Введение

Человек повсюду встречается с процессами и явлениями, результаты которых не могут быть предсказаны заранее. Мы не можем предугадать точно оценку контрольной работы или урожай картофеля, курс доллара через неделю или объём продаж магазина, погоду в течение дня или результат соревнований, выигрыш в лотерею или срок службы электрической лампочки. Основываясь на своих представлениях или на предыдущем опыте, мы пытаемся всякий раз прогнозировать результат таких «опытов». Чаще всего наш прогноз будущего носит вероятностный характер: «завтра, скорее всего, будет солнечно», «рост курса доллара на этой неделе маловероятен», «шансы на выигрыш в лотерею малы».

Теория вероятностей изучает закономерности, возникающие в случайных экспериментах (точнее, в их математических моделях). Но результат случайного эксперимента нельзя предсказать заранее. О каких же закономерностях может идти речь в условиях непредсказуемости результата?

Давайте подбросим монету. Мы не сможем ответить на вопрос, выпадет ли монета вверх гербом или решкой. Результат индивидуального опыта непредсказуем. Но подбросим монету снова и снова. Конечно, мы не сможем сказать заранее, сколько гербов выпадет, если мы будем бросать монету 1000 раз. Но количество выпавших гербов не должно очень отличаться от 500 в силу полной симметрии монеты: герб имеет столько же шансов выпасть, сколько и решка. Более того, частота выпадения герба (отношение количества выпавших гербов к числу всех испытаний) при продолжении испытаний должна всё меньше отличаться от 0.5, т. е. от вероятности выпадения герба.

Так, Г.Бюффон в XVIII веке провёл 4040 подбрасываний монеты. Герб выпал 2048 раз, и частота выпадения герба составила 0.508. К. Пирсон в XIX веке бросал монету 24000 раз, при этом герб выпал в 12012 случаях. Частота выпадения герба равна 0.50051. Такая закономерность наблюдается не только при бросании монеты: вероятность любого события можно трактовать как долю случаев, когда это событие происходит при многократном повторении опыта.

Подведём итог. Не все случайные явления (эксперименты) можно изучать методами теории вероятностей, а лишь те, которые могут быть воспроизведены в одних и тех же условиях, хотя бы мысленно, какое угодно число раз. В случайных экспериментах можно обнаружить закономерности, главная из которых — свойство «статистической устойчивости»: если некоторое событие A может произойти в результате эксперимента, то доля $n(A)/n$ экспериментов, в которых данное событие произошло, с ростом общего числа экспериментов приближается к некоторому числу $P(A)$. Это число объективно характеризует «степень возможности» факта, что событие A произошло.

Возникает соблазн принять за вероятность события предел частоты этого события при повторении опытов. Но частотное определение (его связывают с именем Р. Мизеса) неудобно по многим причинам. Прежде всего потому, что частота появления события в большой, но конечной, серии меняется от серии к серии испытаний. А бесконечную последовательность испытаний, в которой частота наконец перешла бы в точную вероятность, реализовать невозможно. В 1900 г. на Международном Математическом Конгрессе в Париже Д. Гильберт сформулировал проблемы, стоящие перед математикой XX века. Шестая из двадцати трёх знаменитых ныне проблем Гильберта касалась теории вероятностей: «аксиоматизировать те науки, в которых математика играет важную роль, в первую очередь теорию вероятностей и механику». Эта проблема (в части, касающейся теории вероятностей) была окончательно решена А.Н. Колмогоровым в 1933 г. Предложенная им аксиоматика трактует вероятность как функцию, заданную на подмножествах некоторого множества и обладающую рядом свойств.

Сформулируем вывод: теория вероятностей изучает не реальные эксперименты, а лишь их математические модели. Вопросы определения вероятностей событий в реальном мире, т.е. выбора математической модели данного случайного явления, лежат за рамками теории вероятностей как математической науки.

Основные понятия теории вероятностей

Пространство элементарных исходов

Определение. *Пространством элементарных исходов* называется множество Ω, содержащее все возможные взимоисключающие результаты данного случайного эксперимента. Элементы множества называются *элементарными исходами* и обозначаются буквой ω.

Определение. *Событиями* называются подмножества множества Ω. Говорят, что *произошло событие* A, если эксперимент завершился одним из элементарных исходов, входящих в множество A.

Итак, элементарный исход — это мельчайший неделимый результат эксперимента, а событие может состоять из одного или сразу нескольких исходов.

