким образхом в условиях того, что стоимость оцениваемого песка вполне детерминирована сама по себе, для оценщиков она носит случайный характер исключительно в силу их незнания части её слагаемых. При этом каждый из оценщиков обладает разным незнанием, что подчёркивает субъективность неопределённости. Вследствие сказанного выше, можно сделать вывод о том, что, с точки зрения байесовского подхода, плотность распределения случайной величины — это способ закодировать субъективное незнание. В случае отсутствия незнания функция плотности выродилась бы в дельта-функцию [5],[2] являющуюся в таком случае пределом функции Гаусса [7] с дисперсией, стремящейся к нулю. Поскольку чаще всего субъективное незнание всё же имеет место, вместо дельта-функции мы можем наблюдать именно функцию плотности распределения.

Следующее различие между подходами заключается в их отношении к величинам параметров. В частотном подходе существует чёткое разделение на случайные и неслучайные параметры. Типичной задачей является оценка тех или иных параметров генеральной совокупности, представляющей собой набор случайных величин на основе детерминированных параметров выборки, например: среднее, мода, дисперсия и т. д. Последние представляют собой конкретные значения, в которых уже нет никакой случайности. В байесовском же подходе все величины являются случайными. Поскольку в этом случае предпосылкой является отсутствие объективной неопределённости, сами параметры выборки также могут быть описаны не конкретными значениями, а плотностями распределения, в т. ч. априорными. В случае известности значения какого-либо параметра вместо плотности распределения будет использоваться δ-функция. Таким образом, вероятностный аппарат распространяется на все без исключения параметры. В случае точного измерения значения одного из которых происходит коллапс функции его распределения.

Ещё одно отличие между подходами заключается в используемом каждом из них методе вывода (методе статистического оценивания). В частотном подходе используется Метод максимального правдоподобия. [2] Теоретические обоснования говорят о том, что оценки максимального правдоподобия являются оптимальными т. е.:

- несмещёнными;
- состоятельными;
- асимптотически эффективными;
- асимптотически нормальными.

Детальное рассмотрение данного метода не входит в периметр настоящего фрагмента. Скажем только, что оптимальность оценок максимального правдоподобия основана

[2] δ-функция, δ-функция Дирака, дираковская дельта, единичная импульсная функция.

на предпосылке о том, что объём исследуемой выборки стремится к бесконечности. Однако при решении практических задач это почти всегда не так. Например в случае проведения оценки оценщик мог бы найти всего два предложения о продаже карьерного песка интересующей его марки. В силу каких-либо обстоятельств указанные в объявлениях цены могли составить 10 и 15 рублей за 1 куб. м (например оценщик мог бы найти объявления о самовывозе ненужного песка с личных участков). С точки зрения *метода наибольшего правдоподобия* оценкой параметра являлось бы среднее значение 12.5 руб. за 1 куб. м. Однако наличие априорных знаний о характерном значении цены (200-300 руб. за 1 куб. м) не позволило бы сделать вывод о корректности и применимости полученной оценки. Однако в случае анализа нескольких сотен или тысяч объявлений уже любое конкретное значение полученной стоимости могло бы считаться корректным. В байесовском подходе методом вывода служит сама Теорема Байеса:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \tag{1}$$

Следующее различие заключается в получаемых оценках. В частотном подходе получаемая оценка является конкретным числом либо интервалом. В байесовском — апостериорным распределением. Общая схема при этом выглядит следующим образом. К изначальном существовавшему субъективному незнанию, закодированному в функцию плотности априорного распределения добавляются знания о различных параметрах, прямо или косвенно связанных с интересующим нас параметром. Вследствие этого посредством применения Теоремы Байеса априорное распределение трансформируется в апостериорное без потери информации. В случае необходимости в дальнейшем возможно получение точечной оценки, например моды либо матожидания. При этом неизбежно возникает потеря части информации. С точки зрения оценки стоимости можно предложить следующую схему: оценщик имеет некоторое изначальное представление о стоимости объекта. Далее путём внесения новой информации (например о площади объекта оценки) это априорное знание трансформируется в апостериорное, которое является априорным для следующего этапа анализа (на котором, например вносится информация о техническом состоянии объекта). Таким образом, применение байесовского подхода позволяет объединять множество простых моделей в более сложные.

