Распределения случайных величин

В.В. Некруткин

кафедра статистического моделирования `http://statmod.ru`, матмех СПбГУ

Материал к практическим занятиям по теории вероятностей, 2018 г.

Содержание

1 Борелевские σ-алгебры

Пусть $(D, \mathcal{A})$ — топологическое пространство, то есть $\mathcal{A}$ (топология) — это семейство подмножеств множества D, обладающее следующими свойствами

1. $D \in \mathcal{A}$, $\varnothing \in \mathcal{A}$;
2. объединение любого числа множеств из $\mathcal{A}$ принадлежит $\mathcal{A}$;
3. пересечение любого конечного числа множеств из $\mathcal{A}$ принадлежит $\mathcal{A}$.

Любое множество $A \in \mathcal{A}$ называется открытым, а его дополнение — замкнутым.

База $\mathcal{A}_0$ топологии $\mathcal{A}$ — это такое подмножество $\mathcal{A}_0 \subset \mathcal{A}$, что любой элемент $A \in \mathcal{A}$ является объединением некоторого (не обязательно конечного или счетного) числа множеств из $\mathcal{A}_0$.

Пусть теперь (D, ρ) — метрическое пространство. Тогда метрика ρ порождает естественную топологию $\mathcal{A}$ в D следующим образом: базой этой топологии объявляется семейство всех открытых шаров $B_r(x) = \{y \in D : \rho(x, y) < r\}$ с центрами во всевозможных точках $x \in D$ и со всевозможными радиусами r.

Таким образом, здесь любое открытое множество — это просто объединение открытых шаров. В этом случае тройку $(D, \rho, \mathcal{A})$ называют метрическим топологическим пространством.

Общее определение борелевской σ-алгебры в топологическом пространстве следующее. Если $(D, \mathcal{A})$ — топологическое пространство, то минимальная σ-алгебра $\mathcal{B}$, содержащая топологию $\mathcal{A}$ (то есть содержащая все открытые множества) называется борелевской, а тройка $(D, \mathcal{A}, \mathcal{B})$ — измеримым топологическим пространством.

Тем самым по определению любое открытое множество является борелевским. Более того, поскольку любая σ-алгебра всегда замкнута относительно не более чем счетных (любых) теоретико-множественных операций над своими элементами, то любое конечное или счетное объединение (или пересечение, или дополнение) открытых и/или замкнутых множеств является борелевским множеством. Могут существовать, однако, гораздо более сложно устроенные борелевские множества.

В случае, когда $D = \mathbb{R}^d$ со стандартной евклидовой метрикой, существует эквивалентное определение борелевской σ-алгебры. В этом случае борелевскую σ-алгебру $\mathcal{B} = \mathcal{B}_d$ можно определить как минимальную σ-алгебру, содержащую все ячейки, то есть всевозможные множества вида $[a, b)$ для $D = \mathbb{R}^1$, всевозможные множества вида $[a, b) \times [c, d)$ для $D = \mathbb{R}^2$ и т.д.

2 Случайные величины и их распределения

Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathrm{P})$, где Ω — пространство элементарных событий,[1] $\mathcal{F}$ — σ-алгебра событий,[2] P — вероятностная[3] мера, определенная на $\mathcal{F}$. Те задачи, которые мы решали до сих пор, были связаны с вычислением вероятностей $\mathrm{P}(A)$ для различных событий $A \in \mathcal{F}$. Сейчас появляется новое понятие.

Определение 2.1. *Случайной величиной* ξ (со значениями в $\mathbb{R}$) называется измеримое отображение $\xi : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B})$, где $\mathcal{B}$ есть борелевская σ-алгебра подмножеств вещественной прямой.[4]

Напомним, что измеримость в данном случае означает, что соотношение

$$\big\{\omega \in \Omega : \xi(\omega) \in B\big\} = \xi^{-1}(B) \in \mathcal{F}$$

верно для любого $B \in \mathcal{B}$.

Понятие «случайная величина» возникает при измерении некоторой переменной, связанной со случайным экспериментом. Действительно, если наш случайный эксперимент состоит в «случайном» выборе элементарного события $\omega \in \Omega$, то измерение, проведенное на этом случайном эксперименте — это некоторая функция, сопоставляющая каждому ω какое-то число, обозначаемое $\xi(\omega)$. Иначе говоря — это функция $\xi : \Omega \mapsto \mathbb{R}$.[5]

Приведем примеры.

1. Бросание трех различных игральных костей.
   Здесь $\Omega = \{(i, j, k)\}$, где $i, j, k \in \{1, \ldots, 6\}$. Если мы интересуемся суммой выпавших чисел, то появляется отображение $\xi : \Omega \mapsto \mathbb{R}$, задаваемое формулой $\xi(i, j, k) = i + j + k$.

