Дискретная вероятность

1 Понятие дискретной вероятности

1.1. Начнем, как всегда, с самых простых, базовых понятий теории вероятности.

1.1.1. Под случайным экспериментом будем понимать математическую модель некоторого реального эксперимента, результат которого невозможно точно предсказать. Простейшим и наиболее хрестоматийным примером случайного эксперимента является подбрасывание идеальной монетки. В результате любого такого эксперимента у нас обязательно выпадает или орел, или решка (на ребро идеальная монетка упасть не может). Заранее предсказать, что именно выпадет, мы не можем. Другим простейшим примером случайного эксперимента является подбрасывание игральной кости, результатом которого можно считать выпадение одного из шести чисел, нанесенных на грани куба.

1.1.2. Любой результат ω случайного эксперимента называется *элементарным событием* или *исходом.* В примере с подбрасыванием монетки элементарным исходом является, например, выпадение орла. В примере с подбрасыванием кубика элементарный исход — это выпадение в результате этого случайного эксперимента одного из шести чисел. Множество всех возможных исходов обозначим через Ω. Далее мы всегда будем считать, что множество Ω конечно или счетно.

Случайным событием или просто *событием* A называется любое подмножество множества Ω. Так, в примере с подбрасыванием игральной кости событиями являются, например, выпадение четного числа или выпадение числа, меньшего тройки. Говорят, что в результате случайного эксперимента *произошло событие* A, если элементарный исход эксперимента является элементом множества A.

Очень важно понимать разницу между элементарным исходом ω и случайным событием A. У одного и того же множества Ω имеется много (а точнее, 2^n, $n = |\Omega|$) различных подмножеств, так что один и тот же элемент $\omega \in \Omega$ может принадлежать разным подмножествам A. На языке теории вероятности это означает, что в результате случайного эксперимента у нас может произойти сразу несколько случайных событий. *Элементарный же исход любого случайного эксперимента может быть только один.*

Так, в примере с игральной костью элементарный исход ω — это выпадение конкретного числа, например, двойки. Этот элементарный исход, однако, отвечает появлению сразу двух описанных выше событий — события A_1, описывающего выпадение четного числа, и события A_2, соответствующего выпадению числа, меньшего тройки.

1.1.3. Предположим теперь, что у нас для заданного случайного эксперимента имеется некоторый набор $\mathcal{A}_0$ событий. С помощью теоретико-множественных операций объединения, пересечения и дополнения можно из этого набора $\mathcal{A}_0$ построить некоторую новую, более полную систему множеств, также являющихся событиями. Присоединяя к этой системе так называемое невозможное ($A = \varnothing$) и достоверное ($A = \Omega$) события, мы получаем систему множеств $\mathcal{A}$, называемую *алгеброй событий,* то есть такую систему подмножеств множества исходов Ω, что, во-первых, само $\Omega \in \mathcal{A}$, и во-вторых, для любой пары $A, B \in \mathcal{A}$ их объединение, пересечение и дополнение также принадлежат $\mathcal{A}$.

Как правило, в качестве $\mathcal{A}_0$ выбирают некоторый набор подмножеств, образующий разбиение множества Ω. В этом случае $\mathcal{A}$ называется алгеброй, порожденной данным разбиением $\mathcal{A}_0$.

В случае конечного множества Ω в качестве $\mathcal{A}$ чаще всего выбирается множество 2^Ω всех подмножеств данного множества Ω.

1.1.4. Припишем теперь любому элементарному событию $\omega \in \Omega$ некоторое вещественное число $\Pr(\omega)$ из диапазона $[0, 1]$. Это отображение $\Pr : \Omega \to \mathbb{R}_+$ называется *вероятностью,* если для него выполняется следующее условие нормировки:

$$\sum_{\omega\in\Omega} \Pr(\omega) = 1. \tag{1}$$

Данное требование достаточно очевидно — оно означает, что в результате случайного эксперимента какой-то из элементарных исходов обязательно произошел.

Рассмотрим, к примеру, случайный эксперимент, связанный с подбрасыванием монетки. Множество Ω состоит в этом случае из двух элементов:

$$\Omega = \{\omega_1 = \{\text{выпадение орла}\},\ \omega_2 = \{\text{выпадение решки}\}\}.$$

Сопоставим элементарному исходу, отвечающему выпадению орла, вещественное число $p \in [0, 1]$, а элементарному исходу, отвечающему выпадению решки, вещественное число $q \in [0, 1]$:

$$\Pr(\omega_1) = p, \qquad \Pr(\omega_2) = q.$$

Тогда отображение $\Pr : \Omega \to [0, 1]$ будет называться вероятностью, если у нас будет выполнено условие нормировки:

$$\Pr(\omega_1) + \Pr(\omega_2) = p + q = 1.$$

1.1.5. Зная вероятность $\Pr(\omega)$ любого элементарного исхода, можно по формуле

$$\Pr(A) = \sum_{\omega\in A} \Pr(\omega) \tag{2}$$

определить вероятность случайного события A. Иными словами, мы с помощью этой формулы можем продолжить отображение $\Pr$ на все элементы A заданной алгебры событий $\mathcal{A}$ и считать, что функция $\Pr$ задана не на Ω, а на некоторой алгебре событий $\mathcal{A}$.

Из определения (2) сразу же вытекают следующие простейшие свойства вероятности:

$$\Pr(\varnothing) = 0, \qquad \Pr(\Omega) = 1;$$

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \qquad \forall\ A, B \in \mathcal{A}.$$

События называются *несовместными,* если $A \cap B = \varnothing$. Для таких событий

$$\Pr(A \cup B) = \Pr(A) + \Pr(B). \tag{3}$$

Следствием формулы (3) является следующее полезное равенство:

$$\Pr(\bar{A}) = 1 - \Pr(A), \qquad \text{где} \qquad \bar{A} = \Omega \setminus A. \tag{4}$$

Формулы (3) и (4) легко обобщаются на случай набора $\{A_1, A_2, \ldots, A_n\}$ несовместных событий, а также на случай набора $\mathcal{A}_0 = \{A_1, A_2, \ldots, A_n\}$ подмножеств, образующих разбиение множества Ω. В первом случае имеем равенство

$$\Pr(A_1 \cup A_2 \cup \ldots \cup A_n) = \Pr(A_1) + \Pr(A_2) + \ldots + \Pr(A_n), \tag{5}$$

называемое *формулой сложения вероятностей несовместных событий.* Во втором случае получаем равенство

$$\Pr(A_1) + \Pr(A_2) + \ldots + \Pr(A_n) = 1.$$

Набор $\mathcal{A}_0 = \{A_1, A_2, \ldots, A_n\}$ называется при этом *полной группой несовместных событий.*

1.1.6. Приведенное в предыдущем пункте определение вероятности годится только лишь для случая конечного или счетного множества Ω. Саму вероятность $\Pr$ называют при этом дискретной вероятностью. В случае несчетного Ω вероятность любого элементарного исхода ω, как правило, равняется нулю. Как следствие, формулы (1) и (2) перестают работать. Поэтому в случае несчетного множества Ω вероятность $\Pr$ определяется не как функция на множестве событий Ω, а сразу как функция на некоторой алгебре событий $\mathcal{A}$. При таком подходе равенства (1) и (5) становятся аксиомами теории вероятности.

