# Случайные величины, имеющие плотности

Мы обсудили случайные величины, принимающие только счетное множество значений. Однако даже на элементарном уровне возникает немало важных вопросов, для ответа на которые необходимо рассматривать случайные величины, не удовлетворяющие данному ограничению. Это означает, что нам придется изучать несчетные вероятностные пространства $\Omega$. В этом случае появляются технические проблемы, связанные с понятием измеримости, которые нельзя удовлетворительно изложить без использования более сложной математики. Данного вида трудности возникают из невозможности назначения вероятности каждому подмножеству несчетного выборочного пространства. Решение проблемы заключается в том, чтобы ограничиться множествами, входящими в достаточно широкий класс, который, мы напомним, называется *сигма-алгеброй* событий $\mathcal{F}$. Мы рассмотрим частную, но очень важную модель, пригодную для большинства приложений и требующую не очень сложной математической техники. Ею описываются случайные величины, обладающие так называемой «плотностью».

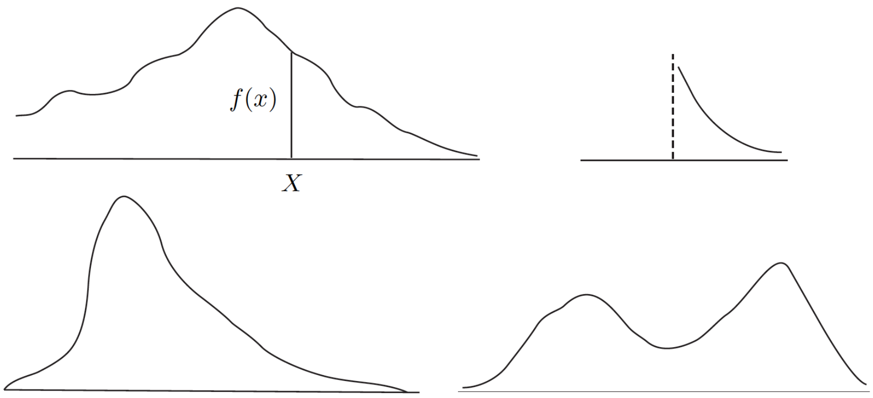
Пусть функция $f : \mathbb{R} \to \mathbb{R}$ удовлетворяет двум условиям:

1. $f(u) \geq 0$ для любого $u \in \mathbb{R}$;
2. $\displaystyle\int_{-\infty}^{+\infty} f(u)du = 1.$

Такая функция называется *плотностью* на $\mathbb{R}$. Надеемся, читатель помнит, что если функция f непрерывна или кусочно-непрерывна, то определенный интеграл $\int_a^b f(u)du$ существует для любого отрезка $[a, b]$. Тем не менее, чтобы был определен несобственный интеграл по неограниченному интервалу $(-\infty; +\infty)$ нужны дополнительные условия, обеспечивающие достаточно быстрое убывание функции $f(u)$ при больших значениях $|u|$. Такие функции называются «интегрируемыми» на $\mathbb{R}$. Требование, чтобы интеграл по всей прямой был равен 1 менее серьезно, чем может показаться. Действительно, если

$$\int_{-\infty}^{+\infty} f(u)du = M < +\infty$$

то достаточно разделить обе части на $M$ и использовать функцию $f/M$ вместо $f$. Приведем несколько возможных графиков плотностей. Какие-то из них гладкие функции, какие-то нет.




Резюмируя, ограничения на плотность таковы: кривая обязана лежать всюду выше оси абсцисс, площадь под графиком должна иметь смысл, а общая площадь равняться 1.

Теперь можно определить интересующий нас класс случайных величин на произвольном выборочном пространстве следующим образом. Пусть $X$ обозначает функцию $X : \Omega \to \mathbb{R}$, но сейчас связанные с ней вероятности для любого отрезка $[a, b]$ зададим с помощью плотности $f$ так:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(u)du.$$

Немного обобщим. Пусть теперь $A$ — не более чем счетное объединение интервалов или отрезков, не обязательно непересекающихся, причем некоторые из них могут быть бесконечными. Такие множества называются борелевскими. Мы положим

$$\mathbb{P}(X \in A) = \int_A f(u)du.$$

Тогда говорят, что случайная величина $X$ имеет плотность $f$. В некоторых книгах такие случайные величины называют абсолютно непрерывными. Отметим, что в приведенном выше определении случайной величины, имеющей плотность, подразумевается, что вероятности приписаны любым событиям вида $\{X \in A\} = \{\omega : \ X(\omega) \in A\}$. Именно такие события обычно включены в сигма-алгебру. Это ограничивает множество рассматриваемых событий, так как, например, мы не включили несчетные объединения интервалов и отрезков.

