Теория вероятности

Дима Трушин

Семинар 1

Теория вероятности

Случайность Во-первых, хорошо бы понять, а что такое случайный объект и как про него правильно думать. Любой случайный объект – это такой черный ящик, который мы можем спросить, в каком состоянии он сейчас находится. И этот черный ящик нам услужливо отвечает, но каждый раз на этот вопрос может давать разные ответы. В качестве примера давайте рассмотрим кубик с шестью гранями. Что мы делаем: мы берем и бросаем кубик, а потом смотрим какой стороной он выпал. Давайте поясню, как на этот процесс смотреть как на черный ящик. Бросание кубика можно рассматривать как процесс задания вопроса черному ящику, а когда на нем выпадает грань, мы на это смотрим как на ответ черного ящика.

Для изучения любого такого черного ящика для начала надо описать все возможные состояния, в котором он может находиться. Обычно множество таких состояний обозначают через Ω, а каждое состояние $\omega \in \Omega$ называют элементарным исходом. В примере с кубиком в качестве Ω можно взять множество из шести граней, нумеруя каждую соответствующим числом.

Случайные события Давайте начнем с примера. В случае кубика, мы можем задать следующее событие: «выпала четная грань» и «выпало число 7». Теперь мы задаем вопрос кубику и смотрим, наступило требуемое событие или нет. Таким образом можно все элементарные исходы поделить на две части: те которые удовлетворяют событию и те, которые не удовлетворяют. Если $\Omega = \{1, 2, 3, 4, 5, 6\}$, то «выпала четная грань» соответствует подмножеству $\{2, 4, 6\}$, а «выпало число 7» соответствует пустому подмножеству $\varnothing$. То есть как мы видим из этого примера, случайные события должны соответствовать подмножеству множества элементарных исходов. Потому в общем случае это кладется в качестве определения.

Пусть Ω – множество элементарных исходов. Тогда событием (или случайным событием) называется любое подмножество $A \subseteq \Omega$. Если у нас есть два события $A, B \subseteq \Omega$, то для них определены все теоретико множественные операции. $A \cap B$ означает, что оба события произойдут одновременно, $A \cup B$ означает, что произойдет хотя бы одно событие, $A \Delta B = (A \setminus B) \cup (B \setminus A)$ означает произойдет одно из событий, но не оба сразу. Пустое множество $\varnothing$ означает невозможное событие (ничего не может произойти, так как ни один элементарный исход не подходит), а все множество Ω означает, что произойти может любое событие. Отрицанием события A называется событие $\bar{A} = \Omega \setminus A$, это значит произошло не A, то есть что угодно, кроме события A.

От математика не математикам Тут я хочу вставить небольшой, но важный разговор. Так часто бывает, что математика бывает сложно устроена, не потому что решаемая задача сложная, а потому что сама математика внутри устроена сложнее, чем задача, для которой она изобреталась. И вот эти самые трудности внутреннего устройства самой математики очень часто сильно переусложняют рассказываемый текст. С одной стороны, это делает строгие математические книжки просто нечитаемыми для обычных смертных, какими бы хорошими эти книжки ни были, с другой стороны, математики не могут просто игнорировать подобные явления, иначе математика внезапно станет противоречивой.[1] Однако, мы же не собираемся создавать математику, мы будем ей пользоваться. Потому я намерено буду обманывать вас относительно технических вещей связанных с теорией множеств, которые возникают в рамках теории вероятности. Говорю я это потому, что вы можете пойти и поглядеть в различную литературу определение вероятностного пространства и увидите, что после разговора о пространстве элементарных исходов Ω люди начинают строить какие-то

[1] Георг Кантор, кстати, сошел с ума именно потому что придумал крутой способ мерить количество элементов в бесконечных множествах, но не сумел сделать этот способ непротиворечивым. Его идею можно легко рассказать другим, но сделать строгой и формально рабочей – задача не самая простая.