Иными словами, в дискретном случае мы можем ограничиться множеством Ω элементарных событий и заданной на нем функцией $\Pr : \Omega \to [0, 1]$, удовлетворяющей условию нормировки (1). В непрерывном же случае нам необходимо рассматривать тройку $(\Omega, \mathcal{A}, \Pr)$, где Ω — множество исходов, $\mathcal{A}$ — некоторая алгебра подмножеств Ω, а $\Pr$ — заданная на $\mathcal{A}$ вероятность. При этом говорят, что тройка $(\Omega, \mathcal{A}, \Pr)$ определяет *вероятностную модель* или *вероятностное пространство* некоторого случайного эксперимента с множеством исходов Ω и алгеброй событий $\mathcal{A}$.

Как мы уже сказали, мы в данном курсе будем рассматривать только случай конечного или счетного множества Ω, то есть будем работать с дискретной вероятностью. Однако иногда для единообразия мы все же будем и в дискретном случае использовать тройку $(\Omega, \mathcal{A}, \Pr)$ для описания случайных экспериментов.

Так, случайный эксперимент с подбрасыванием монетки можно описать чуть более сложно, рассмотрев наряду с Ω алгебру $\mathcal{A}$ событий вида

$$\mathcal{A} = 2^{\Omega} = \{\varnothing, \omega_1, \omega_2, \Omega\}. \tag{6}$$

Задавая на этой алгебре вероятность $\Pr$ соотношениями

$$\Pr(\varnothing) = 0, \qquad \Pr(\omega_1) = p, \qquad \Pr_1(\omega_2) = q, \qquad \Pr_1(\Omega) = 1, \tag{7}$$

мы получим для этой задачи вероятностное пространство $(\Omega, \mathcal{A}, \Pr)$.

1.1.7. Вообще говоря, вероятности элементарных событий можно определять достаточно произвольно. Так, если мы считаем монетку сделанной идеально, то вероятность p выпадения орла совпадает с вероятностью q выпадения решки и равна одной второй. В случае же монетки со смещенным центром тяжести вероятность этих элементарных событий может различаться. Так, ничто не мешает нам, например, определить $p = 1/4$, а $q = 3/4$.

Однако на практике все же часто удобно считать, что все различные элементарные события являются равновероятными. В этом случае для любого $\omega \in \Omega$

$$\Pr(\omega) = \frac{1}{|\Omega|}, \qquad \text{и, как следствие,} \qquad \Pr(A) = \frac{|A|}{|\Omega|}.$$

Такой способ задания вероятности называется классическим. В этом случае подсчет вероятности $\Pr(A)$ события A сводится к подсчету количества исходов, к этому событию приводящих. Иными словами, в этом случае задача становится чисто комбинаторной.

Пример 1.1. Вернемся к одной из классических урновых схем, а именно, к задаче о подсчете упорядоченных выборок с повторениями. Переформулируем эту задачу на языке дискретной вероятности.

Пусть в урне имеется n различимых предметов. Будем поочередно вытаскивать k предметов, записывать, какой из предметов и в каком порядке мы вытащили, а затем возвращать каждый предмет обратно в урну. Данная последовательность действий представляет собой случайный эксперимент, элементарным исходом которого является упорядоченный набор $(a_1, a_2, \dots, a_k)$, $a_i \in X$ для любого $i = 1, \dots, n$. Множество Ω всех возможных исходов выглядит в этом случае так:

$$\Omega = \{\omega : \ \omega = (a_1, a_2, \dots, a_k), \quad a_i \in X\}.$$

Считая, что любые элементарные исходы данного случайного эксперимента являются равновероятными, мы получаем следующую вероятность любого элементарного исхода в этом эксперименте:

$$\Pr(\omega) = \frac{1}{|\Omega|} = \frac{1}{n^k}.$$

Рассмотрим теперь событие A, заключающееся в отсутствии повторений элементов a_i в наборе $(a_1, a_2, \dots, a_k)$:

$$A = \{\omega : \ a_1 \neq a_2 \neq \ldots \neq a_k\}.$$

Из элементарной комбинаторики известно, что $|A| = (n)_k$ — количеству k-перестановок из n элементов без повторений. Как следствие,

$$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{(n)_k}{n^k} = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\dots\left(1 - \frac{k-1}{n}\right).$$

Рассмотренный пример является достаточно характерным — многие задачи элементарной комбинаторики могут быть переформулированы на языке теории вероятности. Например, задача об определении количества трехзначных чисел, содержащих цифры 3 и 6, может рассматриваться и как задача о нахождении вероятности того, что произвольно выбранное трехзначное число содержит цифры 3 и 6.

1.2. Перейдем теперь к важному понятию независимых событий, а также к тесно связанному с ним понятию условной вероятности

1.2.1. Начнем с достаточно характерного примера. Рассмотрим случайный эксперимент, связанный с подбрасыванием игральной кости. Обозначим через A событие, заключающееся в том, что у нас выпало число, большее трех, а через B — событие, состоящее в том, что у нас выпало четное число. Ясно, что

$$\Pr(A) = \frac{3}{6} = \frac{1}{2}, \qquad \Pr(B) = \frac{3}{6} = \frac{1}{2}.$$

Теперь предположим, что мы получили некоторую дополнительную информацию, а именно, нам стало известно, что в результате случайного эксперимента у нас произошло событие A. Мы не знаем, произошло ли у нас событие B, однако появившаяся у нас дополнительная информация позволяет нам утверждать, что вероятность наступления события B увеличилась.

Действительно, тот факт, что у нас произошло событие A, сужает множество возможных исходов случайного эксперимента до подмножества $A = \{4, 5, 6\}$. Два исхода из этих трех — выпадение чисел 4 и 6 — являются благоприятными для наступления события B. Иными словами, вероятность $\Pr(B|A)$ наступления события B при условии, что событие A произошло, возрастает и становится равной $2/3$.

