С.Н. Бернштейн

Теория вероятностей

Москва
«Книга по Требованию»

УДК 51
ББК 22.1
С11

С.Н. Бернштейн

С11 Теория вероятностей / С.Н. Бернштейн – М.: Книга по Требованию, 2015. – 367 с.

ISBN 978-5-458-26197-5

«Курс теории вероятностей» Бернштейна, по отзывам весьма многих специалистов, является выдающимся произведением мировой теоретико-вероятностной литературы. Этот курс написан Сергеем Натановичем Бернштейном на основе его собственной аксиоматики (первое аксиоматическое построение теории вероятностей было выполнено Бернштейном еще в 1917 г.). Курс Бернштейна содержит целый ряд оригинальных исследований автора, большое число весьма интересных задач, и думается, что он до сих пор не утратил своего значения и может быть полезен всем, кто интересуется теорией вероятностей.

ISBN 978-5-458-26197-5

Эта книга является репринтом оригинала, который мы создали специально для Вас, используя запатентованные технологии производства репринтных книг и печати по требованию.

Сначала мы отсканировали каждую страницу оригинала этой редкой книги на профессиональном оборудовании. Затем с помощью специально разработанных программ мы произвели очистку изображения от пятен, клякс, перегибов и попытались отбелить и выровнять каждую страницу книги. К сожалению, некоторые страницы нельзя вернуть в изначальное состояние, и если их было трудно читать в оригинале, то даже при цифровой реставрации их невозможно улучшить.

Разумеется, автоматизированная программная обработка репринтных книг – не самое лучшее решение для восстановления текста в его первозданном виде, однако, наша цель – вернуть читателю точную копию книги, которой может быть несколько веков.

Поэтому мы предупреждаем о возможных погрешностях восстановленного репринтного издания. В издании могут отсутствовать одна или несколько страниц текста, могут встретиться невыводимые пятна и кляксы, надписи на полях или подчеркивания в тексте, нечитаемые фрагменты текста или загибы страниц. Покупать или не покупать подобные издания – решать Вам, мы же делаем все возможное, чтобы редкие и ценные книги, еще недавно утраченные и несправедливо забытые, вновь стали доступными для всех читателей.

Серия Книжный Ренессанс
www.samizday.ru/reprint

ЧАСТЬ ПЕРВАЯ
ОСНОВОПОЛОЖЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ.

ГЛАВА ПЕРВАЯ
ОСНОВНЫЕ СВОЙСТВА И ОПРЕДЕЛЕНИЕ МАТЕМАТИЧЕСКОЙ ВЕРОЯТНОСТИ

1. Наиболее обычная и важная теоретическая схема, которой мы пользуемся для познания внешнего мира и для предвидения отдельных фактов, состоит в том, что на основании предшествующего опыта утверждается достоверность наступления события известного класса A, если осуществлен некоторый определенный комплекс условий α, каковы бы ни были прочие обстоятельства. Поскольку в данном конкретном опыте соблюдены условия α, наступление факта A неизбежно.

В дальнейшем нам придется еще рассмотреть в свете теории вероятностей, на чем покоится и насколько обоснована эта уверенность, и мы увидим, что *абсолютной* достоверностью наши предвидения о ходе реальных явлений никогда не могут обладать; но пока нам нет надобности на этом останавливаться. Заметим лишь, что предполагаемый нами закон, связывающий комплекс α с фактом A, только тогда будет практически полезен, если совокупность условий α не слишком сложна и не слишком трудно поддается наблюдению; если, напротив, неуловимое изменение условий опыта влечет за собой противоположный результат, если, даже при самой совершенной технике, время, нужное для исследования того, соблюдены ли условия α или нет, превышает продолжительность опыта, — то такой закон оказывается непригодным для предвидения факта A, который мы называем тогда случайным. Хотя опыты этого рода мало распространены в лабораториях, но в практической жизни почти ни один наш поступок не мог бы быть совершон, если бы мы хотели знать с безусловной достоверностью, что соблюден весь комплекс условий, обеспечивающий наступление нужного нам факта: мы обычно не выскакиваем на улицу из окна, но предпочитаем спуститься по лестнице, хотя и здесь можно упасть и разбиться на-смерть, между тем как прыжок из окна не всегда смертелен. Поступая так или иначе, мы руководствуемся не уверенностью в наступлении определенного результата, а сознательной или инстинктивной оценкой его вероятности. Таким образом, говоря, что некоторый факт A случаен, т.-е. не может быть предсказан с безусловной уверенностью. мы на этом не успокаиваемся, а стремимся ввести вместо слишком усложнившегося комплекса α, из которого A вытекает всегда, такой более простой комплекс условий β, при наличии которого в данном опыте наступление факта A имеет определенную вероятность. При этом, пока речь идет только об отдельном конкретном опыте, особенно важны те случаи, когда, как в вышеуказанном примере, наступлению рассматриваемого события приписывается очень большая или очень малая вероятность, подразумевая под этим, что факты очень вероятные происходят почти всегда, а маловероятные — почти никогда.