Если $X$ имеет плотность, то в силу определения при $a = b = x$ получаем:

$$\mathbb{P}(X = x) = \int_x^x f(u)du = 0.$$

Геометрически это равенство всего лишь выражает тривиальный факт, что отрезок имеет нулевую площадь. Так как $x$ произвольное число, то случайная величина $X$ принимает любое заданное значение с вероятностью нуль. Это прямо противоположно поведению случайных величин со счетным множеством значений, так как некоторые из значений должны приниматься с положительными вероятностями. Представляется парадоксальным, что с одной стороны $X(\omega)$ есть некоторое число при каждом $\omega$, а с другой стороны любое фиксированное значение принимается с вероятностью нуль. Следующий простой пример поможет прояснить ситуацию.

**Пример 1.** Пусть мы выбираем наудачу число из отрезка $[0, 1]$. Запишем теперь произвольную точку из $[0, 1]$ в десятичном представлении, например,

$$0.141592653589793\ldots. \tag{1}$$

Точка, записываемая с помощью конечного числа знаков, ничем не отличается от других, так как ее запись можно дополнить бесконечной последовательностью нулей, которые равноправны с другими цифрами. Таким образом, выбор точки из $[0, 1]$ сводится к выбору, одной за другой, цифр ее десятичного представления. Шансы получить заданную цифру, скажем, цифру 1, в первом разряде числа, равны $1/10$. Результаты выбора цифр образуют реализацию последовательности независимых испытаний. Следовательно, шансы случайно выбрать 15 цифр, формирующих число (1), составляют

$$\underbrace{\frac{1}{10} \cdot \frac{1}{10} \cdot \ldots \cdot \frac{1}{10}}_{15 \text{ раз}} = \left(\frac{1}{10}\right)^{15}.$$

Так как $10^9$ это 1 миллиард, то данная вероятность представляется настолько малой, что, согласно высказыванию Эмиля Бореля, в земной жизни ею можно пренебречь и фактически считать равной нулю! При этом мы фиксировали только 15 разрядов числа, поэтому вообще не возникает вопроса о том, чтобы выбрать наудачу само это число.

## Независимость случайных величин, имеющих плотности

Что касается случайных величин, обладающих плотностью, то они, по определению, независимы тогда и только тогда, когда

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

при условии, что множества $S_1, \ldots, S_n$ не слишком сложно устроены (являются борелевскими). Мы вернемся к этому понятию немного позже.

## Математическое ожидание и дисперсия случайных величин, имеющих плотности

Математическое ожидание случайной величины $X$, имеющей плотность $f$, определяется следующей формулой:

$$\mathbb{E}X = \int_{-\infty}^{\infty} uf(u)du.$$

А математическое ожидание случайной величины $\phi(X)$, где $\phi : \mathbb{R} \to \mathbb{R}$ — некоторая функция, определяется так:

$$\mathbb{E}\phi(X) = \int_{-\infty}^{\infty} \phi(u)f(u)du.$$

**Упражнение 1**. *Что общего у этих формул с формулами для математического ожидания дискретных случайных величин? Попробуйте объяснить, почему мы ввели именно такое определение.*

В частности, дисперсия случайной величины $X$, имеющей плотность $f$, определяется так же, как и в случае дискретных случайных величин. Пусть $\mu = \mathbb{E}X$, тогда по определению $\mathrm{Var}X = \mathbb{E}(X-\mu)^2$. Дисперсию можно вычислить либо «в лоб»

$$\mathrm{Var}X = \mathbb{E}(X-\mu)^2 = \int_{-\infty}^{\infty} (u-\mu)^2 f(u)du,$$

либо по формуле $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}X)^2$ (вторая формула тоже требует подсчета интеграла для $\mathbb{E}[X^2]$, но во многих задачах оказывается эффективнее).

Свойства математического ожидания и дисперсии для случайных величин, обладающих плотностью, полностью совпадают со свойствами, которые мы проходили для дискретных случайных величин. Повторим их здесь.

**(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 $\mathbb{E}X \geq 0$.

**(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 $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

## Функция распределения

Напомним, что дискретные случайные величины мы задавали с помощью набора вероятностей $p_1, p_2, p_3, \ldots$ с которыми эта случайная величина принимала значения $a_1, a_2, a_3, \ldots$ соответственно. Эту информацию было удобно записывать в следующую таблицу распределения вероятностей

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

Сегодня мы рассмотрели некоторый подкласс случайных величин принимающих уже несчетное количество значений. Случайные величины из этого класса мы задавали с помощью функции плотности.