2. Стрельба по мишени.
   Предполагая, что стрельба ведется без промаха, а пулька представляет собой точку, получаем, что пространство элементарных событий — эта сама мишень. Если нас интересует число очков, которое мы выбили, мы имеем дело с отображением ξ_1, сопоставляющим каждой точке $\omega \in \Omega$ целое число, то есть снова с функцией. Другая функция $\xi_2 : \Omega \mapsto \mathbb{R}$ «отвечает» за расстояние от центра мишени до пробоины и т. д.

3. Ловля крокодилов.
   Рассмотрим случайный эксперимент, состоящий в том, что мы ловим крокодила в реке Нил. Итог эксперимента — это либо пойманный (конкретный) крокодил, либо возвращение[6] домой с пустыми руками.

   Следовательно, множеством элементарных событий естественно считать множество крокодилов, населяющих Нил, плюс особую точку Δ, означающую неудачу мероприятия.

   Вес пойманного крокодила, его длина, число зубов, стоимость — различные «измерения», связанные со случайным экспериментом.[7] В нашей терминологии это различные случайные величины, определенные на одном и том же пространстве элементарных событий.

Разберемся теперь, откуда появляется требование измеримости отображения $\xi : \Omega \mapsto \mathbb{R}$. В эксперименте с бросанием трех костей нас могут интересовать шансы выпадения большого числа очков, скажем, больше 13. На языке теории вероятностей это означает, что мы интересуемся вероятностью $\mathrm{P}(\xi > 13)$. Но отсюда следует, что множество $\{\omega : \xi(\omega) > 13\}$ должно быть событием, то есть должно принадлежать $\mathcal{F}$.[8]

[1] множество элементарных, неделимых исходов случайного эксперимента.
[2] семейство тех подмножеств Ω, которые мы договорились считать «случайными событиями».
[3] нормированная на единицу.
[4] то есть минимальная σ-алгебра, содержащая все ячейки вида $[a, b)$.
[5] Такие функции (то есть случайные величины) в теории вероятностей принято обозначать греческими буквами.
[6] Будем оптимистами!
[7] Если это необходимо, можно доопределить эти функции в точке Δ каким-то отрицательным значением.
[8] По определению мы не имеем права писать $\mathrm{P}(A)$, если $A \notin \mathcal{F}$.

В примере со стрельбой по мишени нас может интересовать шансы того, что пробоина находится на расстоянии не больше 5 мм от центра мишени, и тогда мы приходим к событию $\{\omega : \xi_2(\omega) \leq 5\}$.

А в эксперименте с крокодилом, если нас не волнуют крокодилы-младенцы и крокодилы-старики, мы приходим к необходимости считать событием множество $\{\omega : a < \xi(\omega) < b\}$ для некоторых a и b, где ξ — возраст пойманного крокодила.

Продолжая подобные рассуждения, мы видим, что $\mathcal{F}$-измеримыми должны быть всевозможные множества $\xi^{-1}(< a, b >) = \{\omega : \xi(\omega) \in < a, b >\}$,[9] их конечные или счетные объединения, пересечения, дополнения и т.д., что и приводит к формальному требованию измеримости отображения ξ.[10]

Определение 2.2. Пусть ξ является случайной величиной. Функция $\mathcal{P}_\xi : \mathcal{B} \mapsto \mathbb{R}$, определенная равенством $\mathcal{P}_\xi(B) = \mathrm{P}(\xi \in B)$[11] при каждом $D \in \mathcal{B}$, называется *распределением* случайной величины ξ.

Предложение 2.1. *Распределение $\mathcal{P}_\xi$ случайной величины ξ является вероятностной мерой.*

Тем самым случайная величина ξ порождает еще одно вероятностное пространство $(\mathbb{R}, \mathcal{B}, \mathcal{P}_\xi)$.

Предложение 2.2. *Какова бы ни была вероятностная мера $\mathcal{P}$, определенная на σ-алгебре $\mathcal{B}$ борелевских подмножеств $\mathbb{R}$, существует такое вероятностное пространство $(\Omega, \mathcal{F}, \mathrm{P})$ и такая случайная величина ξ, что распределение ξ равно $\mathcal{P}$.*[12]

Замечание 2.1. Тем самым мы можем употреблять термин *распределение* даже если случайная величина у нас отсутствует, понимая под этим просто вероятностную меру, определенную на борелевских подмножествах прямой.

Тот факт, что случайная величина ξ имеет распределение $\mathcal{P}$, мы будем кратко записывать как $\mathcal{L}(\xi) = \mathcal{P}$. Буква $\mathcal{L}$ расшифровывается как «Law», так как вместо «распределение» иногда говорят «закон распределения». В некоторых случаях вместо $\mathcal{L}(\xi) = \mathcal{P}$ мы будем писать менее точное (но более краткое) $\xi \in \mathcal{P}$. Кроме того, для обозначения распределения случайной величины ξ будет использоваться значок $\mathcal{P}_\xi$ — как мы уже и делали.