Пример 1. Один раз подбрасывают игральную кость. Рассмотрим пространство элементарных исходов $\Omega = \{1, 2, 3, 4, 5, 6\}$. Элементарные исходы здесь соответствуют числу выпавших очков. Событие $A = \{1, 2\}$ произойдёт, если выпадет одно или два очка; событие $B = \{1, 3, 5\}$ означает, что выпадет нечётное число очков. Событие $C = \{6\}$ состоит из одного элементарного исхода и означает появление шести очков.

Пример 2. Подбрасываются две игральные кости. Будем считать их различимыми и назовём одну из них первой, другую — второй. Пространством элементарных исходов Ω является множество пар чисел (i, j), где i — число очков, выпавших на первой кости, j — на второй. В этом множестве $6 \times 6 = 36$ элементарных исходов:

$$\begin{array}{l} (1,1)(1,2)\ldots(1,6) \\ (2,1)(2,2)\ldots(2,6) \\ \ldots\ldots \\ (6,1)(6,2)\ldots(6,6) \end{array}$$

Заметим, что для симметричных костей все эти 36 исходов равновозможны: ни одна из этих комбинаций не имеет больше шансов выпасть, чем другая.

Упражнение 1. *Монета подбрасывается до тех пор, пока не выпадет вверх гербом. Опишите пространство элементарных исходов* Ω. *(В данном случае оно уже является бесконечным, но счётным множеством.)*

Операции над событиями. В теории вероятностей рассматривают те же операции над событиями, что и в теории множеств. Дадим определения новым событиям — результатам этих операций.

Объединением $A \bigcup B$ событий A и B называется событие, состоящее в том, что из двух событий A и B случилось хотя бы одно. *Пересечением* $A \bigcap B$ событий A и B называется событие, состоящее в том, что произошли сразу оба события A и B. Вместо $A \bigcap B$ часто пишут просто AB. *Дополнением* $A \backslash B$ события B до A называется событие, состоящее в том, что произошло событие A, но не произошло B. *Противоположным* (или *дополнительным*) к событию A называется событие $\bar{A} = A$, состоящее в том, что событие A не произошло. «Пустое множество» мы будем обозначать $\varnothing$.

Упражнение 2. *Чему равны следующие множества* $\bar{\Omega}$, $\bar{\varnothing}$, $A \bigcup \bar{A}$, $A \bigcap \bar{A}$.

Упражнение 3. *Пусть событие* A_i *означает, что* i*-я деталь бракованная, где* $1 \leq i \leq 3$ *— номер детали. Запишите с помощью операций над событиями событие* A *— «ровно две из трёх деталей бракованные».*

Сигма-алгебра событий

Определение. Множество событий, которое мы рассматриваем в эксперименте, называется *сигма-алгеброй событий* и обозначается через $\mathcal{F}$.

Множество событий может состоять из всех мыслимых подмножеств Ω, но может быть и более узким набором подмножеств. Сужение набора событий бывает необходимо, если Ω является несчётным (то есть большим) множеством (например, $\Omega = \mathbb{R}$, мы этот случай рассмотрим на следующей лекции), и число всех его подмножеств слишком велико: тогда возможны проблемы с корректным определением их вероятностей. Но сегодня ограничимся рассмотрением дискретных пространств элементарных исходов, где такие проблемы не возникают.

Множество всех подмножеств Ω через 2^Ω. Следующее упражнение объясняет выбор такого обозначения.

Упражнение 4. *Доказать, что если Ω состоит из n элементов, то в множестве всех его подмножеств ровно 2^n элементов.*

Подведем итог: сегодня мы рассматриваем только дискретные (не более, чем счетные) пространства элементарных исходов $\Omega = \{\omega_1, \omega_2, \ldots\}$; для таких Ω мы будем считать, что $\mathcal{F} = 2^\Omega$, то есть событиями являются все возможные подмножества Ω.

Вероятность и её свойства

Определение. *Вероятностью* или *вероятностной мерой* называется функция $\mathbb{P}$, действующая из $\mathcal{F}$ в множество действительных чисел $\mathbb{R}$ и обладающая следующими свойствами (их называют «аксиомами вероятности»):

(P1) Вероятность всего пространства элементарных исходов Ω равна единице: $P(\Omega) = 1$.

(P2) Для любого события $A \in \mathcal{F}$ вероятность неотрицательна и не больше 1: $0 \leq \mathbb{P}(A) \leq 1$.