1.2.2. Данный пример удобно иллюстрировать на диаграммах Эйлера-Венна (рис.1). Как мы знаем, в случае, когда все элементарные исходы равновероятны, вероятности наступления событий A и B равны

$$\Pr(A) = \frac{|A|}{|\Omega|}, \qquad \Pr(B) = \frac{|B|}{|\Omega|}.$$

Информация о том, что у нас произошло событие A, сужает для B пространство возможных исходов с Ω до A. При этом все исходы, принадлежащие $A \cap B$, являются благоприятными для наступления события B, так что вероятность $\Pr(B|A)$ наступления события B при условии, что событие A произошло, становится равной

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}.$$

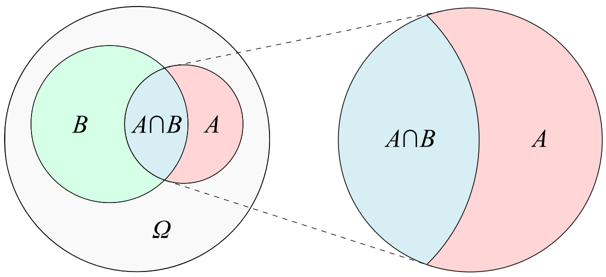

Рис. 1

1.2.3. Теперь мы можем перейти к определению зависимых и независимых событий, а также условной вероятности. По определению, условной вероятностью $\Pr(B|A)$ называется величина

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}, \qquad \Pr(A) > 0, \tag{8}$$

Два события A и B называются зависимыми, если $\Pr(B|A) \neq P(B)$, и независимыми в противном случае. Как следствие, в случае независимых событий

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B).$$

1.2.4. Вернемся к примеру с подбрасыванием игральной кости. Предположим теперь, что нам кто-то сообщил о том, что в результате случайного эксперимента событие A не произошло (или, что то же самое, произошло событие $\bar{A} = \Omega \backslash A$). В этом случае шансы на наступление события B уменьшились — из множества $\bar{A} = \{1, 2, 3\}$ возможных исходов только один исход — выпадение числа 2 — является для B благоприятным. Иными словами, вероятность $\Pr(B|\bar{A})$ при условии, что произошло событие $\bar{A}$, равна $1/3$. Заметим, что у нас при этом выполняется любопытное равенство вида

$$\Pr(A) \cdot \Pr(B|A) + \Pr(\bar{A}) \cdot \Pr(B|\bar{A}) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2} = \Pr(B). \tag{9}$$

Оказывается, и в общем случае двух произвольных событий A и B имеет место равенство

$$\Pr(B) = \Pr(A) \cdot \Pr(B|A) + \Pr(\bar{A}) \cdot \Pr(B|\bar{A}),$$

называемое *формулой полной вероятности.* Для ее доказательства достаточно записать следующее очевидное равенство, которое прекрасно иллюстрируется рис.1:

$$\Pr(B) = \Pr(B \cap A) + \Pr(B \cap \bar{A}).$$

Но по определению условной вероятности,

$$\Pr(B \cap A) = \Pr(A) \cdot \Pr(B|A) \qquad \text{и} \qquad \Pr(B \cap \bar{A}) \Pr(\bar{A}) \cdot \Pr(B|\bar{A}).$$

Тем самым формула полной вероятности доказана.

1.2.5. Формулу полной вероятности обычно записывают в несколько более общем виде. Именно, рассмотрим полную группу несовместных событий $A_1, A_2, \ldots, A_n$. Ясно, что

$$B = B \cap \Omega = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n),$$

откуда на основании свойства (3) следует, что

$$\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \ldots + \Pr(B \cap A_n).$$

С учетом формулы (8) вероятности $\Pr(B \cap A_i)$ можно выразить через условные вероятности $\Pr(B|A_i)$:

$$\Pr(B \cap A_i) = \Pr(B|A_i) \cdot \Pr(A_i). \tag{10}$$

Отсюда окончательно получается следующая обобщенная формула полной вероятности:

$$\Pr(B) = \sum_{i=1}^{n} P(B|A_i) \cdot \Pr(A_i). \tag{11}$$

1.2.6. Приведем достаточно характерный пример использования формулы (11).

Пример 1.2. Пусть в магазине имеется 100 лампочек, 60 из которых сделаны производителем номер 1, 25 — производителем номер 2, и 15 — производителем номер 3. Вероятность выхода из строя лампочки в первую неделю ее работы у первого производителя равна 0.02, у второго — 0.01, и у третьего — 0.03. Какова вероятность того, что случайно купленная в магазине лампочка выйдет из строя в течение первой недели?

Решение. Обозначим через A_i событие, заключающееся в том, что купленная лампочка принадлежит i-му производителю. Очевидно, что событие Ω, отвечающее покупке лампочки, является достоверным событием (то есть $\Pr(\Omega) = 1$), а $A_1 \cup A_2 \cup A_3 = \Omega$, то есть события A_i образуют полную группу несовместных событий. При этом $\Pr(A_1) = 0.6$, $\Pr(A_2) = 0.25$, $\Pr(A_3) = 0.15$.

Далее, вероятность выхода из строя лампочки в первую неделю ее работы у i-го производителя является, очевидно, условной вероятностью $\Pr(B|A_i)$, где B — событие, состоящее в выходе из строя лампочки в первую неделю ее работы. Следовательно, согласно формуле (11) полной вероятности,

$$\Pr(B) = 0.6 \cdot 0.02 + 0.25 \cdot 0.01 + 0.03 \cdot 0.15 = 0.019.$$

1.2.7. В примере 1.2 постановка задачи была в определенном смысле прямой: у нас было известно, что хотя бы одно из трех событий A_i произошло (лампочка была куплена), и мы искали вероятность наступления события B, состоящего в том, что купленная нами в магазине лампочка в первую неделю перегорела. На практике, однако, нас может интересовать и такая постановка задачи: пусть лампочка у нас в первую неделю все же перегорела; какова вероятность того, что в этом случае (то есть при наступлении события B) эта лампочка принадлежит, к примеру, 2-му производителю?

С формальной точки зрения речь идет о вычислении условной вероятности $\Pr(A_i|B)$: событие B произошло, плюс произошло и какое-то из трех возможных событий A_i; нас же интересует вероятность того, что в этом случае до наступления события B случилось именно событие A_2, а не два других события.