**Упражнение 2**. *Почему случайные величины, имеющие плотность, нельзя задать с помощью таблицы распределения вероятностей?*

Оказывается, любую случайную величину можно задать через *функцию распределения.*

**Определение.** *Функцией распределения* случайной величины $X$ называется функция $F_X : \mathbb{R} \to [0,1]$ задаваемая следующей формулой

$$F_X(x) := \mathbb{P}(X \leq x).$$

Значение $F$ в точке $x$ аккумулирует вероятности всех возможных значений $X$ вплоть до $x$ (включительно); по этой причине к названию функции иногда добавляют прилагательное «кумулятивная». Мы иногда будем опускать нижний индекс $X$, когда понятно, о какой случайной величине идет речь. Ясно, что в дискретном случае

$$F_X(x) := \mathbb{P}(X \leq x) = \sum_{i:\ a_i \leq x} p_i,$$

а в непрерывном случае

$$F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(u)du.$$

Функция распределения $F(x)$ определяется через величины $a_i$ и $p_i$ или через функцию $f(u)$. Обратно, если известна функция $F(x)$, т. е. мы знаем значения $F(x)$ при всех $x$, то можно «восстановить» все $a_i$ и $p_i$ или $f(u)$ соответственно. В дискретном случае это довольно очевидное утверждение (Почему?). В случае, когда есть плотность $f(u)$, утверждение тоже очевидно, так как $F$ служит первообразной для плотности $f$. В силу известной теоремы из математического анализа при условии непрерывной плотности функция $f$ является производной от $F$:

$$F'(x) = f(x).$$

Таким образом, в данном случае обе функции взаимно определяют друг друга.

Отметим еще несколько свойств. Легко видеть, что если $a < b$, то

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F_X(b) - F_X(a).$$

**Упражнение 3**. *Пусть $X$ — дискретная случайная величина с функцией распределения $F(x)$. Убедитесь, что для каждого $x$ и $a < b$*

$$\mathbb{P}(a < X < b) = \lim_{\varepsilon \to +0} F(b - \varepsilon) - F(a),$$
$$\mathbb{P}(X = x) = \lim_{\varepsilon \to +0} (F(x + \varepsilon) - F(x - \varepsilon)).$$

**Упражнение 4**. *Нарисуйте графики функций распределения дискретных случайных величин, которые мы изучали на прошлой лекции. Что общего у всех этих графиков?*

## Примеры случайных величин, имеющих плотность

Рассмотрим важнейшие классы случайных величин, имеющих плотность. Начнем мы с равномерного распределения, с которым мы уже сталкивались в Примере 1.

### 1 Равномерное распределение

**Определение.** Говорят, что случайная величина имеет *равномерное распределение* на отрезке $[a, b]$, если её плотность $f(x)$ имеет вид:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

**Упражнение 5**. *Проверьте, что эта функция является плотностью. Найдите функцию распределения.*

**Упражнение 6**. *Найдите математическое ожидание и дисперсию равномерно распределенной на $[a, b]$ случайной величины.*

В следующем примере мы рассмотрим известную задачу Лапласа о последовательных восходах солнца.

**Пример 2.** Пусть солнце всходило $n$ дней подряд. Какова вероятность того, что оно взойдет еще раз?

Предполагается, что априорная вероятность восхода солнца в любой из дней является неизвестной нам постоянной величиной. Ввиду абсолютного отсутствия предпочтений будем считать, что любое из значений внутри отрезка $[0, 1]$ для нее одинаково правдоподобно. То есть данная вероятность рассматривается как случайная величина $\xi$, равномерно распределенная на отрезке $[0, 1]$. Поэтому плотностью для $\xi$ служит функция $f(p) = 1$ при $0 \leq p \leq 1$. Далее, если истинным значением случайной величины $\xi$ является $p$, то при этом предположении вероятность наблюдать $n$ последовательных восходов равна $pn$, так как они считаются независимыми событиями. Пусть $S^n$ обозначает событие «солнце всходило $n$ раз подряд». Тогда эвристически мы можем записать

$$\mathbb{P}(S^n | \xi = p) = p^n.$$

Аналог формулы полной вероятности должен тогда иметь вид

$$\mathbb{P}(S^n) = \sum_{0 \leq p \leq 1} \mathbb{P}(\xi = p)\mathbb{P}(S^n | \xi = p).$$