(P3) Для любого конечного или счётного набора попарно непересекающихся событий $A_1, A_2, \ldots$ имеет место равенство

$$\mathbb{P}\left(\bigcup_i A_i\right) = \sum_i \mathbb{P}(A_i).$$

Упражнение 5. *Докажите, что из аксиом вероятности следуют следующие свойства:*

1. $\mathbb{P}(\varnothing) = 0$.
2. $\mathbb{P}(A) = 1 - \mathbb{P}(\bar{A})$.
3. *Если* $A \subset B$, *то* $\mathbb{P}(A) \leq \mathbb{P}(B)$.
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
5. $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$.

Определение. Тройка объектов $(\Omega, \mathcal{F}, \mathbb{P})$, в которой Ω — пространство элементарных исходов, $\mathcal{F}$ — поле событий и $\mathbb{P}$ — вероятностная мера, называется *вероятностным пространством.*

Задание вероятностного пространства — первый необходимый шаг при построении математической модели любого эксперимента.

Дискретное пространство элементарных исходов

Напомним, что пространство элементарных исходов называется *дискретным*, если множество Ω конечно или счётно: $\Omega = \{\omega_1, \omega_2, \ldots\}$. Кроме того, событием на таком пространстве можно (и мы будем) считать любое подмножество Ω (т.е. $\mathcal{F} = 2^{\Omega}$). Как в данном случая удобнее всего определить вероятность $\mathbb{P}$? Чтобы определить вероятность события, присвоим вероятность каждому элементарному исходу в отдельности. Иначе говоря, снабдим вероятностями мельчайшие «кирпичики» — элементарные исходы, из которых составляется любое событие. Тогда вероятность любого события определяется как сумма вероятностей входящих в него элементарных исходов.

Определение. Сопоставим каждому исходу ω_i число $p_i \in [0, 1]$ так, чтобы $p_1 + p_2 + \ldots = 1$. Вероятностью события A тогда назовем число

$$\mathbb{P}(A) = \sum_{\omega_i \in A} p_i,$$

равное сумме вероятностей элементарных исходов, входящих в множество A. В случае $A = \varnothing$ положим $\mathbb{P}(A) = 0$.

Упражнение 6. *Проверьте, что $\mathbb{P}$ будет удовлетворяет всем аксиомам* ***(P1)*** *–* ***(P3)****.*

Упражнение 7. *Если множество Ω счетно, то нельзя присвоить всем элементарным исходам одну и ту же вероятность. Почему?*

Классическая вероятностная схема

Для конечного множества Ω всегда возможно задать одинаковые вероятности исходов, что мы сейчас и сделаем. Предположим, что мы имеем дело с пространством исходов, состоящим из конечного числа элементов: $\Omega = \{\omega_1, \ldots, \omega_n\}$, и из каких-то соображений можем считать элементарные исходы равновозможными. Равновозможность возникает обычно в силу некоторой симметрии в эксперименте (симметричная монета, хорошо перемешанная колода карт, правильная игральная кость, отсутствие оснований предпочесть один результат другому).

Говорят, что эксперимент описывается *классической вероятностной моделью*, если пространство его элементарных исходов состоит из конечного числа равновозможных исходов. Тогда вероятность любого элементарного исхода равна $1/n$. Если событие $A = \{\omega_{i_1}, \ldots, \omega_{i_k}\}$ состоит из k элементарных исходов, то вероятность этого события равна отношению k/n:

$$\mathbb{P}(A) = p_{i_1} + \ldots + p_{i_k} = k \cdot \frac{1}{n} = \frac{k}{n}.$$

Мы видим, что вычисление вероятности в классической схеме сводится к подсчёту общего числа «шансов» и числа шансов, благоприятствующих событию. Число шансов считают с помощью формул **комбинаторики**.

Условная вероятность

Определение условной вероятности

Пример 3. Игральная кость подбрасывается один раз. Известно, что выпало более трёх очков. Какова при этом вероятность того, что выпало нечётное число очков?

Пусть событие $B = \{4, 5, 6\}$ означает, что выпало более трёх очков, событие $A = \{1, 3, 5\}$ — выпало нечётное число очков. Как понимать вероятность события A, если известно, что B случилось? Знаем, что произошло событие B, но всё равно не знаем, что именно выпало на кости. Однако теперь возможностей осталось только три: могло выпасть 4, 5 или 6 очков. Событию A из этих равновозможных исходов благоприятен единственный исход: выпадение пяти очков. Поэтому искомая вероятность равна $1/3$.