3 Дискретные и абсолютно непрерывные распределения.

Если нам известно распределение $\mathcal{P}_\xi$ случайной величины ξ, то мы имеем полную информацию о том, с какой вероятностью ξ принимает разные значения. Сам объект $\mathcal{P}_\xi$, однако, очень сложен (борелевских множеств очень много), поэтому хочется иметь более простое (но эквивалентное) описание $\mathcal{P}_\xi$. Особенно легко это делается в двух случаях — когда случайные величины имеют дискретное или абсолютно непрерывное распределения.

Дискретные распределения.

Определение 3.1. Если существует такое конечное или счетное множество X, что $\mathrm{P}(\xi \in X) = \mathcal{P}_\xi(X) = 1$, то говорят, что случайная величина ξ является *дискретной* (или обладает *дискретным распределением*).

Пусть $X = \{x_1, \ldots, x_n, \ldots\}$.[13] Кроме этого, предположим, что $\mathrm{P}(\xi = x_i) = p_i > 0$ (если некоторые $p_i = 0$, то просто исключим соответствующие x_i из рассмотрения). Очевидно, $\sum_i p_i = 1$.

[9] Треугольные скобки $< a, b >$ означают, что точки a и b могут принадлежать как этому интервалу, так и его дополнению.

[10] Можно доказать, что, если $\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$ для любых ячеек B, то этот факт верен и для любых $B \in \mathcal{B}$.

[11] Мы будем писать $\mathrm{P}(\xi \in B)$ вместо формально более точного, но и более громоздкого $\mathrm{P}\big(\{\omega : \xi(\omega) \in B\}\big)$.

[12] Идея построения следующая: нужно взять $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}$, $\mathrm{P} = \mathcal{P}$ и $\xi(\omega) = \omega$. Доказательство совсем простое.

[13] Если X конечно, то последнего троеточия нет.

Составим таблицу[14]

$$\mathcal{P}_\xi : \quad \begin{pmatrix} x_1 & x_2 & \dots & x_n & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}. \tag{3.1}$$

Будем называть таблицу (3.1) *таблицей распределения случайной величины* ξ. Очевидно, таблица распределения полностью определяет само распределение: для любого $B \in \mathcal{B}$

$$\mathrm{P}(\xi \in B) = \mathcal{P}_\xi(B) = \sum_{i:\, x_i \in B} p_i$$

и поэтому для описания $\mathcal{P}_\xi$ нам достаточно знать только эту таблицу.

В то же время распределение $\mathcal{P}_\xi$ порождает множество таблиц этого распределения: достаточно в (3.1) переставить 2 столбца, чтобы получить новую таблицу, в то время как само распределение $\mathcal{P}_\xi$ останется прежним. Тем не менее для краткости, рассматривая таблицу (3.1), мы будем иногда говорить о распределении $\mathcal{P}_\xi$, а не о конкретной таблице этого распределения.

Основные[15] дискретные распределения. Перечислим основные дискретные распределения, с которыми нам придется сталкиваться и введем соответствующие обозначения.

1. *Распределение Дирака*[16] δ_a, сосредоточенное в точке a, задается простейшей таблицей

$$\mathcal{P}_\xi : \quad \begin{pmatrix} a \\ 1 \end{pmatrix}.$$

   Распределение δ_a — это распределение случайной величины, тождественно равной a.[17]

2. *Дискретное равномерное распределение* на множестве $X = \{x_1, \dots, x_n\}$ имеет вид

$$\mathcal{P}_\xi : \quad \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ n^{-1} & n^{-1} & \dots & n^{-1} \end{pmatrix}.$$

   Интерпретация очевидна — случайная величина ξ принимает n значений $x_1, \dots, x_n$ с одинаковыми вероятностями. Обозначение — $\mathrm{U}_n(X)$.

3. *Распределение Бернулли* $\mathrm{Ber}(p)$ с параметром $0 < p < 1$. Таблица распределения:

$$\mathcal{P}_\xi : \quad \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}.$$

   Смысл распределения — это распределение числа успехов в одном испытании Бернулли с вероятностью успеха p.

4. *Биномиальное распределение* $\mathrm{Bin}(n, p)$ с параметрами $n \in \{1, 2, \dots\}$ и $0 < p < 1$. Таблица распределения:

$$\mathcal{P}_\xi : \quad \begin{pmatrix} 0 & 1 & \dots & n \\ p_0 & p_1 & \dots & p_n \end{pmatrix},$$

   где $p_k = C_n^k\, p^k\, (1-p)^{n-k}$. Интерпретация — распределение числа успехов в n испытаниях Бернулли с вероятностью успеха p. При $n = 1$ совпадает с распределением Бернулли.