Конечно, строго говоря, эта запись не имеет смысла, но если заменить сумму на интеграл, то в результате получим

$$\mathbb{P}(S^n) = \int_0^1 \mathbb{P}(S^n | \xi = p) dp = \int_0^1 p^n dp = \frac{1}{n+1}.$$

Эта непрерывная версия формулы полной вероятности на самом деле верна, несмотря на нестрогость приведших к ней рассуждений. Считая данную формулу справедливой, применим ее для $n$ и $n + 1$ и, взяв отношение, найдем, что

$$\mathbb{P}(S^{n+1} | S^n) = \frac{\mathbb{P}(S^{n+1} S^n)}{\mathbb{P}(S^n)} = \frac{\mathbb{P}(S^{n+1})}{\mathbb{P}(S^n)} = \frac{n+1}{n+2}.$$

Это и есть ответ Лапласа в задаче о восходах солнца.

Ясно, что перечисленные предположения являются достаточно весомыми. Они вызывают серьезные возражения на разных уровнях. Является ли восход солнца случайным явлением или он детерминирован? Даже если его можно рассматривать как случайное явление, то будет ли он адекватно описываться нашей простой моделью? При допущении, что данная модель в принципе соответствует действительности, остается вопрос, почему априорное распределение истинной вероятности обязано быть равномерным? И как вообще можно установить его вид?

Оставляя эти вопросы в стороне, давайте на минуту вернемся к формуле

$$\mathbb{P}(S^n|\xi = p) = p^n.$$

Мы уже обсуждали, что $\mathbb{P}(\xi = p) = 0$ для любого $p$. Это выходит за рамки определения, которое мы давали на предыдущей лекции. В действительности такую условную вероятность можно совершенно корректно определить с помощью более сложного аппарата (производной Радона-Никодима). Если это сделать, то формула полной вероятности выводится без применения эвристического равенства

$$\mathbb{P}(S^n) = \sum_{0 \le p \le 1} \mathbb{P}(\xi = p)\mathbb{P}(S^n|\xi = p),$$

но это уже выходит за рамки нашего курса.

## 2 Экспоненциальное распределение

**Определение.** Случайная величина $X$ имеет *экспоненциальное* (*показательное*) *распределение* с параметром $\lambda > 0$, если её плотность имеет вид

$$f(u) = \begin{cases} \lambda e^{-\lambda u}, & u \ge 0, \\ 0, & u < 0. \end{cases}$$

**Пример 3.** Предположим, что вы стоите на автобусной остановке, расположенной на относительно спокойной проселочной дороге и наблюдаете за проезжающими мимо автомобилями. С помощью часов с секундомером вы можете засечь время, прошедшее с момента начала наблюдений до появления первого автомобиля. Эту случайную величину $T$ обычно называют *временем ожидания.* При определенных условиях разумная гипотеза заключается в том, что $T$ имеет экспоненциальное распределение.

Соответствующая функция распределения вычисляется интегрированием функции $f$:

$$F(x) = \int_{-\infty}^{x} f(u)du = \int_{0}^{x} \lambda e^{-\lambda u} du = 1 - e^{-\lambda x}.$$

В частности, если положить $x = +\infty$, или, точнее, устремить $x \to +\infty$ в этой формуле, легко увидеть, что интеграл от $f$ будет равен 1. И так как $f$ неотрицательная, то она является плотностью.

Заметим, что время ожидания имеет тенденцию сокращаться при увеличении $\lambda$. На перегруженном транспортом шоссе значение $\lambda$ на самом деле будет большим. Среднее время ожидания задается следующей формулой:

$$\mathbb{E}T = \frac{1}{\lambda}.$$

**Упражнение 7**. *Проверьте, что* $\mathbb{E}T = 1/\lambda$.

Этот результат подтверждает наше предыдущее наблюдение, что $T$ в среднем уменьшается, когда $\lambda$ возрастает. Экспоненциальное распределение служит очень полезной моделью для различных процессов, имеющих дело с временами ожидания: телефонными вызовами, временами обслуживания, распадом радиоактивных частиц и т.п.

**Пример 4.** Предположим, что в задаче из предыдущего примера нас интересует не само время ожидания $T$, а его логарифм (по основанию $e$):

$$S = \ln T.$$

Тогда $S$ также является случайной величиной. Она принимает отрицательные значения, если $T < 1$; равна нулю, если $T = 1$; положительна, если $T > 1$. Каковы связанные с ней вероятности? Интерес могут представлять вероятности вида $\mathbb{P}(a \le S \le b)$. Понятно, что для их нахождения достаточно выписать $\mathbb{P}(S \le x)$, т. е. функцию распределения $F_S$ случайной величины $S$. Так как функция $\ln x$ монотонна, то и обратная к ней функция $e^x$ обладает этим свойством. Поэтому

$$S \le x \Leftrightarrow \ln T \le x \Leftrightarrow T \le e^x.$$

Следовательно,

$$F_S(x) = \mathbb{P}(S \le x) = \mathbb{P}(T \le e^x) = 1 - e^{-\lambda e^x}.$$

Отсюда дифференцированием находим плотность $f_S$:

$$f_S(x) = F_S'(x) = \lambda e^x e^{-\lambda e^x} = \lambda e^{x-\lambda e^x}.$$

Формула выглядит довольно громоздко, но, как вы только что убедились, легко выводится.