непонятные σ-алгебры событий, потом измеримые функции, интегралы Лебега и прочую хренатень. Так получилось, что без всех этих серьезных наворотов построить теорию вероятности невозможно, иначе она будет противоречивой. Грубо говоря, основная проблема следующая: когда нам дали бесконечное множество элементарных исходов Ω, то нельзя в качестве события взять произвольное подмножество, потому что, когда мы будем вводить понятие вероятности, у нас внезапно вылезет критический баг в теории. Потому надо сначала проредить правильным образом все множества. Потом надо будет прореживать функции, потом строить специальную теорию интегрирования этих функций и т.д. Но самое забавное заключается в том, что если вы попытаетесь найти те самые плохие множества, от которых мы пытались избавиться, у вас это ни за что не получится. Оказывается, что они существуют только тогда, когда вы верите в аксиому выбора[2], а если вы в нее не верите, то нельзя ни доказать, что таких плохих множеств нет, ни привести примера. Проще говоря, в нормальное жизни не математика они попросту не встречаются, и их вообще говоря никто никогда не видел и не увидит. Потому я при рассказе о теории вероятности вырежу всю эту трудно перевариваемую лабуду и буду рассказывать почти честно, но зато во много раз (надеюсь) понятнее.

Вероятность Пусть у нас есть какое-то пространство элементарных исходов Ω. Неформально, вероятность – это правило, которое каждому событию (то есть подмножеству) ставит в соответствие неотрицательное число (вероятность события) и ставит это некоторым хорошим способом. Давайте введем строгое определение.

Пусть Ω – пространство элементарных исходов (то есть произвольное множество), через 2^Ω обозначим множество всех подмножеств Ω.[3] Тогда вероятность на Ω это отображение $P\colon 2^\Omega \to \mathbb{R}_+$ (из множества всех подмножеств Ω в множество неотрицательных вещественных чисел), такое что выполнены следующие правила

1. $P(\varnothing) = 0$, $P(\Omega) = 1$.
2. $P(A \cup B) = P(A) + P(B)$ при условии, что $A \cap B = \varnothing$.
3. Пусть есть последовательность непересекающихся событий $A_i \subseteq \Omega$, $A_i \cap A_j = \varnothing$, тогда

$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i)$$

В этом случае значение $P(A)$ называется вероятностью события $A \subseteq \Omega$. Множество Ω вместе с вероятностью P называется вероятностным пространством. Именно эта пара полностью описывает любой случайный объект. Отображение вероятности P еще называют вероятностной мерой.

Вероятность любого события обязательно находится между нулем и единицей. Условие $P(\varnothing) = 0$ означает, что вероятность невозможного события равна нулю, а условие $P(\Omega) = 1$ означает, что вероятность, что что-нибудь да произойдет равна единице.

Третье условие – это лишь более сильная версия второго. Потому второе условие можно было бы откинуть, но оно очень полезно для понимания, что происходит. Второе условие означает, что если у вас есть два события, которые не могут случиться одновременно, то вероятность того, что случится одно из них равна сумме вероятностей. Еще полезно понимать следующие свойства

1. Если есть два события $A \subseteq B \subseteq \Omega$, то есть событие A всегда влечет событие B, то

   $$P(B \setminus A) = P(B) - P(A)$$

   Это значит, что если A влечет B, то вероятность, что произойдет B, но A при этом не случится, равна разности вероятностей $P(B) - P(A)$.

2. Если есть два произвольных события $A, B \subseteq \Omega$, то

   $$P(B \setminus A) = P(B) - P(A \cap B)$$

   Это значит, что если есть два произвольных события A и B, то вероятность того, что произойдет B, но не A, равна разности вероятности того, что произойдет B вероятности одновременного выпадения A и B.

[2] Это одна из очень популярных и не очевидных аксиом теории множеств. Проблема в том, что в абстрактной математике она очень полезная.

[3] То есть запись $A \in 2^\Omega$ означает, что $A \subseteq \Omega$.

3. Для любого события $A \subseteq \Omega$, верно
$$P(\bar{A}) = 1 - P(A)$$
То есть вероятность отрицания события A есть дополнение вероятности A до единицы.

4. Если есть два события $A \subseteq B \subseteq \Omega$, то есть A влечет B, то
$$P(A) \leqslant P(B)$$
То есть у не меньшего события, не меньше вероятность. Может так оказаться, что у большего события вероятность такая же, но это будет означать, что вероятность их разности нулевая.