Теория вероятностей, дающая правила для вычисления вероятностей одних событий, когда известны вероятности некоторых других, имеет одной из главных своих целей установление таких фактов, вероятности которых очень велики; поэтому в том обстоятельстве, что эти последние факты в действительности почти всегда осуществляются, можно видеть экспериментальное подтверждение объективной правильности общих предпосылок, лежащих в основе теории вероятностей. Вместе с тем и всякая специальная схема теории вероятностей, служащая для описания или объяснения определенной группы явлений, может считаться объективно соответствующей действительности тогда и только тогда, когда факты, теоретически обладающие очень большой вероятностью, на самом деле осуществляются почти всегда.

2. Основное допущение теории вероятностей (постулат существования математической вероятности) состоит в том, что существуют такие комплексы условий β, которые (теоретически по крайней мере) могут быть реализованы неограниченное число раз, при наличности которых в данном опыте наступление факта A имеет определенную вероятность, выражающуюся математическим числом. Таким образом, если факт B также обладает при соответствующих условиях определенной вероятностью, то всегда имеет место одно из трех соотношений:

$$\text{вер. } A = \text{вер. } B; \quad \text{вер. } A > \text{вер. } B; \quad \text{вер. } A < \text{вер. } B.$$

Так же, как и в случае закономерных связей, о которых мы говорили вначале, опыт имеет решающий голос в вопросе о том, возможно ли при осуществлении данного комплекса условий β и полной неопределенности прочих обстоятельств приписать факту A определенную вероятность. Одной из важнейших целей прикладной теории вероятностей, или статистики, именно и является установление a posteriori, при помощи применения соответствующих математических теорем, совместимо ли с наблюденными статистическими данными допущение, что в ряде опытов, где были осуществлены известные одинаковые условия β, факт A имел определенную постоянную вероятность. Таким образом можно, например, считать довольно точно проверенным экспериментально, что при бросании с достаточной высоты симметрично сделанной игральной кости появление на верхней ее грани определенного числа очков имеет во всех опытах ту же самую вероятность. С более специальными статистическими примерами этого рода мы встретимся в дальнейшем, и на них нам пока незачем останавливаться. Но следует заметить, что во многих простых опытах — между прочим в обычных играх в карты, кости, лото и т. д. — мы без всяких статистических наблюдений усматриваем a priori, что некоторые результаты A и B данного опыта следует признать **равновероятными** (равновозможными), т.-е. имеющими равные вероятности. Основанием для этого суждения служит в большинстве случаев симметрия, а именно допущение, что при самом внимательном анализе фактических условий каждого опыта нельзя указать ни одного обстоятельства, которое нужно было бы изменить в постановке опыта, если бы мы взаимно заменили названия фактов A и B. Например, тщательным и добросовестным перемешиванием карт полной колоды мы достигаем — или по крайней мере одинаково перед каждой игрой стремимся достигнуть — полной независимости в расположении карт от их значения или масти, так, чтобы объективно учитываемые условия предстоящей игры не изменились от того, назовем ли мы козырем (т.-е. привилегированной в каком-нибудь отношении мастью) бубновую или какую-нибудь другую масть; благодаря указанной симметрии при вынимании карты из колоды считают одинаково вероятным появление бубновой масти и какой-нибудь другой, например трефовой. Точно так же, если в урне имеется несколько перенумерованных шаров, не различимых на ощупь и тщательно перемешанных, то физические условия опыта вынимания шара симметричны по отношению к любым номерам, и появлению шара с одним номером мы приписываем ту же вероятность, что и появлению шара со всяким другим.