Определение. Пусть $\mathbb{P}(B) \neq 0$. Условной вероятностью события A при условии, что произошло событие B, называется число
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(AB)}{\mathbb{P}(B)}.$$
В случае $\mathbb{P}(B) = 0$ будем считать, что условная вероятность $\mathbb{P}(A|B)$ не определена.

Формула полной вероятности

Определение. *Разбиением* пространства Ω называется конечный или счётный набор $H_1, H_2, \ldots$ попарно непересекающихся (несовместных) событий, объединение которых есть всё Ω,

События $H_1, H_2, \ldots$, образующие разбиение, называют *гипотезами.* Обычно для некоторого события A можно вычислить $P(A|H_i)$ (вероятность событию A произойти при выполнении «гипотезы» H_i) и $\mathbb{P}(H_i)$ (вероятность выполнения «гипотезы» H_i). Как, используя эти данные, посчитать вероятность события A? Ответ на этот вопрос дает следующая теорема.

Теорема 1. *Пусть $H_1, H_2, \ldots$ — разбиение Ω. Тогда вероятность любого события A может быть вычислена по формуле:*
$$\mathbb{P}(A) = \sum_i \mathbb{P}(H_i)\mathbb{P}(A|H_i).$$

Доказательство. Разложим событие A в сумму попарно несовместных событий:
$$A = A \cap \Omega = A \cap (H_1 \cup H_2 \cup \ldots) = AH_1 \cup AH_2 \cup \ldots$$
Далее применим свойство вероятности **(P3)**
$$\mathbb{P}(A) = \sum_i \mathbb{P}(AH_I) = \sum_i \mathbb{P}(H_i)\mathbb{P}(A|H_i).$$
Теорема доказана. □

Формула Байеса

Теорема 2. *Пусть $H_1, H_2, \ldots$ — разбиение Ω. Тогда условная вероятность события H_k, если в результате эксперимента наблюдалось событие A, может быть вычислена по формуле:*
$$\mathbb{P}(H_k|A) = \frac{\mathbb{P}(H_k)\mathbb{P}(A|H_k)}{\sum_i \mathbb{P}(H_i)\mathbb{P}(A|H_i)}.$$

Доказательство. По определению условной вероятности,
$$\mathbb{P}(H_k|A) = \frac{\mathbb{P}(H_k A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(H_k)\mathbb{P}(A|H_k)}{\sum_i \mathbb{P}(H_i)\mathbb{P}(A|H_i)}.$$
□

Вероятности $\mathbb{P}(H_i)$, известные заранее, до проведения эксперимента, называют *априорными вероятностями* (a'priori — «до опыта»). Условные вероятности $P(H_i|A)$, называют *апостериорными вероятностями* (a'posteriori — «после опыта»). Формула Байеса позволяет переоценить априорные вероятности после того, как получено знание о результате эксперимента. Эта формула находит многочисленные применения в экономике, статистике, социологии и т.п.

Пример 4. Два стрелка подбрасывают монетку и выбирают, кто из них будет стрелять по мишени (одной пулей). Первый стрелок попадает по мишени с вероятностью 1, второй стрелок — с вероятностью 10^{-5}.

Можно сделать два предположения об эксперименте: H_1 — стреляет 1-й стрелок (выпал герб) и H_2 — стреляет 2-й стрелок (выпала решка). Априорные вероятности этих гипотез одинаковы: $\mathbb{P}(H_1) = \mathbb{P}(H_2) = 1/2$.

Как изменятся вероятности гипотез после проведения опыта? Рассмотрим событие A — пуля попала в мишень. Известно, что

$$\mathbb{P}(A|H_1) = 1, \quad \mathbb{P}(A|H_2) = 10^{-5}.$$

По формуле полной вероятности, вероятность пуле попасть в мишень равна

$$\mathbb{P}(A) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 10^{-5}.$$

Предположим, что событие A произошло. Тогда по формуле Байеса

$$\mathbb{P}(H_1|A) = \frac{\frac{1}{2} \cdot 1}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 10^{-5}} = \frac{1}{1 + 10^{-5}} \approx 0.99999,$$
$$\mathbb{P}(H_2|A) = \frac{\frac{1}{2} \cdot 10^{-5}}{\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 10^{-5}} = \frac{10^{-5}}{1 + 10^{-5}} \approx 0.00001.$$

Попадание пули в мишень сделало выпадение герба в 10^5 раз более вероятным, чем выпадение решки.