Оказывается, на такой вопрос также достаточно легко ответить. Именно, предположим, что вероятность $\Pr(B)$ события B строго больше нуля. Заметим, что тогда наряду с формулами

$$\Pr(B|A_i) = \frac{\Pr(B \cap A_i)}{\Pr(A_i)} \qquad \Longleftrightarrow \qquad \Pr(B \cap A_i) = \Pr(B|A_i) \cdot \Pr(A_i)$$

для любого $i = 1, 2, \ldots, n$ мы можем написать и аналогичные равенства

$$\Pr(B \cap A_i) = \Pr(A_i|B) \cdot \Pr(B) \qquad \Longleftrightarrow \qquad \Pr(A_i|B) = \frac{\Pr(B \cap A_i)}{\Pr(B)}.$$

Вспоминая теперь, что $\Pr(B \cap A_i)$ вычисляется по формуле (10), а также то, что для $\Pr(B)$ справедлива формула полной вероятности (1.2), мы для $\Pr(A_i|B)$ получаем соотношение

$$\Pr(A_i|B) = \frac{\Pr(A_i) \cdot \Pr(B|A_i)}{\Pr(A_1) \cdot \Pr(B|A_1) + \Pr(A_2) \cdot \Pr(B|A_2) + \ldots + \Pr(A_n) \cdot \Pr(B|A_n)}. \tag{12}$$

В частности, в нашей задаче вероятность того, что лампочка была изготовлена вторым производителем, равна

$$\Pr(A_2|B) = \frac{0.25 \cdot 0.01}{0.6 \cdot 0.02 + 0.25 \cdot 0.01 + 0.03 \cdot 0.15} = \frac{0.01 \cdot 0.25}{0.019} = \frac{5}{38} = 0.13.$$

1.2.8. Формула (12) носит название *теоремы Байеса* и играет достаточно важную роль в разного рода практических задачах. В этих задачах события A_i часто называют *гипотезами,* вероятность $\Pr(A_i)$ — *априорной* вероятностью гипотезы A_i, а вероятность $\Pr(A_i|B)$ трактуется как *апостериорная* вероятность наступления события A_i, то есть вероятность этого события *после* наступления события B.

Одна из наиболее популярных задач в этой области — это нахождение так называемой *наиболее вероятной гипотезы,* то есть события A_i, для которого $\Pr(A_i|B)$ будет наибольшим среди всех A_i. В нашей задаче таковой будет, очевидно, гипотеза, состоящая в том, что перегоревшая лампочка была изготовлена первым предприятием:

$$\Pr(A_1|B) = \frac{0.6 \cdot 0.02}{0.019} = \frac{12}{19} = 0.63.$$

Так как знаменатель в формуле Байеса (12) равен $\Pr(B)$ и не зависит от A_i, то общем случае для определения наиболее вероятной гипотезы следует найти такую гипотезу A_i, для которой величина $\Pr(B|A_i) \cdot \Pr(A_i)$ будет максимальной.

Довольно часто в такого рода задачах все априорные вероятности считаются одинаковыми и равными $1/n$. В этом случае наиболее вероятной гипотезой будет, очевидно, событие A_i с наибольшей условной вероятностью $\Pr(B|A_i)$.

1.3. Вернемся к чрезвычайно важному в теории вероятности понятию независимости случайных событий.

1.3.1. Обычно определить, являются ли какие-то два отдельных события A и B независимыми, довольно затруднительно. Например, не проводя соответствующих вычислений, не очень понятно, являются ли независимыми события "выпадение на кубике нечетного числа" и "выпадение на кубике числа, меньшего или равного трем". Однако существует довольно распространенная конструкция, в которой пары независимых событий возникают довольно естественно, по построению. Для того, чтобы эту конструкцию проще было понять, мы опишем вначале схему ее построения на достаточно простом, но в то же время важном примере — на так называемой схеме Бернулли.

1.3.2. Рассмотрим случайный эксперимент, состоящий из последовательности n одинаковых независимых случайных испытаний, то есть таких испытаний, результаты каждого из которых никак не зависят от результатов прочих испытаний. Предположим, что в каждом таком испытании возможны ровно два исхода, называемые успехом и неудачей, причем вероятность успеха в каждом эксперименте одинакова и равна $p \in (0,1)$, а вероятность q неудачи равна, соответственно, $1-p$. Такого рода случайный эксперимент и называется *схемой Бернулли.*

Данная схема моделирует, например, случайный эксперимент, заключающийся в n-кратном подбрасывании монетки. Если монетка правильная, то $p = q = 1/2$. В противном случае имеем так называемую несимметричную монетку, для которой $p \neq q$. Кроме этого простейшего примера, схема Бернулли моделирует множество других случайных экспериментов [**?**].

1.3.3. С целью формального описания схемы Бернулли вернемся к простейшему случайному эксперименту, заключающемуся в подбрасывании монетки. Напомним, что вероятностное пространство для такого эксперимента описывается соотношениями (6)–(7).

Теперь несколько усложним его, а именно, предположим, что мы n раз подбрасываем данную монету. Результат любого такого эксперимента можно записать в виде битовой строки вида $(a_1, a_2, \dots, a_n)$, в которой $a_i = 1$ в случае, если, например, у нас выпал орел, и $a_i = 0$ в случае выпадения решки. При таком подходе пространство всех исходов Ω_n и алгебра событий $\mathcal{A}_n$ имеют вид

$$\Omega_n = \{\omega :\ \omega = (a_1, a_2, \dots, a_n)\}, \qquad \mathcal{A}_n = 2^{\Omega_n}. \tag{13}$$

Покажем, что вероятность $\Pr_n$ любого элементарного события ω можно в этом случае задать с помощью соотношения

$$\Pr_n(\omega) = p^k q^{n-k}, \qquad k = \sum_{i=1}^{n} a_i, \qquad p, q > 0, \qquad p + q = 1. \tag{14}$$

Действительно, очевидно, что $\Pr_n(\omega) \in (0,1)$. Кроме того, количество всех элементарных исходов, для которых $\sum_i a_i = k$, совпадает с количеством $\binom{n}{k}$ битовых строк длины n, содержащих k единиц. Поэтому

$$\sum_{\omega \in \Omega_n} \Pr_n(\omega) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1.$$

Следовательно, описываемая приведенными выше соотношениями (13)–(14) тройка $(\Omega_n, \mathcal{A}_n, \mathrm{Pr}_n)$ определяет вероятностную модель, описывающую n-кратное подбрасывание монеты.

Обозначим через A_k событие, состоящее в том, что в результате серии n испытаний произошло ровно k успехов. Как мы только что показали, вероятность такого события равна

$$\mathrm{Pr}_n(A_k) = \binom{n}{k} p^k\, q^{n-k}, \qquad\qquad k = 0, 1, \ldots, n. \tag{15}$$

Очевидно, что набор $\{A_0, A_1, \ldots, A_n\}$ образует полную группу несовместных событий. Соответствующий этим событиям набор вероятностей $\{\mathrm{Pr}_n(A_0), \mathrm{Pr}_n(A_1), \ldots, \mathrm{Pr}_n(A_n)\}$ называется *биномиальным распределением* количества успехов в n испытаниях.