5. Если есть два события $A, B \subseteq \Omega$, то
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
Если вы знаете, что события A и B могут произойти одновременно, то вероятность того, что случится одно из них считается по формуле выше, надо сложить вероятности событий и вычесть вероятность их одновременного выпадения. Это называется формула включения исключения.

Опять пример с кубиком Давайте переключимся на наш пример с кубиком. В этом случае $\Omega = \{1, 2, 3, 4, 5, 6\}$. Откуда брать меру? Ее надо задать. Из каких соображений? Здесь надо исходить из физических условий задачи. Все зависит от того, какими свойствами обладает наш кубик. Если мы считаем, что кубик хорошо сбалансирован, то это означает, что вероятность получить любую грань одинаковая. То есть $P(i) = p \in \mathbb{R}_+$, где $1 \leqslant i \leqslant 6$. С другой стороны
$$1 = P(\Omega) = P(\{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{5\} \cup \{6\}) = P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\}) = 6p$$
Значит $p = \frac{1}{6}$. Но если мы сомневаемся, что кубик был правильным, то вообще говоря, чтобы задать вероятность на всех событиях мы должны задать вероятности каждого элементарного исхода произвольно $P(i) = p_k$, где $1 \leqslant i \leqslant 6$, при этом $0 \leqslant p_i$ и $\sum_{i=1}^{6} p_i = 1$. Теоретически подходят любые наборы чисел p_i удовлетворяющие последним двум условиям. Все они будут отвечать кубикам с разной степенью балансировки граней.

Как можно было бы экспериментально измерить подобные числа p_i? Можно было бы сделать так: берем кубик и кидаем его n раз. Считаем сколько раз выпало число 1 и получаем число $k_n(1)$, для числа 2 получаем $k_n(2)$ и т.д. Важно, что $\sum_{i=1}^{6} k_n(i) = n$. Теперь посмотрим на частоту выпадения грани i после n экспериментов и получим $\frac{k_n(i)}{n}$. Если бы мы могли провести эксперимент бесконечно долго, то мы могли бы найти предел
$$p_i = \lim_{n \to \infty} \frac{k_n(i)}{n}$$
Это и были бы те самые вероятности, которые надо было бы приписать к данному конкретному кубику.

Пример с монетками Пусть мы бросаем монетку n раз. В случае выпадения орла мы будем писать 0, а в случае выпадения решки – 1. Тогда в качестве результата нашего эксперимента (или ответа черного ящика) у нас будут последовательности из 0 и 1 длины n. Всего таких последовательностей будет 2^n. Таким образом $\Omega = \{(a_1, \ldots, a_n) \mid a_i \in \{0, 1\}\}$. Вероятность зададим так: $P(a_1, \ldots, a_n) = 1/2^n$. То есть все элементарные исходы равновероятные.

Конечные множества элементарных исходов Пусть у нас некоторый случайный объект описывается множеством исходов Ω и пусть это множество конечно, то есть у нас возможно всего лишь конечное число возможных состояний у нашего случайного объекта. Пусть $\Omega = \{\omega_1, \ldots, \omega_n\}$. Оказывается, что чтобы задать вероятность на таком пространстве, то достаточно каждой точке ω_i приписать ее вероятность, а именно положить $P(\omega_i) = p_i$, где числа p_i удовлетворяют следующим двум условиям
$$0 \leqslant p_i \text{ и } \sum_{i=1}^{n} p_i = 1$$
Тогда вероятность $P\colon 2^\Omega \to \mathbb{R}_+$ задается по правилу, для любого $A \subseteq \Omega$ положим
$$P(A) = \sum_{\omega_i \in A} p_i$$

То есть, чтобы посчитать вероятность события A, надо сложить вероятности всех элементарных исходов этого события. Это самый простой способ задавать вероятности. На конечных множествах других способов задавать вероятности не существует. Потому, когда описывают случайный объект с конечным числом состояний, всегда указывают вероятности всех элементарных исходов. Еще в подобных случаях принято считать, что $p_i > 0$, так как если у элементарного исхода вероятность ноль, то его можно выкинуть из Ω ибо это бы означало, что этот исход просто не может произойти.