Случайные величины и их распределения

Мы уже видели, что для многих экспериментов нет никаких различий в подсчёте вероятностей событий, тогда как элементарные исходы в этих экспериментах очень различаются. С вероятностью 0.5 при броске монеты выпадает герб, на игральной кости — чётное число очков и т.д. Пора во всех таких «похожих» экспериментах для обозначения элементарных исходов использовать, например, числа. Иначе говоря, пора каждый элементарный исход заменить действительным числом, не обязательно уникальным, и работать только с числами.

Пусть задано дискретное вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$.

Определение. *Случайной величиной* X называется функция, которая каждому элементарному исходу $\omega \in \Omega$ ставит в соответствие действительное число, то есть $X : \Omega \to \mathbb{R}$.

Пример 5. Подбрасывают один раз правильную игральную кость. Рассмотрим $\Omega = \{1, 2, 3, 4, 5, 6\}$ и определим следующие случайные величины:

1. $X(\omega) = \omega$: эта случайная величина равна числу выпавших на кости очков,
2. $Y(\omega) = 1$, если $\omega = 2, 4, 6$, иначе $Y(\omega) = 0$: эта случайная величина служит индикатором того, выпало ли на кости чётное или нечётное число очков. В первом случае она становится равна единице, во втором — нулю.

Случайная величина X принимает значения $1, 2, 3, 4, 5, 6$ с вероятностями по одной шестой каждое. Например, $X = 1$, когда на кости выпало одно очко, т.е. с вероятностью $1/6$. Можно записать соответствие между значениями случайной величины X и вероятностями принимать эти значения в виде таблицы распределения вероятностей или, коротко, таблицы распределения:

X	1	2	3	4	5	6
$\mathbb{P}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Пример 6. Пусть Ω обозначает человеческую популяцию, состоящую из n человек. Это можно записать так:

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}.$$

Предположим, нас интересует распределение людей по возрасту. Обозначим через $A(\omega)$ возраст ω. Таким образом, каждой точке ω приписано число $A(\omega)$, выраженное в некоторых единицах измерения, скажем, в годах. Другими словами, отображение $\omega \mapsto A(\omega)$ является функцией с областью определения Ω. В качестве значений функции выступают целые числа.

Давайте введем удобный символизм для обозначения подмножеств выборочного пространства Ω, порождаемых случайными величинами. Например, множество тех ω из Ω, чей возраст заключен между 20 и 40 годами, будем обозначать так:

$$\{\omega |\ 20 \leq A(\omega) \leq 40\}$$

или более кратко, когда нет опасности неправильного понимания,

$$\{20 \leq A \leq 40\}.$$

Это множество является событием.

Определение. Случайная величина X имеет *дискретное распределение*, если множество ее значений не более, чем счетное. То есть существует конечный или счётный набор чисел $a_1, a_2, \ldots$ и соотвествующий набор вероятностей $p_1 = \mathbb{P}(X = a_1)$, $p_2 = \mathbb{P}(X = a_2)$, $\ldots$ таких, что $p_1 + p_2 + \ldots = 1$.

Замечание. Обратите внимание, $\{X = a\}$ тоже является краткой записью события $\{\omega |\ X(\omega) = a\}$.

Итак, случайная величина X имеет дискретное распределение, если множество её значений конечно или счётно. Дискретное распределение удобно задавать следующей таблицей, в которой $p_i = \mathbb{P}(X = a_i)$:

X	a_1	a_2	a_3	$\ldots$
$\mathbb{P}$	p_1	p_2	p_3	$\ldots$

Независимость случайных величин

Обычно вероятностную модель необходимо построить не для одного эксперимента, а для серии опытов. В этом случае нередко можно предполагать отсутствие взаимного влияния разных опытов друг на друга, их независимость.

Определение. Случайные величины $X_1, \ldots, X_n$, принимающие счетное число значений, называются независимыми тогда и только тогда, когда для любых действительных чисел $x_1, \ldots, x_n$ выполняется равенство

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdot \ldots \cdot \mathbb{P}(X_n = x_n).$$

Оно тривиально, если хотя бы один из сомножителей справа равен нулю, поэтому в определении можно ограничиться только такими x_i, которые принадлежат счетному множеству возможных значений случайных величин X_i.