## 3 Нормальное распределение

**Определение.** Случайная величина $X$ имеет *нормальное распределение* (или *распределение Гаусса-Лапласа*), если её плотность имеет вид

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

Будем писать $X \sim \mathcal{N}(0, 1)$. Иногда к названию добавляют уточнение «стандартное» для того, чтобы выделить данное распределение из всего семейства нормальных распределений, получающихся из него с помощью линейных преобразований, см. ниже.

Функцию распределения мы будем обозначать через $\Phi$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du.$$

Функция $\Phi$ не является элементарной, то есть, интеграл не может быть сведен к табличным и быть композицией элементарных функций. (Но ее значения, конечно, можно численно вычислить.)

Нам еще следует убедиться, что $\varphi$ действительно является плотностью в соответствии с определением, а именно, что

$$\int_{-\infty}^{+\infty} \varphi(x) dx = 1.$$

Доказательство этого факта содержится в большинстве учебников по математическому анализу, но мы воспроизведем его здесь из-за присущей ему оригинальности. Фокус состоит в том, чтобы возвести в квадрат интеграл из левой части формулы и затем преобразовать его в двойной интеграл:

$$\left(\int_{-\infty}^{+\infty}\varphi(x)dx\right)\left(\int_{-\infty}^{+\infty}\varphi(y)dy\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\varphi(x)\varphi(y)dxdy$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Для вычисления последнего интеграла используем переход к полярным координатам:

$$r^2 = x^2+y^2, \quad dxdy = r\,drd\phi.$$

$$\frac{1}{2\pi}\int_0^{2\pi}\int_0^{+\infty} re^{-\frac{r^2}{2}}drd\phi = \frac{1}{2\pi}\int_0^{2\pi} -e^{-\frac{r^2}{2}}\Big|_0^{+\infty}d\phi = \frac{1}{2\pi}\int_0^{2\pi}1\,d\phi = 1.$$

Извлечение квадратного корня приводит к необходимому равенству.

Плотность $\varphi$ нормального распределения обладает множеством замечательных аналитических свойств. В действительности, К. Гаусс определил ее путем выбора некоторых из них в качестве характеристических для «закона распределения ошибок наблюдений».

В общем случае для любых действительных чисел $a$ и $\sigma^2$ говорят, что случайная величина $X$ имеет нормальное распределение $\mathcal{N}(a,\sigma^2)$ тогда и только тогда, когда $X = \sigma Y + a$, где $Y$ — стандартная нормальная случайная величина $\mathcal{N}(0,1)$.

**Упражнение 8**. *Докажите, что плотностью нормального распределения $\mathcal{N}(a,\sigma^2)$ служит функция*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-a)^2}{2\sigma^2}} = \frac{1}{\sigma}\,\varphi\left(\frac{x-a}{\sigma}\right)$$

**Упражнение 9**. *Докажите, что для случайной величины $X \sim \mathcal{N}(a,\sigma^2)$ имеет место*

$$\mathbb{E}X = a, \quad \mathrm{Var}X = \sigma^2.$$

## Как образуются случайные величины?

Напомним, что определенная на пространстве $\Omega$ и принимающая числовые значения функция $X$ от $\omega$, называется случайной величиной (на $\Omega$). Прилагательное «случайный» напоминает, что мы имеем дело с выборочным пространством и пытаемся описывать явления, обычно называемые случайными или зависящими от случая событиями. Важно отметить, что вся случайность функции $X(\omega)$ заключается в выборе наудачу аргумента $\omega$, как при бросании игральной кости или случайном отборе индивидуума из популяции. Коль скоро аргумент $\omega$ выбран, значение функции $X(\omega)$ немедленно становится полностью определенным, всякая случайность исчезает.