Примеры «непрерывных» мер Пусть у нас $\Omega = [0,1]$ – отрезок. Тогда для любого интервала $[a,b] \subseteq [0,1]$ положим $P([a,b]) = b - a$, то есть вероятность равна длине отрезка. Для любого подмножества $X \subseteq [0,1]$ мы можем задать меру как

$$P(X) = \int\limits_X 1\,dx$$

Если $X = [a,b]$, то мы получим введенное выше правило – длину отрезка. Но это правило применимо для кучи других множеств. Можно взять точку, или объединение отрезков, или полуинтервалы. На самом деле можно брать любое подмножество на отрезке, но тогда придется строить правильную версию интеграла. Подобная вероятность «равномерно распределена» по отрезку $[0,1]$. Еще обратите внимание, что вероятность попасть в точку равна нулю, то есть $P(\{a\}) = 0$ для любой $a \in [0,1]$. Тем не менее, сама вероятностная мера P при этом не нулевая. Этот пример поясняет, почему не достаточно меру задавать только на элементарных исходах.

Условная вероятность Пусть у нас есть вероятностное пространство (Ω, P), описывающее какой-либо случайный объект и пусть у нас есть фиксированное событие $B \subseteq \Omega$. Теперь предположим, что нас интересуют не все события случайного объекта, а только те, что случаются внутри события B. Про это можно думать так, мы игнорируем любой ответ нашего черного ящика до тех пор, пока не попался ответ из события B. То есть у такого нового черного ящика пространство всех состояний – это B. Вопрос, а как же там устроена вероятность? Эта вероятность называется условной вероятностью и считается так: пусть $A \subseteq \Omega$ – любое событие на Ω, тогда

$$P(A \mid B) = P_B(A) = \frac{P(A \cap B)}{P(B)}$$

Последнее выражение называется условной вероятностью A при условии B. Про это надо думать так. Так как мы теперь интересуемся только теми исходами, что лежат внутри B, то когда мы проверяем событие A у нас обязательно случается событие $A \cap B$ (то есть одновременно A и B), но $P(A \cap B)$ – это вероятность этого события в исходном эксперименте. Если мы забываем все события вне B, то вероятность B должна стать единицей. Потому мы дополнительно нормируем на вероятность B. Как видно из формулы, подобную конструкцию можно провернуть только если $P(B) \neq 0$.

Еще про это можно думать так: $P(\Omega) = 1$ – это как бы доля всех возможных элементарных исходов в Ω и разумно, что все исходы дают долю 1. Тогда $P(B)$ – это доля тех событий, что лежат в B, а $P(A \cap B)$ – это доля событий одновременно в A и B. Тогда $P_B(A)$ показывает долю события $A \cap B$ в событии B.

Еще обратите внимание, что P_B – это вероятностная мера на исходном Ω, а не только на B. Просто если событие A не может наступить одновременно с B (это значит $A \cap B = \varnothing$ или что то же самое $A \subseteq \bar{B}$), то $P_B(A) = 0$. Вообще, $P_B(A) = 0$ тогда и только тогда, когда $P(A \cap B) = 0$, то есть если вероятность наступления обоих событий одновременно равна нулю.

Еще эту формулу полезно переписать так

$$P(A \cap B) = P(B)P(A \mid B)$$

То есть, чтобы посчитать вероятность одновременного наступления событий A и B надо умножить вероятность наступления события B на условную вероятность наступления события A при условии наступления B. Причем в этом виде нам не надо требовать, чтобы $P(B) \neq 0$. Она верна даже в случае $P(B) = 0$, в этом случае обе части равенства нулевые. В таком виде эта формула полезна в задачах.

Независимые события Пусть (Ω, P) – некоторое вероятностное пространство, описывающее какой-то случайный объект. Пусть $A, B \subseteq \Omega$ – два события. Они называются независимыми, если $P(A \cap B) = P(A)P(B)$. Давайте разберемся с этим формальным определением. Во-первых, если $P(A) = 0$, то A независимо с любым событием B и наоборот, если $P(B) = 0$, то оно независимо с любым событием A. Это важный частный случай, который надо иметь в виду. Помимо это остается случай, когда обе вероятности не ноль. В этом случае мы можем переписать условие независимости в виде

$$P(A) = \frac{P(A \cap B)}{P(B)} = P(A \mid B)$$

Это равенство можно прочитать в несимметричной форме. Событие A не зависит от события B, если его условная вероятность $P(A \mid B)$ не зависит от B и равна обычной вероятности $P(A)$. Но заметим, что фраза «A не зависит от B» – это лишь способ прочитать формулу, реально это всегда означает симметричное условие, что оба события независимы, то есть равенство $P(A \cap B) = P(A)P(B)$.