Замечание. Заметим, что если мы посчитаем, например, условную вероятность события $\{X_2 = a_2\}$ при условии $\{X_1 = a_1\}$, то, в случае независимости X_1 и X_2, мы получим просто вероятность события $\{X_2 = a_2\}$:

$$\mathbb{P}(X_2 = a_2 | X_1 = a_1) = \frac{\mathbb{P}(X_2 = a_2 \text{ и } X_1 = a_1)}{\mathbb{P}(X_1 = a_1)} = \frac{\mathbb{P}(X_2 = a_2)\mathbb{P}(X_1 = a_1)}{\mathbb{P}(X_1 = a_1)} = \mathbb{P}(X_2 = a_2).$$

То есть знание исходов предыдущих испытаний не оказывает влияния на «истинные» вероятности исходов в последующих испытаниях, и в этом смысле испытания не зависят друг от друга.

Утверждение. *Пусть $X_1, \ldots, X_n$ — независимые случайные величины. Тогда произвольных счетных множеств $S_1, \ldots, S_n$*

$$\mathbb{P}(X_1 \in S_1, \ldots, X_n \in S_n) = \mathbb{P}(X_1 \in S_1) \times \ldots \times \mathbb{P}(X_n \in S_n).$$

Упражнение 8. *Докажите это несложное утверждение.*

Примеры дискретных распределений

Вырожденное распределение (константа). Случайная величина X имеет вырожденное распределение в точке c, если X принимает лишь одно значение c, т. е. $\mathbb{P}(X = c) = 1$.
Распределение Бернулли $\mathbf{B}_p$. Случайная величина X имеет распределение Бернулли с параметром $p \in (0, 1)$, если X принимает значение 1 с вероятностью p и значение 0 с вероятностью $q = 1-p$. Таблица распределения случайной величины $X \sim \mathbf{B}_p$ имеет вид:

X	0	1
$\mathbb{P}$	$1-p$	p

Распределение Пуассона Π_λ. Случайная величина X имеет распределение Пуассона с параметром $\lambda > 0$, если X принимает целые неотрицательные значения $k = 0, 1, 2, \ldots$ с вероятностями

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Биномиальное распределение $\mathbf{B}_{n,p}$. Это распределение связано с так называемой схемой Бернулли.

Определение. *Схемой Бернулли* называется последовательность независимых испытаний, в каждом из которых возможны лишь два исхода — «успех» и «неудача», при этом успех в каждом испытании происходит с одной и той же вероятностью $p \in (0, 1)$, а неудача — с вероятностью $q = 1 - p$. Другими словами, схема Бернулли — последовательность независимых бернуллиевских случайных величин с одинаковой вероятностью «успеха» p.

Обозначим через Y_n число успехов, случившихся в n испытаниях схемы Бернулли. Эта (случайная) величина может принимать целые значения от 0 до n в зависимости от результатов испытаний. Например, если все n испытаний завершились неудачей, то величина Y_n равна нулю.

Теорема 3 (формула Бернулли). *Для любого $k = 0, 1, \ldots, n$ вероятность получить в n испытаниях k успехов равна*

$$\mathbb{P}(Y_n = k) = C_n^k p^k q^{n-k}, \quad \text{где } C_n^k = \frac{n!}{k!(n-k)!}.$$

Доказательство. Событие $A = \{Y_n = k\}$ означает, что в n испытаниях схемы Бернулли произошло ровно k успехов. Рассмотрим один элементарный исход из события A:

$$(\underbrace{y, y, \ldots, y}_{k}, \underbrace{n, n, \ldots, n}_{n-k}),$$

когда первые k испытаний завершились успехом, остальные неудачей. Поскольку испытания независимы, вероятность такого элементарного исхода равна $p^k(1-p)^{n-k}$. Другие элементарные исходы из события A отличаются лишь расположением k успехов на n местах. Есть ровно C_n^k способов расположить k успехов на n местах (почему?). Поэтому событие A состоит из C_n^k элементарных исходов, вероятность каждого из которых равна $p^k q^{n-k}$. $\square$

Упражнение 9. *Правильная монета подбрасывается* 10 *раз. Найти вероятность того, что герб выпадет от* 4 *до* 6 *раз.*

Числовые характеристики распределений

Математическое ожидание

Определение. *Математическим ожиданием* (*средним значением, первым моментом*) случайной величины X, имеющей дискретное распределение со значениями $a_1, a_2, \ldots$ и соответствующими вероятностями $p_1, p_2, \ldots$, называется число $\mathbb{E}X$, которое вычисляется по формуле

$$\mathbb{E}X = \sum_i a_i \cdot \mathbb{P}(X = a_i) = \sum_i a_i p_i,$$

если данный ряд абсолютно сходится, т.е. если $\sum |a_i| p_i < \infty$. В противном случае говорят, что математическое ожидание не существует.