1.3.4. Перейдем теперь к описанию независимых событий в схеме Бернулли.

Определение 1.3. Говорят, что событие B_k зависит только от k-го испытания, если появление этого события определяется только лишь значением a_k.

Самым простым и самым важным событием такого рода является событие B_k вида

$$B_k = \{\omega : \ \ a_k = 1\}, \qquad\qquad \bar{B}_k = \{\omega : \ \ a_k = 0\}.$$

Иными словами, событие B_k заключается в появлении успеха, а $\bar{B}_k$ — неудачи на k-м испытании. Ясно, что

$$\mathrm{Pr}(B_k) = p, \qquad\qquad \mathrm{Pr}(\bar{B}_k) = q.$$

Несложно с помощью формальных выкладок также показать, что для всех $k \neq l$

$$\mathrm{Pr}(B_k \cap B_l) = p^2, \qquad \mathrm{Pr}(B_k \cap \bar{B}_l) = p \cdot q, \qquad \mathrm{Pr}(\bar{B}_k \cap \bar{B}_l) = q^2.$$

Следовательно, все такие события являются попарно независимыми. Более того, оказывается, что эти события являются также независимыми в совокупности, то есть такими событиями, для которых равенство

$$\mathrm{Pr}(B_{k_1} \cap B_{k_2} \cap \ldots \cap B_{k_i}) = \mathrm{Pr}(B_{k_1}) \cdot \mathrm{Pr}(B_{k_2}) \cdot \ldots \cdot \mathrm{Pr}(B_{k_i})$$

выполняется для любых $i = 1, \ldots, n$, $1 \leqslant k_1 < k_2 < \ldots < k_i \leqslant n$. Независимость событий B_k и дает формальное основание называть построенную вероятностную модель $(\Omega_n, \mathcal{A}_n, \mathrm{Pr}_n)$ моделью, отвечающую n независимым испытаниям с двумя исходами, или схемой Бернулли.

1.3.5. Заметим теперь, что формально та же самая схема Бернулли может быть построена и несколько по-другому. Именно, вернемся к элементарной вероятностной модели (6)–(7), описывающей однократное подбрасывание монетки, и рассмотрим декартовое произведение n множеств Ω_1:

$$\Omega_1 \times \Omega_1 \times \ldots \times \Omega_1 := \Omega.$$

Элементами ω этого множества Ω будут, как и в предыдущем случае, упорядоченные наборы вида $(a_1, a_2, \ldots, a_n)$, в которых $a_i \in \Omega_1$. Алгебра $\mathcal{A}$ ее подмножеств строится из множеств вида $A_1 \times \ldots \times A_1$, $A_1 \in \mathcal{A}_1$. Наконец, для любого $\omega = (a_1, a_2, \ldots, a_n)$ положим

$$\mathrm{Pr}(\omega) = \mathrm{Pr}_1(a_1) \cdot \mathrm{Pr}_1(a_2) \cdot \ldots \cdot \mathrm{Pr}_1(a_n).$$

Несложно убедиться, что получаемая в итоге тройка $(\Omega, \mathcal{A}, \mathrm{Pr})$ образует вероятностную модель, полностью эквивалентную введенной выше модели $(\Omega_n, \mathcal{A}_n, \mathrm{Pr}_n)$, то есть также описывающую схему Бернулли.

1.3.6. Вернемся теперь к началу данного пункта и опишем общую схему построения такой вероятностной модели, в которой достаточно естественно возникают пары независимых событий. Сразу заметим, что эта схема легко обобщается на случай n независимых событий.

Пусть $(\Omega_1, \mathcal{A}_1, \mathrm{Pr}_1)$ и $(\Omega_2, \mathcal{A}_2, \mathrm{Pr}_2)$ — пара вероятностных моделей. Образуем декартово произведение пары множеств элементарных событий $\Omega_1 \times \Omega_2$, которое обозначим через Ω. Затем из набора подмножеств вида $A_1 \times A_2$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$ стандартным образом построим алгебру событий $\mathcal{A}$. Наконец, вероятность на этой алгебре событий определим так: для любых $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$

$$\mathrm{Pr}(A_1 \times A_2) = \mathrm{Pr}_1(A_1) \cdot \mathrm{Pr}_2(A_2).$$

Полученная тройка $(\Omega, \mathcal{A}, \mathrm{Pr})$ в этом случае будет представлять собой некоторую вероятностную модель, в которой любые события вида (A_1, Ω_2) и (Ω_1, A_2) гарантированно являются независимыми. На практике большинство примеров независимых событий имеют именно такую природу.

2 Случайные величины

2.1. Начнем с определения случайной величины.

Определение 2.1. Случайной величиной называется произвольная вещественная функция $\xi = \xi(\omega)$, определенная на (конечном) пространстве Ω элементарных событий.

Так как обычно в каждой конкретной задаче имеется лишь одно пространство Ω, то вместо $\xi(\omega)$ мы, как правило, будем просто писать ξ.

В дальнейшем мы многие понятия будем объяснять на следующем простейшем примере.

Пример 2.2. Рассмотрим случайный эксперимент, состоящий в подбрасывании двух игральных костей. Случайную величину $\xi = \xi(\omega)$ для такого эксперимента можно определить, например, как арифметическую сумму очков на выпавших гранях.

2.1.1. Пусть $X = \{x_1, x_2, \ldots, x_n\}$ есть множество всех значений, которые может принимать произвольная случайная величина ξ, и пусть

$$A = \{\omega \in \Omega : \ \xi(\omega) = x_k\}$$

есть событие, заключающееся в том, что ξ принимает заданное значение x_k. Вычислим для каждого $k = 1, \ldots, n$ вероятность такого события

$$\mathrm{Pr}(A) = \mathrm{Pr}(\xi = x_k) = \sum_{\omega :\, \xi(\omega) = x_k} \mathrm{Pr}(\omega).$$

Тогда набор $\{\mathrm{Pr}(\xi = x_1), \mathrm{Pr}(\xi = x_2), \ldots, \mathrm{Pr}(\xi = x_n)\}$ называется *распределением вероятностей* случайной величины ξ. Очевидно, что при любом исходе случайная величина какое-то значение обязательно примет. Как следствие, сумма всех вероятностей из данного набора обязана быть равной единице:

$$\sum_{k=1}^{n} \mathrm{Pr}(\xi = x_k) = 1.$$

Например, случайная величина ξ из примера 2.2 принимает 11 возможных значений $x_k \in X$, $X = \{2, 3, \dots, 12\}$, и распределение вероятностей значений ξ здесь таково:

Значение x_k случайной величины ξ:	2	3	4	5	6	7	8	9	10	11	12
Вероятность $\Pr(\xi = x_k)$:	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Отдельно взятая случайная величина ξ полностью характеризуется множеством своих значений, а также распределением вероятностей этих значений. Детальная информация о структуре множества Ω всех элементарных событий нам в этом случае уже не важна. По сути, мы можем определить новое множество $\widetilde{\Omega} = X$, множество $\mathcal{B}$ всех его подмножеств, и ввести на $(X, \mathcal{B})$ вероятность $\Pr_\xi(B)$ по формуле

$$\Pr_\xi(B) = \Pr(\xi \in B) = \sum_{x_k \in B} \Pr(\xi = x_k).$$

Тройка $(X, \mathcal{B}, \Pr_\xi)$ в этом случае полностью определяет нам соответствующую вероятностную модель, в которой событие вида "случайная величина ξ принимает заданное значение $x_k \in X$" является элементарным событием.

2.1.2. Предположим теперь, что на одном и том же множестве Ω заданы две случайные величины ξ_1 и ξ_2. Для того, чтобы охарактеризовать их поведение, не зная ничего о структуре исходного вероятностного пространства, нам нужно, вообще говоря, задать их совместное распределение вероятностей, то есть сосчитать вероятности вида

$$\Pr(\xi_1 = x_k \text{ и } \xi_2 = y_j) \qquad \text{для всех} \quad x_k \in X, \quad y_j \in Y.$$

Иными словами, нам в данном случае в качестве множества значений нужно взять декартово произведение $X \times Y$ множества значений случайных величин ξ_1 и ξ_2, а затем определить вероятность для каждого элемента этого декартова произведения.

В общем случае сосчитать совместное распределение вероятностей непросто. Однако существует важный частный случай, когда вычисление этого распределения сводится к задаче вычисления распределения вероятности каждой отдельно взятой случайной величины.

Определение 2.3. Говорят, что ξ_1 и ξ_2 являются *независимыми* случайными величинами, если

$$\Pr(\xi_1 = x_k \text{ и } \xi_2 = y_j) = \Pr(\xi_1 = x_k) \cdot \Pr(\xi_2 = y_j) \qquad \text{для всех} \quad x_k \in X, \quad y_j \in Y.$$

В более общем случае рассмотрим некоторый набор $\{\xi_1, \xi_2, \dots, \xi_n\}$ случайных величин, заданных на одном и том же множестве Ω и принимающие значения в некотором конечном или счетном множестве X.

Определение 2.4. Случайные величины $\xi_1, \xi_2, \dots, x_n$ называются *независимыми (в совокупности)*, если для любых $x_1, x_2, \dots, x_n \in X$

$$\Pr(\xi_1 = x_1, \xi_2 = x_2, \dots, \xi_n = x_n) = \Pr(\xi_1 = x_1) \cdot \Pr(\xi_2 = x_2) \cdot \ldots \cdot \Pr(\xi_n = x_n).$$

2.1.3. Простейший пример независимых в совокупности случайных величин можно получить, рассматривая описанную в конце первого параграфа схему Бернулли. Именно, пусть

$$\Omega = \{\omega : \ \omega = (a_1, a_2, \dots, a_n), \quad a_i \in \{0, 1\}\}, \qquad \Pr(\omega) = p^k q^{n-k}, \quad k = \sum_{i=0}^{n} a_i.$$

Мы заметили, что события вида

$$B_1 = \{\omega:\ a_1 = 1\}, \qquad B_2 = \{\omega:\ a_2 = 1\}, \qquad \dots \qquad B_n = \{\omega:\ a_n = 1\}$$

являются независимыми в совокупности. Введем множество $X = \{0, 1\}$ и рассмотрим n случайных величин вида $\xi_k(\omega) = a_k$, принимающих значения в этом множестве. Каждая такая случайная величина ξ_k характеризует результат случайного испытания в схеме Бернулли на k-м шаге. Свойство независимости событий B_k влечет тогда тот факт, что события $\xi_1, \dots, \xi_n$ также являются независимыми в совокупности событиями. Кроме того, все эти n случайных величин имеют одинаковое распределение вероятностей

$$\Pr(\xi_k = 1) = p, \qquad\qquad \Pr(\xi_k = 0) = q.$$

Введенная таким образом совокупность $\xi_1, \xi_2, \dots, \xi_n$ независимых и одинаково распределенных случайных величин носит название последовательности бернулиевских случайных величин.

2.1.4. Рассмотрим еще несколько простых примеров независимых и зависимых случайных величин.

Пример 2.5. В примере 2.2 мы можем ввести случайную величину ξ_1 как число очков на первом кубике, и случайную величину ξ_2 как количество очков на втором кубике. Ясно, что эти случайные величины являются независимыми.

Пример 2.6. В том же примере 2.2 мы наряду с ξ можем ввести случайную величину η как произведение чисел, выпавших на двух кубиках. Распределение вероятностей данной случайной величины таково:

Значение y_j:	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
$\Pr(\eta = y_j)$:	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Случайные величины ξ и η независимыми уже не являются. Это видно и чисто интуитивно: если нам сказали, что $\xi = 2$, то мы тут же можем сказать, что $\eta = 1$. Для формального доказательство данного факта достаточно, например, заметить, что произведение вероятностей $\Pr(\xi = x_k)$ и $\Pr(\eta = y_j)$ для любых допустимых значений x_k и y_j строго больше нуля, тогда как вероятность того, что $\xi = x_k$ и одновременно $\eta = y_j$, достаточно часто равна нулю — например, если $x_k = 2$, а $y_j > 1$.

2.1.5. Вернемся к примеру 2.5. Понятно, что любое значение случайной величины ξ из примера 2.2 представляет собой сумму значений случайных величин ξ_1 и ξ_2. Возникает вопрос — возможно ли по известным распределениям вероятностей случайных величин ξ_1 и ξ_2 вычислить распределение вероятностей случайной величины ξ.