[14] Предполагается, что элементы множества X уже как-то упорядочены.
[15] В том смысле, что именно они, как правило, будут использоваться при решении задач.
[16] В другой терминологии — *дельта-мера*, сосредоточенная в точке a.
[17] Более точно, $\mathrm{P}(\xi = a) = 1$.

5. *Геометрическое распределение* $\mathrm{Geom}(p)$ с параметром $0 < p < 1$:

$$\mathcal{P}_\xi : \quad \begin{pmatrix} 0 & 1 & \dots & k & \dots \\ p_0 & p_1 & \dots & p_k & \dots \end{pmatrix},$$

где $p_k = p\,(1-p)^k$, $k \geq 0$. Имеет смысл распределения числа неудач до первого успеха в испытаниях Бернулли с вероятностью p.

Иногда под геометрическим распределением подразумевают распределение полного числа испытаний до первого успеха. Тогда вероятности остаются теми же, а к значениям случайной величины прибавляется единица. У нас (по определению) геометрическое распределение всегда соответствует числу неудач.

6. Распределение Пуассона $\Pi(\lambda)$ с параметром $\lambda > 0$:

$$\mathcal{P}_\xi : \quad \begin{pmatrix} 0 & 1 & \dots & k & \dots \\ p_0 & p_1 & \dots & p_k & \dots \end{pmatrix},$$

где $p_k = \lambda^k\, e^{-\lambda}/k!$. Это очень распространенное распределение, одна из простейших интерпретаций[18] — (предельное) распределение числа успехов в очень большом числе испытаний Бернулли с очень маленькой вероятностью успеха.

7. Отрицательное биномиальное распределение $\mathrm{NB}(m,p)$ с параметрами $m \in \{1,2,\dots\}$ и $0 < p < 1$:

$$\mathcal{P}_\xi : \quad \begin{pmatrix} 0 & 1 & \dots & k & \dots \\ p_0 & p_1 & \dots & p_k & \dots \end{pmatrix},$$

где $p_k = C^k_{m+k-1}\, p^m\,(1-p)^k$. Смысл — распределение числа неудач до достижения m-го успеха.

Обратите внимание, что в последних трех примерах число значений случайных величин бесконечно (счетно).

Абсолютно непрерывные распределения.

Определение 3.2. Говорят, что случайная величина ξ имеет *абсолютно непрерывное*[19] *распределение*, если существует такая функция p_ξ, что для любого борелевского множества $B \in \mathcal{B}$ выполняется равенство

$$\mathrm{P}(\xi \in B) = \mathcal{P}_\xi(B) = \int_B p_\xi(x)dx. \tag{3.2}$$

Функция p_ξ при этом называется *плотностью распределения* случайной величины ξ.[20] А саму случайную величину ξ для краткости тоже называют абсолютно непрерывной.

Таким образом, в абсолютно непрерывном случае нам достаточно знать всего лишь одну функцию p_ξ, что бы иметь возможность восстановить все распределение.[21]

Отметим, что абсолютно непрерывное распределение не является дискретным, так как для него

$$\mathrm{P}(\xi = a) = \int_{\{a\}} p_\xi(dx) = 0$$

при любом $a \in \mathbb{R}$.

Перечислим общие свойства плотности.

[18] Вспомните предельную теорему Пуассона!

[19] сокращенно — а.н.

[20] Для тех, кто знает производные Радона-Никодима: плотность p_ξ является производной Р. -Н. распределения $\mathcal{P}_\xi$ по мере Лебега, условия существования плотности — это условия абсолютной непрерывности меры $\mathcal{P}_\xi$ по мере Лебега.

[21] А в дискретном случае достаточно знать таблицу распределения.

1. Если плотность распределения существует, то она единственна с точностью до множества Лебеговой меры ноль.[22]

2. Плотность является п. в. (по мере Лебега) неотрицательной функцией.[23]

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

   (для доказательства достаточно положить в (3.2) $B = \mathbb{R}$).

4. Если существует такая функция $\varphi(x)$, что для любых $a < b$

   $$\mathrm{P}(a < \xi < b) = \int_a^b \varphi(x)dx,$$

   то $\varphi(x)$ — плотность распределения случайной величины ξ.

5. Если существует такая функция $\varphi(x)$, что для любого $y \in \mathbb{R}$

   $$\mathrm{P}(\xi < y) = \int_{-\infty}^{y} \varphi(x)dx,$$

   то $\varphi(x)$ — плотность распределения случайной величины ξ.