Математическое ожидание имеет простой физический смысл: если на прямой, как на невесомом стержне, распределить единичную массу, поместив в точки a_i массу p_i, то точка $\mathbb{E}X$ будет координатой «центра тяжести» прямой. Если математическое ожидание существует, то стержень, подвешенный в этой точке, будет находиться в состоянии равновесия.

Пример 7. Пусть случайная величина X равна числу очков, выпадающих при одном подбрасывании кубика. Тогда

$$\mathbb{E}X = \sum_{k=1}^{6} k \cdot \frac{1}{6} = 3.5.$$

В среднем при одном подбрасывании кубика выпадает 3.5 очка. Это несмотря на то, что ни при одном подбрасывании 3.5 очка выпасть не может.

Свойства математического ожидания. Во всех свойствах ниже предполагается, что рассматриваемые математические ожидания существуют.

(E1) Математическое ожидание постоянной равно ей самой: $\mathbb{E}c = c$.

(E2) Постоянный множитель можно вынести за знак математического ожидания: $\mathbb{E}(cX) = c\mathbb{E}X$.

(E3) Математическое ожидание суммы *любых* случайных величин равно сумме их математических ожиданий: $\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y$.

(E4) Математическое ожидание произведения *независимых* случайных величин равно произведению их математических ожиданий: $\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y$, если X и Y независимы.

(E5) Если $X \geq 0$, то $\mathbb{E}X \geq 0$.

Упражнение 10. *Докажите свойства **(E1)**, **(E2)** и **(E3)**.*

Из этих свойств можно вывести много следствий. Например:

Следствие. *Если $X \leq Y$, то $\mathbb{E}X \leq \mathbb{E}Y$.*

Следствие. *Если $a \le X \le b$, то $a \le \mathbb{E}X \le b$.*

Второе следствие особенно очевидно: центр тяжести стержня не может находиться вне отрезка, если вся масса сосредоточена на этом отрезке.

Если нам необходимо вычислить математическое ожидание величины $g(X)$ для произвольной функции $g : \mathbb{R} \to \mathbb{R}$, то можно воспользоваться формулой:

$$\mathbb{E}g(X) = \sum_i g(a_i)\mathbb{P}(X = a_i).$$

Упражнение 11. *Почему эта формула верна?*

Дисперсия и моменты старших порядков

Определение. Пусть $E|X|^k < \infty$. Число EX^k называется моментом порядка k или k-м моментом случайной величины X, число $E|X|^k$ — абсолютным k-м моментом, число $E(X - \mathbb{E}X)^k$ — центральным k-м моментом, и число $E|X - \mathbb{E}X|^k$ — абсолютным центральным k-м моментом случайной величины X. Число $\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2$ (центральный момент второго порядка) называется *дисперсией* случайной величины X.

Дисперсия $\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2$ есть среднее значение квадрата отклонения случайной величины X от её среднего. Посмотрим, за что эта величина отвечает.

Пример 8. Пусть случайная величина X принимает значения ± 1 с равными вероятностями, а случайная величина Y — значения ± 10 с равными вероятностями. Тогда $\mathbb{E}X = \mathbb{E}Y = 0$, поэтому $\mathrm{Var}(X) = \mathbb{E}X^2 = 1$, $\mathrm{Var}(Y) = \mathbb{E}Y^2 = 100$. Говорят, что дисперсия характеризует *степень разброса* значений случайной величины вокруг её математического ожидания.

Теорема 4. *Дисперсия может быть вычислена по формуле:* $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}X)^2$.

Доказательство. Обозначим для удобства $a = \mathbb{E}X$. Тогда

$$\mathrm{Var}(X) = \mathbb{E}(X - a)^2 = \mathbb{E}(X^2 - 2aX + a^2) = \mathbb{E}X^2 - 2a\mathbb{E}X + a^2 = \mathbb{E}X^2 - a^2.$$

Теорема доказана. □

Свойства дисперсии. Свойства дисперсии следуют из соответствующих свойств математического ожидания. Во всех свойствах предполагается существование вторых моментов случайных величин.

(V1) При умножении случайной величины на постоянную c дисперсия увеличивается в c^2 раз: $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$.

(V2) Дисперсия всегда неотрицательна: $\mathrm{Var}(X) > 0$.