Указанные примеры, по отношению к которым все следствия теории можно считать экспериментально проверенными, являются типичными образцами опытов, при которых соответствующие факты имеют определенную вероятность. Таким образом утверждение, что некоторый факт A (при опытах, где соблюден определенный комплекс условий β) имеет известную вероятность, а именно ту же самую, что и появление данного шара в определенном опыте с урной вышеуказанного образца, означает, что объективные условия для возникновения в рассматриваемых опытах факта A равноценны по существу условиям, осуществляемым в соответствующем схематическом опыте с шарами.

Например, если мы говорим, что при некоторых определенного рода опытах вероятность наступления события A равна вероятности извлечения данного шара из урны, в которой находятся 100 одинаковых шаров, то мы этим желаем выразить, что в каждом отдельном опыте мы так же мало можем рассчитывать на наступление факта A, как на появление указанного шара из рассматриваемой урны; в частности, при многократном повторении опыта нет никаких объективных оснований ожидать, что факт A будет появляться чаще (или реже), чем данный шар при повторении извлечения из нашей урны (полагая, что вынутый шар всегда кладется обратно и тщательно смешивается с остальными).

3. В дальнейшем, опыты с шарами и картами будут служить нам иллюстрациями для пояснения общих правил, которыми нужно руководиться при установлении соотношений между вероятностями одних фактов, когда известны определенные соотношения между вероятностями некоторых других связанных с ними фактов. Правила эти, из которых дедуктивным путем выводятся все положения теории вероятностей, сводятся к трем аксиомам: 1) аксиома сравнения вероятностей, 2) аксиома о несовместимых событиях, 3) аксиома о совмещении событий.

Следует заметить, что первые две аксиомы, говоря о вероятности какого-нибудь факта A, всегда имеют в виду существование некоторого неизменного комплекса условий β опыта; напротив, последняя аксиома (о совмещении) связывает вероятность факта A при одних данных условиях α с вероятностью того же факта при другом соответствующем комплексе условий β. Для того, чтобы лучше уяснить самостоятельное значение последней аксиомы, мы отложим ее формулировку до второй главы, ограничившись пока рассмотрением первых двух аксиом и простых, но важных следствий, которые вытекают из них.

Введем предварительно несколько основных понятий и терминов, которыми нам неоднократно придется пользоваться. Положим, что из колоды карт вынимается одна карта, и интересующее нас событие A заключается в появлении туза безразлично какой масти; в таком случае каждый из фактов a_1, a_2, a_3, a_4 появления туза соответственно бубновой, трефовой, пиковой и червонной масти называется частным случаем события A. Таким образом всякий частный случай a события A характеризуется тем, что наступление a означает также и наступление A. Понятно, что появление A_1 туза черных мастей a_2 или a_3 также является частным случаем появления A туза вообще.

Часто бывает полезно делать различие между частными случаями в узком и широком смысле слова: если мы знаем, например, что в колоде кроме тузов черных мастей есть и красные тузы, так что событие A возможно и без наступления его частного случая A_1, то A_1 называется частным случаем A в узком смысле слова (или, короче, видом события A); напротив, если бы наличность красных тузов нам не была известна, т.-е. если бы мы не были уверены в возможности появления A помимо A_1, мы ограничились бы утверждением, что A_1 есть частный случай A (в широком смысле слова), не исключая, таким образом, допущения, что A_1 вполне тождественно с событием A. Итак, утверждая, что a частный случай в узком смысле слова (или вид) события A, мы выражаем не только то, что A несомненно осуществляется, коль скоро наступило a, но также и то, что осуществление A возможно и без появления a.