Замечание 3.1. 1. Таким образом, у абсолютно непрерывной случайной величины существует не одна плотность распределения, а целый класс эквивалентных плотностей. При решении задач, конечно, из этого класса функций выбирается какая-то одна, наиболее удобная. Например, если существует непрерывный (или кусочно-непрерывный) вариант плотности распределения, но с ним и нужно иметь дело.
2. Последние два свойства говорят о том, что для проверки существования плотности не нужно проверять равенство (3.2) для всех борелевских множеств B. Достаточно в качестве B брать всевозможные интервалы или даже всевозможные полупрямые.
3. Пусть в нас есть неотрицательная измеримая функция p, заданная на прямой. Если при этом $\int_{\mathbb{R}} p(x)dx = 1$, то равенства

$$\mathcal{P}(B) = \int_B p(x)dx, \quad B \in \mathcal{B},$$

задают вероятностную меру на измеримом пространстве $(\mathbb{R}, \mathcal{B})$. Тем самым (см. Предложение 2.2 и Замечание 2.1) мы можем говорить о распределении с плотностью p, не имея в виду никакой конкретной случайной величины.[24]

Основные абсолютно непрерывные распределения.

1. *Равномерное распределение* $\mathrm{U}(a, b)$ на отрезке $[a, b]$, $-\infty < a < b < \infty$. Задается плотностью

   $$p_\xi(x) = \begin{cases} \dfrac{1}{b-a} & \text{при } a \le x \le b, \\ 0 & \text{иначе.} \end{cases} \tag{3.3}$$

   Смысл понятен: вероятность попасть случайной величине в отрезок длины ε, находящийся внутри $[a, b]$, не зависит от расположения этого отрезка (а только от его длины). Отметим, что значения плотности в точках a и b не существенны.

[22] Свойство понятно: если изменить плотность в нескольких точках, то все интегралы (3.2) останутся теми же.
[23] Иначе при каком-то B интеграл в правой части (3.2) был бы отрицательным.
[24] На самом деле аналогичное замечание может быть сделано и о дискретных распределениях с заменой плотности распределения на его таблицу.

2. *Показательное*[25] распределение $\mathrm{Exp}(\lambda)$ с параметром $\lambda > 0$. Плотность имеет вид

$$p_\xi(x) = \begin{cases} \lambda e^{-\lambda x} & \text{при } x \geq 0, \\ 0 & \text{иначе.} \end{cases} \tag{3.4}$$

Такое распределение часто имеют случайные величины типа «времени жизни» (в общем случае — типа «времени до наступления какого-то события»). В частности, время полного распада атома радиоактивного вещества имеет именно показательное распределение.

3. *Нормальное* (гауссовское) распределение $\mathbb{N}(a, \sigma^2)$ с параметрами $a \in \mathbb{R}$ и $\sigma^2 > 0$. Плотность следующая:[26]

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-a)^2/2\sigma^2}\ , \quad x \in \mathbb{R}. \tag{3.5}$$

Это очень важное и распространенное (как в теоретических, так и практических исследованиях) распределение, но о содержательном смысле его мы поговорим потом — здесь все не так просто.[27]

Иногда для нормального распределения использую обозначение $\mathbb{N}(a, \sigma)$. У нас второй параметр — σ^2, а не σ. Поэтому распределение $\mathrm{N}(0, 4)$ имеет плотность

$$p_\xi(x) = \frac{1}{2\sqrt{2\pi}}\, e^{-x^2/8}\ , \quad \text{а не} \ \ p_\xi(x) = \frac{1}{4\sqrt{2\pi}}\, e^{-x^2/32}\ .$$

Распределение $\mathrm{N}(0, 1)$ носит название *стандартного нормального.*

4 Функция распределения

Существует еще одна функция, полностью определяющая распределение.

Определение 4.1. Пусть ξ — случайная величина. Функция $F : \mathbb{R} \mapsto \mathbb{R}$, определяемая равенством

$$F_\xi(x) = \mathrm{P}(\xi < x) = \mathrm{P}\big(\xi \in (-\infty, x)\big) = \mathcal{P}_\xi\big((-\infty, x)\big) \tag{4.1}$$

называется *функцией распределения* случайной величины ξ.

Очевидно,[28] функция распределения однозначно определяется распределением. Верно и обратное.

Предложение 4.1. *Пусть ξ — некоторая случайная величина, $\mathcal{P}_\xi$ — ее распределение и F_ξ — функция распределения. Тогда $\mathcal{P}_\xi$ однозначно определяется по F_ξ.*

Таким образом, если мы знаем функцию распределения, то мы знаем все распределение.[29]

Замечание 4.1. В некоторых книгах (например, А.Н. Ширяев, Вероятность, разные издания) функция распределения определяется равенством, аналогичным (4.1), но в правой части стоит $\mathrm{P}(\xi \leq x)$. Это вопрос договоренности, привычки и удобства. У нас всегда используется знак «меньше».

Важно отметить, что функция распределения определяет распределение для любого типа распределений.

Свойства функции распределения.