Последнее различие заключается в возможностях применения в зависимости от числа наблюдений. Частотный подход в общем случае требует значительное их число. Байесовский работает при любом. Даже тогда, когда число наблюдений равно нулю. В этом случае апостериорное распределение будет совпадать с априорным. Поступление данных о хотя бы одном наблюдении позволяет скорректировать распределение в правильную сторону. При этом, в случае, когда $n \longrightarrow \infty$ байесовский подход переходит в частотный, что является доказательством высказанного ранее утверждения о непротиворечивости этих подходов относительно друг друга. С математической точки зрения это означает, что при очень большом числе наблюдений апостериорное распределение начинает переходить в дельта-функцию в точке максимального правдоподобия. Общая формула связи между подходами на примере подбрасывания монеты выглядит следующим образом:

$$\frac{m+1}{n+2}, \tag{2}$$

где m — число выпадений «орла», n — общее число бросков. Очевидно, что при $n = 0$ число выпадений «орла» также будет равно нулю, что приводит к тому, что апостериорная вероятность равна априорной и составляет $\frac{1}{2}$. При этом, в случае если $n \longrightarrow \infty$, то и $m \longrightarrow \infty$, что означает, что:

$$\frac{m}{n}. \tag{3}$$

В выражении 3 несложно узнать функцию правдоподобия. В промежуточных случаях, когда число наблюдений больше нуля, но существенно меньше бесконечности байесовский подход позволяет находить компромисс между априорными знаниями и новой вносимой информацией.

Таблица 1. Сводные сведения об основных различиях между подходами к вероятности

	Частотный (Frequentist)	Байесовский (Bayesian)
Интерпретация случайности	Объективная неопределённость	Субъективное незнание
Величины	Чёткое деление на *случайные* и *неслучайные*	Все величины случайны
Метод вывода	Метод максимального правдоподобия	Теорема Байеса
Оценки	Точечные либо интервальные	Апостериорное распределение
Применимость	$n \longrightarrow \infty$ либо, как минимум, $n \gg 1$	$\forall$ n

Источники информации

[1] Computer Science Center. *Введение в вероятностный язык построения моделей машинного обучения.* 2015. URL: `https://www.youtube.com/playlist?list=PLlb7e2G7aSpR8mbaShVBods-hGaFGifkl` (дата обр. 09.10.2021).

[2] Machinelearning.ru. *Метод наибольшего правдоподобия.* URL: `http://www.machinelearning.ru/wiki/index.php?title=%D0%9C%D0%B5%D1%82%D0%BE%D0%B4_%D0%BD%D0%B0%D0%B8%D0%B1%D0%BE%D0%BB%D1%8C%D1%88%D0%B5%D0%B3%D0%BE_%D0%BF%D1%80%D0%B0%D0%B2%D0%B4%D0%BE%D0%BF%D0%BE%D0%B4%D0%BE%D0%B1%D0%B8%D1%8F` (дата обр. 31.12.2021).

[3] Machinelearning.ru. *Нормальное распределение.* URL: `http://www.machinelearning.ru/wiki/index.php?title=%D0%9D%D0%BE%D1%80%D0%BC%D0%B0%D0%BB%D1%8C%D0%BD%D0%BE%D0%B5_%D1%80%D0%B0%D1%81%D0%BF%D1%80%D0%B5%D0%B4%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5` (дата обр. 02.03.2021).

[4] Wikipedia. *Байесовская вероятность*. URL: https://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%B9%D0%B5%D1%81%D0%BE%D0%B2%D1%81%D0%BA%D0%B0%D1%8F_%D0%B2%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C (дата обр. 09.09.2021).

[5] Wikipedia. *Дельта-функция*. URL: https://ru.wikipedia.org/wiki/%D0%94%D0%B5%D0%BB%D1%8C%D1%82%D0%B0-%D1%84%D1%83%D0%BD%D0%BA%D1%86%D0%B8%D1%8F (дата обр. 10.10.2021).

[6] Wikipedia. *Кардано, Джероламо*. URL: https://ru.wikipedia.org/wiki/%D0%9A%D0%B0%D1%80%D0%B4%D0%B0%D0%BD%D0%BE,_%D0%94%D0%B6%D0%B5%D1%80%D0%BE%D0%BB%D0%B0%D0%BC%D0%BE (дата обр. 10.10.2021).

[7] Wikipedia. *Фукнция Гаусса*. URL: https://ru.wikipedia.org/wiki/%D0%93%D0%B0%D1%83%D1%81%D1%81%D0%BE%D0%B2%D0%B0_%D1%84%D1%83%D0%BD%D0%BA%D1%86%D0%B8%D1%8F (дата обр. 12.10.2021).

[8] Wikipedia. *Частотная вероятность*. URL: https://ru.wikipedia.org/wiki/%D0%A7%D0%B0%D1%81%D1%82%D0%BE%D1%82%D0%BD%D0%B0%D1%8F_%D0%B2%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%8C (дата обр. 09.09.2021).