Давайте приведем пример зависимых и независимых событий. Возьмем правильный игральный кубик, то есть $\Omega = \{1, 2, 3, 4, 5, 6\}$ и $P(i) = 1/6$. В этом случае рассмотрим события $A = \{1, 2\}$ («выпадет один или два») и $B = \{1, 3\}$ («выпадет один или три»). Тогда эти события зависимы

$$P(A \cap B) = P(\{1, 2\} \cap \{1, 3\}) = P(\{1\}) = \frac{1}{6}$$
$$P(A)P(B) = P(\{1, 2\})P(\{1, 3\}) = \left(\frac{1}{6} + \frac{1}{6}\right)\left(\frac{1}{6} + \frac{1}{6}\right) = \frac{1}{9}$$

Еще полезно понимать, если у вас есть два не пересекающихся события A и B, имеющие ненулевые вероятности ($P(A) \neq 0$ и $P(B) \neq 0$), то они всегда зависимы, действительно $P(A \cap B) = P(\varnothing) = 0$, но $P(A)P(B) \neq 0$.

Теперь приведем пример независимых событий. Пусть $A = \{1, 2, 3\}$ и $B = \{1, 4\}$, тогда

$$P(A \cap B) = P(\{1, 2, 3\} \cap \{1, 4\}) = P(\{1\}) = \frac{1}{6}$$
$$P(A)P(B) = P(\{1, 2, 3\})P(\{1, 4\}) = \left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right)\left(\frac{1}{6} + \frac{1}{6}\right) = \frac{1}{6}$$

Вообще говоря, это очень не очевидный вопрос, как, глядя в глаза двум событиям, выпытать у них независимы они или нет. Единственный способ – подставить в формулу и проверить равенство. Ничего лучше не существует. Но зато когда вы знаете, что какие-то события независимы – это обычно огромный повод для радости.

Обычная ситуация, когда получаются независимые события – когда вы проводите два разных случайных эксперимента. А именно, пусть у вас есть два черных ящика X и Y и каждый дает свой ответ. Теперь вы строите новый черный ящик Z как такой, что выдает пару ответов от двух прежних ящиков. Оказывается, что события сформулированные в терминах X и события сформулированные в терминах Y будут независимы для ящика Z. Я не буду углубляться в математическую формальность этой конструкции и надеюсь, что смог хоть какую-то интуицию по поводу независимости тут навести.

Независимость в совокупности Пусть (Ω, P) – наш вероятностный черный ящик и в нем заданы события $A_1, \ldots, A_n \subseteq \Omega$. Тогда они называются независимыми в совокупности, если

$$P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$$

для любого набора подмножеств $A_{i_1}, \ldots, A_{i_k}$, где $1 \leqslant i_1 < \ldots < i_k \leqslant n$ и $1 \leqslant k \leqslant n$. Я хочу обратить внимание, что если события независимы в совокупности, то любое подмножество событий $A_{i_1}, \ldots, A_{i_k}$ тоже будет независимым в совокупности. В частности все события будут попарно независимыми. Однако, из попарной независимости не следует независимость в совокупности. Давайте рассмотрим следующий пример.

Рассмотрим независимое бросание двух правильных монеток. Тогда $\Omega = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$, где 0 означает выпадение орла, а 1 – выпадение решки. Вероятность каждого из этих событий полагается $1/4$. Тогда рассмотрим три события:

- $A_1 = \{(1, 0), (1, 1)\}$, что значит «у первой монеты выпала решка».
- $A_2 = \{(0, 1), (1, 1)\}$, что значит «у второй монеты выпала решка».
- $A_0 = \{(0, 1), (1, 0)\}$, что значит «монеты выпали по-разному».