[25] Другое название — *экспоненциальное.*
[26] Конечно, здесь σ — это положительный квадратный корень из σ^2.
[27] Впрочем, вспомните интегральную теорему Муавра-Лапласа!
[28] Действительно?
[29] Снова см. Предложение 2.2 и Замечание 2.1.

1. Общие свойства.

   - $0 \le F_\xi(x) \le 1$.[30]
   - $F_\xi(x) \to 0$ при $x \to -\infty$, $F_\xi(x) \to 1$ при $x \to +\infty$.[31]
   - F_ξ монотонно неубывает[32] и непрерывна слева.[33]

     Поэтому, если $F_\xi(x_0) = 0$, то $F_\xi(x) = 0$ при любом $x < x_0$. Аналогично, если $F_\xi(x_0) = 1$, то $F_\xi(x) = 1$ при любом $x > x_0$.
   - $\mathrm{P}(a \le \xi < b) = F_\xi(b) - F_\xi(a)$ и $\mathrm{P}(\xi \ge a) = 1 - F_\xi(a)$.

     Отсюда следует, что если случайная величина не принимает значений на отрезке $[a, b)$, то функция распределения постоянна на этом отрезке. Верно и обратное.
   - Поскольку функция распределения ограничена, она может иметь разрывы только первого рода. Так как она непрерывна слева, то величины таких разрывов («скачки») имеют вид $F_\xi(x+0) - F_\xi(x)$. Скачки функции распределения имеют вероятностный смысл: для любого $x \in \mathbb{R}$

     $$\mathrm{P}(\xi = x) = F_\xi(x+0) - F_\xi(x). \tag{4.2}$$

     Таким образом,

     a) для того, чтобы $\mathrm{P}(\xi = x) = 0$ необходимо и достаточно, чтобы функция распределения F_ξ была непрерывна в точке x;

     b) если функция распределения F_ξ имеет скачок в точке x, то величина этого скачка равна $\mathrm{P}(\xi = x)$, то есть ξ принимает значение x с положительной вероятностью, равной правой части (4.2).

Замечание 4.2. Можно доказать, что любая монотонно неубывающая непрерывная слева функция F, обладающая пределами $F(x) \to 0$ при $x \to -\infty$ и $F(x) \to 1$ при $x \to +\infty$, является функцией распределения некоторого распределения $\mathcal{P}$ в том смысле, что $\mathcal{P}\big((-\infty, x)\big) = F(x)$. Поэтому (снова см. Предложение 2.2 и Замечание 2.1) мы можем говорить о любой функции, обладающей перечисленными свойствами, как о некоторой функции распределения.

Далее — более специфические свойства функций распределения.

2. Функция распределения дискретных случайных величин.

   Пусть распределение дискретной случайной величины задано конечной[34] таблицей

   $$\mathcal{P}_\xi : \quad \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}, \tag{4.3}$$

   причем $x_1 < x_2 < \ldots < x_n$. Обозначим $s_i = p_1 + \ldots + p_i$. Тогда

   $$F_\xi(x) = \begin{cases} 0 & \text{при } x \le x_1, \\ s_i & \text{при } x_i < x \le x_{i+1}, \quad i = 1, \ldots, n-1, \\ 1 & \text{при } x > x_n. \end{cases} \tag{4.4}$$

[30] Очевидно, так как $F_\xi(x)$ — вероятность некоторого события.

[31] Это следуют из непрерывности меры снизу и сверху.

[32] так как при $x < y$ имеет место включение $\{\omega : \xi(\omega) < x\} \subset \{\omega : \xi(\omega) < y\}$.

[33] Снова по той же непрерывности меры сверху. Заметим, что если определять ф. р. с помощью знака $\le$, то функция распределения будет непрерывна справа.

[34] С бесконечными распределениями ситуация может быть более хитрой: ничего не мешает случайной величине принимать, например, все рациональные значения. У нас такой экзотики, конечно, не будет.

3. Функция распределения а. н. случайных величин.

- Функция распределения выражается через плотность следующим образом:

$$F_\xi(x) = \int_{-\infty}^{x} p_\xi(t)dt.$$

Отсюда сразу же следует, что для а. н. распределений функция распределения является непрерывной.

- Если существует такая функция φ, что для любого x

$$F_\xi(x) = \int_{-\infty}^{x} \varphi(t)dt,$$

то φ является плотностью распределения случайной величины ξ.[35]

- Так как в абсолютно непрерывном случае $\mathrm{P}(\xi = a) = \mathrm{P}(\xi = b) = 0$, то

$$\mathrm{P}(a \le \xi \le b) = \mathrm{P}(a \le \xi < b) = \mathrm{P}(a < \xi \le b) = \mathrm{P}(a < \xi < b) = F_\xi(b) - F_\xi(a).$$

- Если ξ абсолютно непрерывна с плотностью p_ξ, то при почти всех по мере Лебега $x \in \mathbb{R}$ $F'_\xi(x) = p_\xi(x)$.