Итак, предположим, что мы имеем пару случайных величин ξ_1 и ξ_2, таких, что ξ_1 принимает значения из множества $X = \{x_1, x_2, \dots, x_n\}$, а ξ_2 — из множества $Y = \{y_1, y_2, \dots, y_m\}$. Введем новую случайную величину ξ как арифметическую сумму случайных величин ξ_1 и ξ_2. Ясно, что случайная величина $\xi = \xi_1 + \xi_2$ принимает значения из множества

$$Z = \{z:\ z = x_k + y_j, \quad k = 1, 2, \dots, n, \quad j = 1, 2, \dots, m\}.$$

Распределение же вероятностей случайной величины ξ выражается через совместное распределение вероятностей случайных величин ξ_1 и ξ_2:

$$\Pr(\xi = z) = \Pr(\xi_1 + \xi_2 = z) = \sum_{(k,j):\, x_k + y_j = z} \Pr(\xi_1 = x_k \text{ и } \xi_2 = y_j).$$

Теперь понятно, что в случае, когда случайные величины ξ_1 и ξ_2 являются независимыми, последнее соотношение можно упростить и выразить распределение вероятностей ξ через известные распределения вероятностей случайных величин ξ_1 и ξ_2:

$$\Pr(\xi = z) = \Pr(\xi_1 + \xi_2 = z) = \sum_{(k,l):\, x_k + y_j = z} \Pr(\xi_1 = x_k) \cdot \Pr(\xi_2 = y_j) = \sum_{k=1}^{n} \Pr(\xi_1 = x_k) \cdot \Pr(\xi_2 = z - x_k).$$

В последней сумме вероятность $\Pr(\xi_2 = z - x_k)$ считается равной нулю, если $z - x_k \notin Y$.

Рассмотрим в качестве примера случайную величину $\xi = \xi_1 + \xi_2$, где ξ_1, ξ_2 есть пара независимых бернуллиевских случайных величин. В этом случае $Z = \{0, 1, 2\}$, а распределение вероятностей случайной величины ξ описывается следующими соотношениями:

$$\Pr(\xi = 0) = \Pr(\xi_1 = 0) \cdot \Pr(\xi_2 = 0) = q \cdot q = q^2,$$

$$\Pr(\xi = 1) = \Pr(\xi_1 = 0) \cdot \Pr(\xi_2 = 1) + \Pr(\xi_1 = 1) \cdot \Pr(\xi_2 = 0) = 2\, p\, q,$$
$$\Pr(\xi = 2) = \Pr(\xi_1 = 1) \cdot \Pr(\xi_2 = 1) = p \cdot p = p^2.$$

Аналогичные рассуждения справедливы и в случае, когда мы имеем сумму n случайных величин. Так, в случае n независимых бернуллиевских случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ случайная величина $\xi = \xi_1 + \xi_2 + \ldots + \xi_n$, принимающая значения из множества $Z = \{0, 1, \ldots, n\}$, описывает количество успехов и неудач в схеме Бернулли. Распределение вероятностей этой случайной величины совпадает с биномиальным распределением:

$$\Pr(\xi = k) = \binom{n}{k} p^k q^{n-k}.$$

Здесь $\Pr(\xi = k)$ есть, по смыслу задачи, вероятность выбросить ровно k решек при n бросаниях монетки. Данную формулу можно доказать, например, по индукции.

2.2. Перейдем теперь к таким важным характеристикам случайной величины, как математическое ожидание и дисперсия.

2.2.1. Рассмотрим произвольный случайный эксперимент. Предположим, что на множестве Ω элементарных исходов этого случайного эксперимента задана случайная величина ξ, принимающая значения из множества X, $|X| = n$. Будем повторять этот случайный эксперимент N раз, и следить за значениями, которые принимает случайная величина ξ в этих экспериментах. Интуитивно ясно, что каждое значение $x_k \in X$ случайной величины ξ появляется в этих испытаниях $N \cdot p_k$ раз, где $p_k = \Pr(\xi = x_k)$. Тогда среднее значение, которое принимает случайная величина ξ в результате N случайных экспериментов, должно быть примерно равно

$$\frac{1}{N}[N\, p_1\, x_1 + N\, p_2\, x_2 + \ldots + N\, p_n\, x_n] = \sum_{k=1}^{n} p_k \cdot x_k.$$

Данные рассуждения приводят нас к следующему определению.

Определение 2.7. Математическим ожиданием или средним значением случайной величины ξ называется сумма вида

$$\mathrm{E}(\xi) = \sum_{k=1}^{n} p_k \cdot x_k. \tag{16}$$

Пример 2.8. Колесо рулетки имеет равномерно расположенные по колесу ячейки, которые обозначены числами от 0 до 36. Предположим, что игрок ставит на некоторое число сумму, равную одному доллару. Если эта цифра выигрывает, то он получает 36 долларов, если проигрывает — то не получает ничего. Пусть ξ есть случайная величина, равная количеству денег, которые игрок выигрывает. Подсчитаем, сколько денег в среднем за n испытаний игрок может получить обратно.

Для этого заметим, что множество X возможных значений в данном случайном эксперименте есть $X = \{0, 36\}$. Первое значение случайная величина ξ принимает с вероятностью, равной 36/37, а второе — с вероятностью 1/37. Следовательно, по формуле (16) имеем

$$\mathrm{E}(\xi) = 36 \cdot \frac{1}{37} + 0 \cdot \frac{36}{37} = \frac{36}{37} = 0.973.$$

Как видим, в среднем за каждую попытку игрок теряет 0.027 долларов.

2.2.2. Поговорим теперь об основных свойствах математического ожидания. Прежде всего, заметим, что иногда нам все же полезно вспомнить об исходном вероятностном пространстве $(\Omega, \mathcal{A}, \mathrm{Pr})$ и переписать формулу (16) в следующем эквивалентном виде:

$$\mathrm{E}(\xi) = \sum_{\omega \in \Omega} \xi(\omega)\ \mathrm{Pr}(\omega). \tag{17}$$

Действительно, для всех ω, для которых $\xi(\omega) = x_k$, имеем

$$\sum_{\omega\colon\, \xi(\omega) = x_k} \mathrm{Pr}(\omega) = \mathrm{Pr}(\xi = x_k),$$

и мы от (17) приходим к (16).

Пусть у нас теперь имеются две произвольные случайные величины ξ_1 и ξ_2, определенные на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, \mathrm{Pr})$. В этом случае, согласно формуле (17), математическое ожидание суммы $\xi = \xi_1 + \xi_2$ этих двух случайных величин равно сумме математических ожиданий $\mathrm{E}(\xi_1)$ и $\mathrm{E}(\xi_2)$:

$$\mathrm{E}(\xi_2 + \xi_2) = \sum_{\omega \in \Omega} (\xi_1(\omega) + \xi_2(\omega))\ \mathrm{Pr}(\omega) = \mathrm{E}(\xi_1) + \mathrm{E}(\xi_2).$$

Данное равенство выражает собой свойство линейности математического ожидания.