Очевидно, что, применяя вышеуказанным образом термин вид, мы должны в соответствии с нашим интуитивным представлением о „вероятности" принять, что всякое событие A всегда более вероятно, чем некоторый его вид (частный случай в узком смысле слова) a, т.-е. *вер.* $A >$ *вер.* a. Естественно поэтому принять следующую аксиому:

Аксиома сравнения вероятностей. Если a есть вид (частный случай в узком смысле слова) события A, то *вер.* $a <$ *вер.* A;

обратно, если между вероятностями фактов a_1 и A существует неравенство *вер.* $a_1 <$ *вер.* A, то оно означает, что *вер.* $a_1 =$ *вер.* a, где a есть некоторый вид события A.

Согласно первой части нашей аксиомы мы принимаем, например, что в некоторых определенных условиях вероятность рождения ребенка вообще больше, чем вероятность рождения мальчика.

Из первой части этой аксиомы вытекают два следствия, которые и непосредственно очевидны:

1. Если факт A достоверен (т.-е. необходимо должен наступить), а факт B только возможен, то наша уверенность в наступлении A больше, чем в наступлении B (*вер.* $A >$ *вер.* B), ибо B является лишь одним из видов осуществления A.

2. Если факт B возможен, а событие C невозможно, то мы больше рассчитываем на появление B, чем на появление C (*вер.* $B >$ *вер.* C), так как событие B может произойти и без C.

Иначе говоря, достоверность есть максимальная вероятность, а невозможность есть минимальная вероятность. Таким образом все достоверные факты имеют одну и ту же максимальную, т.-е. наибольшую вероятность, а невозможные факты имеют минимальную, т.-е. наименьшую вероятность.

Поясним теперь вторую часть высказанной аксиомы. Когда мы говорим, что вероятность смерти в течение года больше, чем вероятность смерти в течение месяца, то последний факт есть частный случай первого. Но что выражает утверждение, что вероятность смерти A в течение года для лиц одной категории (например для стариков 70 лет) больше, чем вероятность смерти a_1 в течение года для лиц другой категории (например для юношей 20 лет)? Очевидно, что здесь не может быть речи о том, что a_1 является частным случаем A; но если мы введем частный случай a события A, заключающийся в смерти старика в течение некоторого более или менее короткого промежутка времени, то, согласно первой части аксиомы, по мере уменьшения рассматриваемого промежутка времени уменьшается и вероятность факта a: в согласии со второй частью аксиомы утверждение, что *вер.* $A >$ *вер.* a_1 означает существование такого срока (примерно три недели), для которого *вер.* $a =$ *вер.* a_1, т.-е. вероятность старику умереть в течение некоторого определенного срока, меньшего, чем год, та же, что вероятность юноше 20 лет умереть в продолжение года.

Разумеется, фактическое указание такого вида a события A, что *вер.* $a =$ *вер.* a_1, может представлять более или менее значительные затруднения, но тем не менее теоретическая возможность этого принципиально необходима для того, чтобы утверждение *вер.* $A >$ *вер.* a_1 имело смысл.

4. Переходим к следующей аксиоме.

Аксиома о несовместимых событиях. Если известно, что события A и A_1 несовместимы между собой, и, с другой стороны, события B и B_1 также между собой несовместимы, при чем *вер.* $A =$ *вер.* B и *вер.* $A_1 =$ *вер.* B_1, то вероятность факта C, заключающегося в наступлении события A или события A_1, равна вероятности факта C_1, заключающегося в наступлении B или B_1, т.-е. *вер.* (A или A_1) $=$ *вер.* (B или B_1).

Понятие несовместимости двух фактов A и A_1 вряд ли нуждается в особом пояснении: оно означает, что наступление в данном опыте одного из них невозможно в случае наступления другого.

Высказанную аксиому можно формулировать еще иначе: вероятность наступления какого-нибудь одного из двух несовместимых событий A или A_1 есть величина, которая вполне определяется вероятностями каждого из них в отдельности, т.-е. представляет собой определенную функцию *вер.* A и *вер.* A_1, вовсе не зависящую от природы самих фактов A и A_1; иными словами,

$$\text{вер. } (A \text{ или } A_1) = f\,(\text{вер. } A,\ \text{вер. } A_1). \qquad (1)$$

Например, если по тем или иным основаниям мы знаем, что для лица данной категории, вступающего в брак, вероятность овдоветь в течение трех лет равна вероятности, что при вынимании шара из данной урны появится белый шар, а вероятность овдоветь после трех лет равна вероятности появления черного шара, то вероятность вообще овдоветь для лица указанной группы равна вероятности появления из урны шара белого либо черного цвета.