(V3) Дисперсия не зависит от сдвига случайной величины на постоянную: $\mathrm{Var}(X + c) = \mathrm{Var}(X)$.

(V4) Если X и Y *независимы*, то $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Упражнение 12. *Докажите свойства **(V1)**, **(V2)** и **(V3)**.*

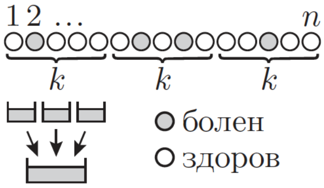
Замечание. Обратите внимание, что дисперсия суммы случайных величин равна сумме дисперсий только в случае, если случайные величины независимы!

Поиск больных

Применим элементарную теорию вероятностей к решению одной проблемы выявления больных.

Во время второй мировой войны всех призывников в армию США подвергали медицинскому обследованию. Реакция Вассермана позволяет обнаруживать в крови больных сифилисом определенные антитела. Р. Дорфманом была предложена простая методика, на основе которой необходимое для выявления всех больных число проверок удалось уменьшить в 5 раз!

Методика. Смешиваются пробы крови k человек и анализируется полученная смесь. Если антител нет, то этой одной проверки достаточно для k человек. В противном случае кровь каждого человека из этой группы нужно исследовать отдельно, и для k человек всего потребуется $k+1$ раз провести анализ.



Вероятностная модель. Предположим, что вероятность обнаружения антител p одна и та же для всех n обследуемых, и результаты анализов для различных людей независимы, т.е. моделью является последовательность из n испытаний Бернулли с вероятностью «успеха» p.

Допустим для простоты, что n делится нацело на k. Тогда надо проверить n/k групп обследуемых. Пусть X_j — количество проверок, потребовавшихся в j-й группе, $j = 1, \ldots, n/k$. Тогда

$$X_j = \begin{cases} 1 & \text{с вероятностью } (1-p)^k \text{ (все } k \text{ человек здоровы)}, \\ k+1 & \text{с вероятностью } 1-(1-p)^k \text{ (есть больные)}. \end{cases}$$

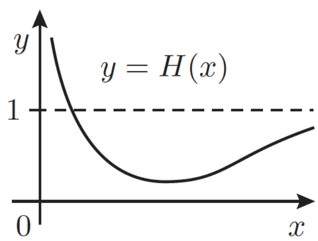
Обозначим общее число проверок $X_1 + \ldots + X_{n/k}$ через Z. Задача заключается в том, как для заданного значения p определить размер группы $k_0 = k_0(p)$, минимизирующий $\mathbb{E}Z$. Согласно определению математического ожидания находим

$$\mathbb{E}X_j = 1 \cdot (1-p)^k + (k+1) \cdot (1-(1-p)^k) = k + 1 - k(1-p)^k.$$

Отсюда по свойству математического ожидания **(E3)** имеем

$$\mathbb{E}Z = \mathbb{E}X_1 + \ldots + \mathbb{E}X_{n/k} = \frac{n}{k}\,\mathbb{E}X_1 = n(1 + 1/k - (1-p)^k).$$

Положим $H(x) = 1 + 1/x - (1-p)^x$ при $x > 0$.



Для близких к нулю значений p минимум функции $H(x)$ достигается в точке x_0, где x_0 — наименьший из корней уравнения $H(x) = 0$, т. е. уравнения

$$1/x^2 + (1-p)^x \ln(1-p) = 0.$$

Его нельзя разрешить явно относительно x. Поэтому, используя формулу $(1-p)^x \approx 1 - px$ при малых p, заменим $H(x)$ на функцию

$$\tilde{H}(x) = 1 + 1/x - 1 + px = 1/x + px,$$

имеющую точку минимума $\tilde{x}_0 = 1/\sqrt{p}$, причем $\tilde{H}(\tilde{x}_0) = 2\sqrt{p}$. Для $p = 0,01$ получаем $\tilde{x}_0 = 10$ и $\tilde{H}(\tilde{x}_0) = 1/5$, т.е. $\mathbb{E}Z \approx n/5$.

Список литературы

[1] Н. И. Чернова. Теория вероятностей. Учебное пособие. Новосибирск, 2009.

[2] К. Л. Чжун, Ф. АитСахлиа. Элементарный курс теории вероятностей. Стохастические процессы и финансовая математика. Бином, 2011.

[3] М. Я. Кельберт, Ю. М. Сухов. Вероятность и статистика в примерах и задачах. Том 1. МЦНМО, 2007.