**Пример 5.** Страховая компания время от времени получает заявки, содержащие требования о возмещении ущерба. Как моменты поступления требований, так и суммы страховых выплат заранее не известны и определяются случаем. При этом, конечно, общее количество заявок, скажем, за год, также случайно. Понятно, что оно будет определено, как только мы узнаем ответы на вопросы «когда» и «сколько» обо всех требованиях. Занумеруем требования в порядке поступления. Пусть $S_n$ дата поступления $n$-й заявки. Например, $S_3 = 33$ означает, что третье требование пришло 2-го февраля. Итак, мы имеем:

$$1 \le S_1 \le S_2 \le \ldots \le S_n \le \ldots$$

и равенства встречаются, когда в течение одного дня поступают несколько требований. Обозначим через $C_n$ размер выплаты (в рублях), выставленной в $n$-й заявке. Сколько требований поступит за год? Ответ $N$ задается следующей формулой:

$$N = \max\{n|\ S_n \le 365\}.$$

Очевидно, величина $N$ случайна. Однако она полностью определяется последовательностью из $S_n$. Теоретически нам надо знать всю последовательность, поскольку значение N может оказаться сколь угодно большим. Зная $N$ и последовательность из $C_n$, мы сумеем найти общую сумму всех выплат за год $C_1 + \ldots + C_N$. Конечно, эта сумма также случайна. Она зависит как от $S_n$, так и от $C_n$.

В данном случае легко представить, что требования поступают в офис одно за другим и регистрируются в виде записей в книге учета, как в дневнике. В какие-то из дней совсем не будет поступлений, в другие сразу несколько требований на разные суммы выплат. Такая книга обычно хранится годами и может выглядеть по-разному в различные временные периоды. Другая страховая компания ведет свою книгу учета, которая в чем-то похожа, но и в чем-то отличается от первой. Каждый мыслимый список требований в таких книгах может рассматриваться в качестве $\Omega$, и достаточно большое число таких списков могут служить выборочным пространством. К примеру, списки, содержащие миллион заявок за один день или выплаты в размере 95 центов, следует исключить из рассмотрения. Это позволит сохранить наше представление о выборочном пространстве в рамках реальности.

Мы уже обсуждали, как можно задавать дискретные случайные величины и случайные величины, имеющие плотность. Но, кроме этого, имея в запасе несколько случайных величин, можно получать другие величины, производя разные операции над имеющимися.

**Утверждение.** *Если $X$ и $Y$ случайные величины, то таковыми являются и*

$$X + Y,\ X - Y,\ XY,\ X/Y\ (Y \neq 0),$$

*а также $aX + bY$, где $a$ и $b$ действительные числа.*

Это утверждение немедленно вытекает из общего определения, так как, например,

$$\omega \mapsto X(\omega) + Y(\omega)$$

есть функция в силу того, что она является суммой функций $X$ и $Y$. Ситуация здесь совершенно такая же, как в математическом анализе: если $f$ и $g$ функции, то функциями являются и $f + g$, $f - g$, $fg$ и т.д. Единственное отличие заключается в том, что в математическом анализе их аргументом служит действительное число $x$, а случайные величины — функции от точки выборочного пространства.

**Утверждение.** *Если $\phi$ есть функция двух переменных, а $X$ и $Y$ случайные величины, то отображение*

$$\omega \mapsto \phi(X(\omega), Y(\omega))$$

*также является случайной величиной, которая более кратко записывается в виде $\phi(X, Y)$.*

Эти утверждения можно обобщить на случай нескольких случайных величин.

## Зависимость и связанные с ней понятия

В предыдущем разделе мы рассматривали случайные величины, которые были функциями от других случайных величин. Естественно предположить, что применяя функцию к набору случайных величин, мы получим случайную величину, которая не будет являться независимой в совокупности с исходным набором случайных величин. Чтобы дать характеристику «степени зависимости» случайных величин, нам нужно будет познакомиться с формализмом для изучения распределений случайных векторов. Как мы увидим ниже, этот формализм вполне аналогичен рассмотрению распределения одной (скалярной) случайной величины.

## Случайные векторы. Базовые понятия

Для простоты обозначений мы станем рассматривать только две случайные величины $X$ и $Y$, но обобщение на случай произвольного конечного числа случайных величин не должно вызвать затруднений. Прежде всего, займемся случаем, когда $X$ и $Y$ имеют счетные множества значений. Пусть $X$ принимает значения $\{x_i\}$, $Y$ принимает значения $\{y_j\}$, и положим

$$\mathbb{P}(X = x_i, Y = y_j) = p(x_i, y_j).$$

Когда $x_i$ и $y_j$ пробегают весь возможный диапазон значений, множество «элементарных вероятностей» $p(x_i, y_j)$ представляет собой *совместное распределение* компонент случайного вектора $(X, Y)$. Чтобы найти распределение $X$ отдельно, позволим $y_j$ принимать любое из возможных значений; при этом

$$\mathbb{P}(X = x_i) = \sum_{y_j} p(x_i, y_j).$$

Когда $x_i$ пробегает весь возможный диапазон значений, множество вероятностей $\mathbb{P}(X = x_i)$ образует *маргинальное распределение случайной величины* $X$. Аналогично определяется маргинальное распределение случайной величины $Y$. Отметим, что эти маргинальные распределения в общем случае не определяют совместное распределение двух случайных величин.