Высказанную аксиому нетрудно распространить на какое угодно число несовместимых событий. А именно:

Следствие 1. Если $A_1, A_2, A_3, \ldots, A_n$ между собой несовместимы и факты $B_1, B_2, B_3, \ldots, B_n$ также между собой несовместимы, при чем *вер.* $A_1 =$ *вер.* B_1, *вер.* $A_2 =$ *вер.* $B_2, \ldots,$ *вер.* $A_n =$ *вер.* B_n, то *вер.* (A_1 или $A_2 \ldots$ или A_n) $=$ *вер.* (B_1 или $B_2 \ldots$ или B_n).

В самом деле, наше утверждение для $n = 2$ соответствует выше формулированной аксиоме; поэтому для доказательства его в общем виде достаточно показать, что если утверждение справедливо для некоторого значения n, то оно верно также и для числа $n+1$. Итак, дано, что A_{n+1} несовместимо ни с одним из событий $A_1, A_2, \ldots, A_n$; следовательно A_{n+1} несовместимо также и с наступлением факта α, состоящего в появлении A_1 или $A_2 \ldots$ или A_n, точно так же B_{n+1} несовместимо с фактом β, состоящим в наступлении B_1 или $B_2 \ldots$ или B_n. Но мы полагаем уже известным, что *вер.* (A_1 или $A_2 \ldots$ или A_n) $=$ *вер.* (B_1 или $B_2 \ldots$ или B_n), т.-е. *вер.* $\alpha =$ *вер.* β, и кроме того, по условию, *вер.* $A_{n+1} =$ *вер.* B_{n+1}. Следовательно, в силу нашей аксиомы

$$\text{вер. } (\alpha \text{ или } A_{n+1}) = \text{вер. } (\beta \text{ или } B_{n+1}),$$

т.-е. *вер.* (A_1 или $A_2 \ldots$ или A_{n+1}) $=$ *вер.* (B_1 или $B_2 \ldots$ B_{n+1}), что и требовалось доказать.

С другой стороны, сопоставляя обе аксиомы, получаем

Следствие 2. Если A и A_1 несовместимы между собой, а B и B_1 несовместимы между собой, при чем *вер.* $A =$ *вер.* B и *вер.* $A_1 >$ *вер.* B_1, то *вер.* (A или A_1) $>$ *вер.* (B или B_1).

В самом деле, согласно аксиоме сравнения вероятностей, можно указать такой вид a_1 факта A_1, что *вер.* $a_1 =$ *вер.* B_1. В таком случае наступление A или a_1 будет частным случаем в узком смысле слова факта (A или A_1), и следовательно *вер.* (A или A_1) $>$ $>$ *вер.* (A или a_1) $=$ *вер.* (B или B_1).

Следствие второе, очевидно, равнозначно утверждению, что функция f (*вер.* A, *вер.* A_1), которая стоит в формуле (1), возрастает при возрастании одной из двух вероятностей, от которых она зависит, если другая из них остается неизменной. Тем более, конечно, функция f возрастает, если обе вероятности увеличиваются. Таким образом получаем:

Следствие 3. Если (при сохранении прежних условий несовместимости) *вер.* $A >$ *вер.* B и *вер.* $A_1 >$ *вер.* B_1, то *вер.* (A или A_1) $>$ *вер.* (B или B_1).

Применяя затем тот же способ математической индукции, которым мы воспользовались для доказательства следствия (1), нетрудно распространить следствия (2) и (3) на произвольное число фактов, и получаем

Следствие 4. Если события A_1, $A_2, \ldots, A_n$ между собой несовместимы, а, с другой стороны, события B_1, $B_2, \ldots, B_n$ несовместимы между собой, при чем *вер.* $A_1 \geqslant$ *вер.* $B_1, \ldots,$ *вер.* $A_{n-1} \geqslant$ *вер.* B_{n-1}, *вер.* $A_n >$ *вер.* B_n, то *вер.* (A_1 или $A_2 \ldots$ или A_n) $>$ *вер.* (B_1 или $B_2 \ldots$ или B_n).