В случае произведения случайных величин простая формула, связывающая математическое ожидание $\mathrm{E}(\xi_1 \cdot \xi_2)$ с математическими ожиданиями $\mathrm{E}(\xi_1)$ и $\mathrm{E}(\xi_2)$, имеет место лишь в том случае, когда ξ_1 и ξ_2 являются независимыми случайными величинами. В этом случае

$$\mathrm{E}(\xi_1 \cdot \xi_2) = \mathrm{E}(\xi_1) \cdot \mathrm{E}(\xi_2).$$

Действительно,

$$\mathrm{E}(\xi_1 \cdot \xi_2) = \sum_{\omega \in \Omega} \xi_1(\omega)\, \xi_2(\omega)\ \mathrm{Pr}(\omega) = \sum_{k,j} x_k y_j\ \mathrm{Pr}(\xi_1 = x_k \text{ и } \xi_2 = y_j) =$$

$$= \sum_{k,j} x_k y_j \Pr(\xi_1 = x_k) \Pr(\xi_2 = y_j) = \sum_k x_k \Pr(\xi_1 = x_k) \cdot \sum_j y_j \Pr(\xi_2 = y_j) = \mathrm{E}(\xi_1) \cdot \mathrm{E}(\xi_2).$$

Так, несложно убедиться, математические ожидания случайных величин ξ_1 и ξ_2 из примера 2.5 одинаковы и равны $\mathrm{E}(\xi_1) = \mathrm{E}(\xi_2) = 7/2$. Следовательно, математическое ожидание $\mathrm{E}(\xi)$ случайной величины $\xi = \xi_1 + \xi_2$ из примера 2.2 равно 7, а математическое ожидание $\mathrm{E}(\eta)$ случайной величины $\eta = \xi_1 \cdot \xi_2$ из примера 2.6 равно 49/4.

Для любой из n бернуллиевских случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ математическое ожидание $\mathrm{E}(x_k)$ равно, очевидно, p. Следовательно, математическое ожидание случайной величины

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n$$

оказывается равным $\mathrm{E}(\xi) = n \cdot p$.

2.2.3. Перейдем теперь к понятию дисперсии $\mathrm{Var}(\xi)$ случайной величины ξ. Она характеризует степень разброса значений ξ относительно ее математического ожидания и определяется как средний квадрат отклонения от среднего:

$$\mathrm{Var}(\xi) := \mathrm{E}((\xi - \mathrm{E}(\xi))^2).$$

Так как

$$\mathrm{E}((\xi - \mathrm{E}(\xi))^2) = \mathrm{E}(\xi^2 - 2\xi\,\mathrm{E}(\xi) + (\mathrm{E}(\xi))^2) = \mathrm{E}(\xi^2) - 2\,\mathrm{E}(\xi) \cdot \mathrm{E}(\xi) + (\mathrm{E}(\xi))^2,$$

то

$$\mathrm{Var}(\xi) = \mathrm{E}(\xi^2) - (\mathrm{E}(\xi))^2. \tag{18}$$

Далее, с использованием последней формулы несложно проверяется, что в случае независимых случайных величин ξ_1 и ξ_2 дисперсия их суммы равна сумме их дисперсий:

$$\mathrm{Var}(\xi_1 + \xi_2) = \mathrm{Var}(\xi_1) + \mathrm{Var}(\xi_2).$$

Так, в примере 2.5 для дисперсии случайной величины ξ_1 согласно (18) имеем

$$\mathrm{Var}(\xi_1) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2 = \frac{35}{12},$$

поэтому дисперсия случайной величины ξ из примера 2.2 равна

$$\mathrm{Var}(\xi) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}.$$

В случае бернуллиевской случайной величины ξ_1, принимающей два значения 0 и 1 с вероятностями q и p соответственно,

$$\mathrm{Var}(\xi_1) = \mathrm{E}(\xi_1^2) - (\mathrm{E}(\xi_1))^2 = (p \cdot 1^2 + q \cdot 0^2) - p^2 = p - p^2 = p(1 - p) = pq.$$

Как следствие, если $\xi_1, \xi_2, \ldots, \xi_n$ есть последовательность бернуллиевских случайных величин и ξ есть их сумма, то $\mathrm{Var}(\xi) = npq$.

2.2.4. Теперь мы можем более строго обосновать тот факт, что дисперсия есть мера разброса случайной величины ξ.

Лемма 2.9. *Для любого вещественного $\alpha > 0$ справедливо неравенство*

$$\Pr\left((\xi - \mathrm{E}(\xi))^2 \geqslant \alpha\right) \leqslant \frac{\mathrm{Var}(\xi)}{\alpha}, \tag{19}$$

называемое неравенством Чебышева.

Доказательство. Обозначим через $\mu := \mathrm{E}(\xi)$. По определению,

$$\mathrm{Var}(\xi) = \mathrm{E}((\xi-\mu)^2) = \sum_{\omega\in\Omega} \Pr(\omega)\,(\xi(\omega)-\mu)^2 \geqslant \sum_{\omega\in\Omega\colon (\xi(\omega)-\mu)^2\geqslant\alpha} \Pr(\omega)\,(\xi(\omega)-\mu)^2 \geqslant$$

$$\geqslant \sum_{\omega\in\Omega\colon (\xi(\omega)-\mu)^2\geqslant\alpha} \Pr(\omega)\cdot\alpha = \alpha\cdot\Pr\left((\xi-\mu)^2\geqslant\alpha\right).$$

□

Смысл данного неравенства следующий: в случае, если дисперсия случайной величины ξ мала, то ξ в большинстве случаев принимает значения, близкие к своему среднему значению μ.

Часто неравенство (19) записывают несколько в ином виде. Введем наряду с $\mathrm{Var}(\xi)$ так называемое стандартное отклонение $\sigma = \sqrt{\mathrm{Var}(\xi)}$ случайной величины ξ и перейдем от параметра α к новому параметру c по формуле $\alpha = c^2\,\mathrm{Var}(\xi) = c^2\sigma^2$. Условие $(\xi-\mu)^2 \geqslant \alpha$ в этом случае принимает вид $(\xi-\mu)^2 \geqslant c^2\sigma^2$, что равносильно тому, что $|\xi-\mu| \geqslant c\sigma$. В результате неравенство можно переписать так:

$$\Pr\left(|\xi-\mu| \geqslant c\sigma\right) \leqslant \frac{1}{c^2}.$$

Последнее неравенство можно трактовать следующим образом: ξ лежит в пределах своего c-кратного стандартного отклонения σ от своего среднего значения μ, за исключением случаев, вероятность которых не превышает $1/c^2$. Например, в случае $c = 2$ имеем $1/c^2 = 0.25$, и поэтому отклонение случайной величины ξ от μ не превосходит 2σ по крайней мере для 75% испытаний. В случае $c = 10$ этот процент возрастает до 99%.