Подобно тому как выражалось математическое ожидание произвольной функции от $X$ с помощью вероятностного распределения, выразим теперь математическое ожидание произвольной функции $\phi : \mathbb{R}^2 \to \mathbb{R}$ от вектора $(X, Y)$:

$$\mathbb{E}(\phi(X, Y)) = \sum_{x_i} \sum_{y_j} \phi(x_i, y_j) p(x_i, y_j).$$

Теперь рассмотрим случай, когда существует плотность. Говорят, что случайный вектор $(X, Y)$ имеет (совместную) плотность $f$, если

$$\mathbb{P}(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, dudv$$

для всех $(x, y)$. Тогда можно показать, что для любого не слишком сложно устроенного (борелевского) подмножества $S$ плоскости верно равенство

$$\mathbb{P}((X, Y) \in S) = \iint_S f(u, v)\, dudv.$$

Плотность $f$ совместного распределения двух случайных величин должна удовлетворять следующим условиям:

1. $f(u, v) \geq 0$ для любых $u, v \in \mathbb{R}$;
2. $\displaystyle\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(u, v) dudv = 1.$

Конечно, мы подразумеваем, что функция $f$ интегрируема на всей плоскости. Определим маргинальное распределение $\mathbb{P}(X \leq x)$ и маргинальную плотность $f_1(u)$ следующим образом:

$$\mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_1(u) du, \quad \text{где} \quad f_1(u) = \int_{-\infty}^{+\infty} f(u, v) dv.$$

Маргинальное распределение и маргинальная плотность для компоненты $Y$ определяется аналогично. Формула вычисления математического ожидания преобразуется в случае существования плотности в следующую: для любой «обыкновенной» функции $\phi : \mathbb{R}^2 \to \mathbb{R}$

$$\mathbb{E}(\phi(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi(u,v) f(u,v)\, dudv.$$

Класс «обыкновенных» функций содержит все ограниченные непрерывные функции от переменных $(u,v)$; индикаторы «не слишком сложных» множеств; функции, непрерывные всюду, за исключением гладких линий разрыва, для которых интеграл существует и т.п.

В самом общем случае (совместная) функция распределения $F$ случайного вектора $(X,Y)$ определяется формулой

$$F(x,y) = \mathbb{P}(X \leq x, Y \leq y).$$

Обозначая $\lim_{y\to+\infty} F(x,y)$ через $F_1(x)$, запишем, что

$$F_1(x) = \mathbb{P}(X \leq x, Y < +\infty) = \mathbb{P}(X \leq x),$$

так как условие «$Y < +\infty$» не накладывает никаких ограничений на $Y$. Таким образом, $F_1(x)$ есть маргинальная функция распределения случайной величины $X$. Аналогично определяется маргинальная функция распределения случайной величины $Y$.

## Плотность независимых случайных величин

Напомним, что независимость случайных величин $X_1, \ldots, X_n$, имеющих плотность, определялась так:

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

при условии, что множества $S_1, \ldots, S_n$ не слишком сложно устроены (являются борелевскими). Используя понятие совместной функции распределения $F$ случайного вектора $(X_1, \ldots, X_n)$, подставляя $S_1 = (-\infty; x_1], \ldots, S_n = (-\infty; x_n]$, мы для не зависимых случайных величин получим

$$F(x_1, \ldots, x_n) = F_1(x_1) \times \ldots \times F_n(x_n),$$

где $F_j$ маргинальное распределение случайной величины $X_j$, $1 \leq j \leq n$. Таким образом, в случае независимости компонент маргинальные распределения определяют совместное распределение.