Заметим, что все эти события одновременно не совместны, то есть $A_0 \cap A_1 \cap A_2 = \varnothing$. Потому равенство $P(A_0 \cap A_1 \cap A_2) = P(A_0)P(A_1)P(A_2)$ не выполняется. Но, прямые вычисления, показывают, что все эти события попарно независимы. Проверим это в случае A_1 и A_2. Видим, что $P(A_1) = 1/2$, $P(A_2) = 1/2$ и $P(A_1 \cap A_2) = P(\{(1,1)\}) = 1/4 = P(A_1)P(A_2)$. Аналогично делаются остальные два случая.

Думать про этот пример надо так. Для каждой пары события наступают независимо. То есть вероятность того случится A_1 или нет никак не зависит от того произошло A_0 или нет. Однако, если мы хотим сформулировать более сложное событие, например, $A_0 \cap A_1$, что означает «монеты показали разный результат и на первой монете выпала решка». В этом случае мы видим, что результат выпадения второй монеты предопределен, он обязан быть орлом, тут нет никакой независимости.

Еще раз про монетки Пусть мы n раз бросаем монетку. Тогда элементарными исходами у нас являлась последовательность из нулей и единиц длины n. Мы полагали, что $P(a_1, \ldots, a_n) = 1/2^n$. Пользуясь понятием независимости, можно дать естественное объяснение тому, что это означает для нашего эксперимента. Действительно, событие $(a_1, \ldots, a_n)$ означает: « первая монета показала a_1, вторая монета показала a_2,…, n-я монета показала a_n». То есть это пересечение событий вида «i-я монета показала a_i». Тогда видно, что мы положили нашу вероятность по правилу

$$P(a_1, \ldots, a_n) = P(\text{«1-я монета показала } a_1\text{»}) \ldots P(\text{«}n\text{-я монета показала } a_n\text{»}) = \frac{1}{2} \cdots \frac{1}{2} = \frac{1}{2^n}$$

Таким образом предположение о равномерности распределения вероятности по всем бросаниям монеток означает, что выпадения орлов и решек у разных монет происходит независимо друг от друга. В этом случае говорят, что мы бросаем n монеток независимо друг от друга.

Формулы полной вероятности Пусть у вас есть вероятностное пространство (Ω, P), описывающее какой-то случайный объект. Пусть у вас задан набор событий $A_1, \ldots, A_n \subseteq \Omega$ такой, что $A_i \cap A_j = \varnothing$ для любых $i \neq j$ и $\Omega = A_1 \cup \ldots \cup A_n$. То есть события не пересекаются и покрывают все пространство элементарных исходов. Тогда такой набор событий называется разбиением Ω. Неформально про разбиение надо думать так: вы перечисляете взаимоисключающие состояния в которых может быть система. Например, в случае кубика вы можете выбрать разбиение из двух событий: «выпало четное» и «выпало нечетное». Вы знаете, что всегда произойдет одно из двух и ничего другого быть не может. Еще можно разбить на события $A_1 = \{1, 2\}$, $A_2 = \{3, 4\}$ и $A_3 = \{5, 6\}$. Эти события сложно описать словесно, точнее можно, но будет звучать глупо. Но тем не менее, это тоже вполне себе разбиение.

Пусть теперь у нас есть событие A, тогда выполняется следующая формула

$$P(A) = P(A \mid A_1)P(A_1) + \ldots + P(A \mid A_n)P(A_n) = \sum_{i=1}^{n} P(A \mid A_i)P(A_i)$$

Эта формула называется формулой полной вероятности. Она бывает полезна, когда вы знаете условные вероятности $P(A \mid A_i)$ и знаете вероятности $P(A_i)$, с которыми ваше пространство бьется на эти события.

Формулы Байеса Предыдущая формула обычно идет в связке с еще одной популярной формулой. Она связана с тем, как пересчитать $P(A \mid B)$ через $P(B \mid A)$ и наоборот. Эти формулы называются формулами Байеса. Пусть у нас есть два события A и B в Ω, тогда

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{P(A \mid B)P(B)}{P(A)}$$

Итоговое равенство

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

и называется формулой Байеса.