Предупреждение: можно доказать, что любая функция распределения F при почти всех x имеет производную. Иначе говоря, равенство $F'_\xi(x) = f(x)$ для какой-то функции f и почти всех x выполняется для любой функции распределения (то есть не обязательно только для абсолютно непрерывных распределений).

Если при этом

$$\int_{\mathbb{R}} f(x)dx = 1,$$

то у распределения с функцией распределения F есть плотность, равная f. Если интеграл в левой части последнего неравенства не равен единице,[36] то распределение не является абсолютно непрерывным.

Например, пусть распределение дискретной случайной величины задается таблицей

$$\mathcal{P}_\xi : \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}.$$

Тогда, согласно (4.4), ее функция распределения имеет вид

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \le 0, \\ 1/2 & \text{при } 0 < x \le 1, \\ 1 & \text{при } x \ge 1. \end{cases}$$

Тем самым производная этой функции распределения равна 0 при всех x, кроме нуля и единицы (а этих точках производная не существует). Иначе говоря, при почти всех (по мере Лебега) x выполнено равенство $F'_\xi(x) = f(x) = 0$. Конечно $\int_{\mathbb{R}} f(x)dx = 0$, а не единице.

[35] Мы уже обсуждали это свойство, когда говорили об абсолютно непрерывных распределениях.

[36] На самом деле он всегда не превосходит 1.

Если $\xi \in \mathrm{U}(a,b)$ (см. формулу (3.3) для плотности этого распределения), то соответствующая функция распределения имеет вид

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \le a, \\ \dfrac{x-a}{b-a} & \text{при } a < x \le b, \\ 1 & \text{при } x \ge 1, \end{cases}$$

то есть эта функция является линейной на отрезке $[a,b]$, причем в точке a она равна нулю, а в точке b — единице.[37]

При $\xi \in \mathrm{EXP}(\lambda)$ (см. формулу (3.4))

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \le 0, \\ 1 - e^{-\lambda x} & \text{при } x \ge 0. \end{cases}$$

Что касается нормального распределения $\mathrm{N}(a,\sigma^2)$ (см. (3.5)), то в этом случае формула для функции распределения явно не выписывается.

5 Смеси распределений

Пусть $\mathcal{P}_1, \ldots, \mathcal{P}_n, \ldots$ — некоторые распределения, а $q_1, \ldots, q_n, \ldots$ — неотрицательные числа, сумма которых равна 1. Тогда

$$\mathcal{P} = \sum_{i=1}^{\infty} q_i \mathcal{P}_i \tag{5.1}$$

также является распределением и называется *смесью распределений $\mathcal{P}_i$ с весами q_i*.[38]

Если обозначить F_i функции распределения $\mathcal{P}_i$, то (5.1) перепишется в виде

$$F(x) = \sum_{i=1}^{\infty} q_i F_i(x),$$

где F является функцией распределения для $\mathcal{P}$.

Если же распределения $\mathcal{P}_i$ обладают плотностями p_i, то распределение $\mathcal{P}$ также является абсолютно непрерывным, причем для почти всех (по мере Лебега) значений $x \in \mathbb{R}$

$$p(x) = \sum_{i=1}^{\infty} q_i p_i(x),$$

где p — плотность распределения $\mathcal{P}$.

Замечание 5.1. Если взять $q_i = 0$ при $i > n$, то все эти определения и формулы автоматически перенесутся на конечное число распределений $\mathcal{P}_i$.

6 Как искать распределения

6.1 Общие соображения

Стандартная задача, которой мы будем заниматься, может быть сформулирована следующим образом.

[37] Обратите внимание на то, что здесь функция распределения не дифференцируема в точках a и b.

[38] В терминах математического анализа это, конечно, выпуклая линейная комбинация распределений $\mathcal{P}_i$.

Пусть случайная величина ξ обладает плотностью распределения p_ξ и φ — некоторая измеримая функция. Найти распределение случайной величины $\eta = \varphi(\xi)$.

Общий (неконструктивный) ход решения следующий. Пусть $B \in \mathcal{B}$. Тогда[39]

$$\mathrm{P}(\eta \in B) = \mathrm{P}(\varphi(\xi) \in B) = \mathrm{P}\left(\xi \in \varphi^{-1}(B)\right) = \int_{\varphi^{-1}(B)} p_\xi(z)dz = \int_{z:\,\varphi(z)\in B} p_\xi(z)dz. \tag{6.1}$$

Конечно, надеяться на вычисление правой части (6.1) для любого борелевского множества B глупо. Более разумно поступать следующим образом.