Заметим, что если $f_k$ — плотность $X_k$ для всех $k = 1, \ldots, n$, то используя независимость

$$\begin{aligned}\mathbb{P}(X_1 \in S_1, \ldots, X_n \in S_n) &= \left(\int_{S_1} f_1(u)du\right) \times \ldots \times \left(\int_{S_n} f_n(u)du\right)\\ &= \int_{S_1} \ldots \int_{S_n} f_1(u_1) \times \ldots \times f_n(u_n)\, du_1 \ldots du_n.\end{aligned}$$

Но, с другой стороны, используя совместную плотность $f$, вероятность в левой части равна

$$\mathbb{P}(X_1 \in S_1, \ldots, X_n \in S_n) = \int_{S_1} \ldots \int_{S_n} f(u_1, \ldots, u_n)\, du_1 \ldots du_n.$$

Сравнение этих двух выражений дает соотношение

$$f(u_1, \ldots, u_n) = f_1(u_1) \times \ldots \times f_n(u_n).$$

То есть совместная плотность независимых случайных величин распадается в произведение маргинальных плотностей.

**Упражнение 10**. *Если две точки выбираются наудачу из отрезка* $[0,1]$, *то какова вероятность того, что расстояние между ними окажется меньше 1/2?*

### Ковариация

Ковариация — та самая обещанная количественная характеристика «степени зависимости» случайных величин. Введем сначала определение.

**Определение.** Пусть задано совместное распределение случайных величин $X$ и $Y$. Тогда число $\mathrm{Cov}(X,Y)$ называется их *ковариацией* определяется по следующей формуле:

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)].$$

Интерпретация ковариации: если ковариация положительна, то с ростом значений одной случайной величины, значения второй имеют тенденцию возрастать, а если знак отрицательный — то убывать.

**Упражнение 11**. *Пользуясь свойствами математического ожидания, покажите, что ковариацию можно вычислять по следующей формуле:*

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y.$$

**Следствие.** *Если случайные величины $X$ и $Y$ независимы, то* $\mathrm{Cov}(X,Y) = 0$.

Обратное утверждение не верно: если $\mathrm{Cov}(X,Y)=0$, то не обязательно $X$ и $Y$ будут независимыми.

**Упражнение 12**. *Пусть случайная величина $Z$ принимает значения* $0, \frac{\pi}{2}, \pi$, *каждое с вероятностью* $\frac{1}{3}$. *Посчитайте ковариацию величин* $\sin Z$ *и* $\cos Z$. *Покажите, что*

$$P(\sin Z = 1, \cos Z = 1) \neq P(\cos Z = 1)P(\sin Z = 1).$$

Пользуясь свойствами математического ожидания, можно показать, что для произвольных случайных $X$ и $Y$ величин

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathbb{E}(X+Y-\mathbb{E}(X+Y))^2 \\
&= \mathbb{E}((X-\mathbb{E}X)+(Y-\mathbb{E}Y))^2 \\
&= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)] \\
&= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y).
\end{aligned}$$

В случае произвольного количества случайных величин верна следующая формула для дисперсии:

$$\mathrm{Var}(X_1+\ldots+X_n) = \mathrm{Var}(X_1)+\ldots+\mathrm{Var}(X_n) + 2\sum_{i<j}\mathrm{Cov}(X_i,X_j),$$

## Примеры случайных величин, имеющих плотность (продолжение)

### 4 Распределение хи-квадрат

**Упражнение 13**. *Пусть $X$ имеет стандартное нормальное распределение . Найдите плотность распределения случайной величины $X^2$ и соответствующую ей функцию распределения. В статистике оно известно под именем «распределения хи-квадрат» (с одной степенью свободы).*

**Hint.** *Продифференцируйте* $\mathbb{P}(X^2 < x) = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{x}} e^{-\frac{u^2}{2}} du.$

Распределение хи-квадрат с $k$ степенями свободы определяется следующим образом.

**Определение.** Пусть $X_1, \ldots, X_k$ — независимые стандартные нормальные случайные величины, то есть $X_i \sim \mathcal{N}(0, 1)$. Тогда случайная величина

$$Z = X_1^2 + \ldots + X_k^2$$

имеет распределение *хи-квадрат с $k$ степенями свободы*. Мы будем обозначать это так: $Z \sim \chi_k^2$.

К распределению хи-квадрат мы вернемся в курсе статистики.

## 5 Распределение Коши

**Определение.** *Распределением Коши* с параметром сдвига $\mu$ и параметром масштаба $\sigma > 0$ называется распределение с плотностью

$$f(x) = \frac{1}{\pi\sigma} \cdot \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2}.$$

Если $\mu = 0$ и $\sigma = 1$, то такое распределение называется стандартным распределением Коши.

Случайная величина, имеющая распределение Коши, является стандартным примером величины, не имеющей математического ожидания и дисперсии.

**Упражнение 14**. *Покажите, что у случайной величины, имеющей распределение Коши, не существует математического ожидания (и дисперсии).*

## Список литературы

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

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

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