Предположим теперь, в частности, что рассматриваемые нами несовместимые между собой случаи A_1, $A_2, \ldots, A_n$ имеют равные вероятности и, кроме того, единственно возможны, т.-е. один из фактов A_1, $A_2, \ldots, A_n$ непременно должен наступить; предположим также, что и события B_1, $B_2, \ldots, B_n$ несовместимы между собой, равновероятны и единственно возможны. Нетрудно понять, что при таких условиях вероятность каждого из событий A должна быть равна вероятности события B; действительно, если бы этого не было, т.-е. если бы мы допустили, например, что *вер.* $A_1 >$ *вер.* B_1, то мы имели бы также *вер.* $A_2 >$ *вер.* $B_2, \ldots,$ *вер.* $A_n >$ *вер.* B_n, и благодаря следствию (4) пришлось бы признать, что *вер.* (A_1 или $A_2 \ldots$ или A_n) $>$ *вер.* (B_1 или $B_2 \ldots$ или B_n), что невозможно, так как наступление A_1 или $A_2 \ldots$ или A_n достоверно, равно как и наступление B_1 или $B_2 \ldots$ или B_n. Аналогичным образом убеждаемся в недопустимости предположения, что *вер.* $A_1 <$ *вер.* B_1; а потому, какова бы ни была физическая природа фактов A и фактов B, мы вынуждены им приписать одну и ту же вероятность.

Принимая во внимание следствие (1), мы выводим отсюда

Следствие 5. Если событие X происходит в некотором опыте, допускающем n равновероятных, несовместимых и единственно возможных исходов A_1, $A_2, \ldots, A_n$, тогда и только тогда, когда наступает какое-нибудь из m ($m < n$) определенных исходов A_1, $A_2, \ldots, A_m$ (иначе говоря, если факту X благоприятствуют только случаи A_1, $A_2, \ldots, A_m$); если, с другой стороны, и событию Y благоприятствуют m определенных случаев $B_1, \ldots, B_m$ из каких-нибудь n равновероятных, несовместимых и единственно возможных случаев B_1, $B_2, \ldots, B_n$, — то событие X и событие Y имеют равные вероятности.

Последнему утверждению, очевидно, можно дать и такую формулировку:

Если событию X благоприятствуют m случаев из общего числа всех n единственно возможных несовместимых и равновероятных случаев, то вероятность события X зависит только от чисел m и n (а не от природы рассматриваемого опыта), т.-е.

$$\text{вер. } X = F(m, n), \tag{2}$$

где $F(m, n)$ есть некоторая определенная функция, которая должна быть установлена раз навсегда.

Какова же эта функция $F(m, n)$? Можем ли мы ее выбрать по своему усмотрению, или же, напротив, наш выбор в той или иной мере предопределен принятыми нами аксиомами? Ответ на этот вопрос дает следующая основная теорема.

5. *Теорема.* Если событию X благоприятствуют m случаев из общего числа n несовместимых, равновероятных и единственно возможных случаев $A_1, A_2, \ldots, A_n$, а событию Y благоприятствуют m_1 случаев из n_1 несовместимых, равновероятных и единственно возможных случаев $B_1, B_2, \ldots, B_{n_1}$, то *вер.* $X =$ *вер.* Y, если

$$\frac{m}{n} = \frac{m_1}{n_1};$$

напротив, *вер.* $X >$ *вер.* Y, если

$$\frac{m}{n} > \frac{m_1}{n_1}.$$

В самом деле, предположим сначала, что $\frac{m}{n} = \frac{m_1}{n_1}$; общее значение этих обеих дробей можно представить в виде несократимой дроби $\frac{M}{N}$. Иными словами, обозначая соответственно через k и k_1 множители, на которые нужно сократить дроби $\frac{m}{n}$ и $\frac{m_1}{n_1}$, чтобы привести их к несократимому виду $\frac{M}{N}$, можем написать:

$$m = kM, \quad n = kN, \quad m_1 = k_1M, \quad n_1 = k_1N.$$