Если случайная величина $\eta = \varphi(\xi)$ дискретна (то есть, если она принимает конечное или счетное число значений y_i), то, взяв в (6.1) в качестве B одноточечное множество $\{y_i\}$, получим, что

$$\mathrm{P}(\eta = y_i) = \int_{z:\,\varphi(z)=y_i} p_\xi(z)dz. \tag{6.2}$$

Вычисляя такие интегралы, получим таблицу распределения случайной величины η.

Аналогично, для нахождения функции распределения F_η случайной величины η имеем

$$F_\eta(x) = \int_{t:\varphi(t)<x} p_\xi(z)dz, \tag{6.3}$$

то есть задача сводится к вычислению однократного интеграла, зависящего от параметра x.[40]

Конечно, формула (6.3) — это универсальное решение задачи о нахождении распределения случайной величины η. Однако, если у η есть плотность распределения p_η, то естественным ответом будет именно эта плотность. При этом возникает вопрос: если случайная величина ξ обладает плотностью распределения p_ξ, то для каких функций φ случайная величина $\eta = \varphi(\eta)$ тоже будет абсолютно непрерывной?

Для решения задач нам хватит следующего достаточного условия, накладываемого на функцию φ.

Предложение 6.1. *Пусть случайная величина ξ обладает плотностью p_ξ. Предположим, что существуют такие числа $-\infty < a_1 < \ldots < a_n < \infty$, что на любом промежутке (a_i, a_{i+1}), включая бесконечные промежутки $(-\infty, a_1)$ и (a_n, ∞), функция φ обратима и непрерывно дифференцируема. Тогда случайная величина $\eta = \varphi(\xi)$ является абсолютно непрерывной.*

Грубо говоря, условие Предложения 6.1 говорит, что функция φ должна быть кусочно гладкой и не должна иметь ступенек.[41] В этих условиях нахождение плотности случайной величины η может быть осуществлено вычислением $F_\eta(x)$ с последующим дифференцированием. Можно также попытаться непосредственно продифференцировать правую часть (6.3), используя аналоги формулы Барроу.

Приведем простейший пример.

Пример 6.1. Случайная величина ξ имеет плотность $p_\xi(x)$. Найти распределение случайной величины $\eta = a\xi + b$, $a > 0$.

[39] Здесь $\varphi^{-1}(B)$ — полный прообраз множества B при отображении φ, то есть $\varphi^{-1}(B) = \{x : \varphi(x) \in B\}$. Не путайте это обозначение с обратной функцией: φ совершенно не обязательно должна быть взаимно-однозначной!

[40] Конечно, формула (6.3) годится и для дискретной случайной величины η. Однако при решении задач формула (6.2) для этого случая гораздо предпочтительней. Во-первых, для дискретных случайных величин таблица распределения — гораздо более наглядный объект, определяющий само распределение, чем функция распределения — в таблице прямо указано, какие значения и с какими вероятностями принимает случайная величина η. Во-вторых, поскольку интегралы (6.3) должны быть сосчитаны при всех x, это может быть источником никому не нужных ошибок.

[41] Если $\varphi(x) = c$ при $a < x < b$, то $\mathrm{P}(\eta = c) \geq \mathrm{P}(a < \xi < b)$, а последняя вероятность будет положительна, если только p_ξ не является почти всюду тождественным нулем на (a, b).

Решение. Найдем сначала функцию распределения случайной величины ξ. Это легко:

$$F_\eta(x) = \mathrm{P}(\eta < x) = \mathrm{P}(a\xi + b < x) = \mathrm{P}(\xi < (x-b)/a) = F_\xi\big((x-b)/a\big).$$

В терминах функций распределения задача решена: неизвестную функцию распределения F_η выразили через известную F_ξ, при этом существование плотности p_ξ нам не потребовалось. Но, согласно Предложению 6.1, у случайной величины η есть плотность распределения. Поэтому

$$p_\eta(x) = \frac{d}{dx}\,F_\eta(x) = \frac{d}{dx}\,F_\xi\left(\frac{x-b}{a}\right) = \frac{1}{a}\,p_\xi\left(\frac{x-b}{a}\right).$$

А можно и не пользоваться Предложением 6.1. Действительно, сделав замену переменных $t = (s-b)/a$, получим, что

$$F_\eta(x) = F_\xi((x-b)/a) = \int_{-\infty}^{(x-b)/a} p_\xi(t)dt = \frac{1}{a}\int_{-\infty}^{x} p_\xi\big((s-b)/a\big)ds.$$

Следовательно (см. свойства плотности) функция $p_\eta(s) = a^{-1}p_\xi\big((s-b)/a\big)$ является плотностью распределения случайной величины η.

Конечно, эта задача оказалась столь простой только потому, что линейная функция $\phi(x) = ax + b$ является обратимой и обратная к ней является очень простой. Если это не так, то все вычисления оказываются более сложными.

6.2 Некоторые приемы

Перечислим некоторые приемы, облегчающие задачу нахождения распределения случайной величины $\eta = \varphi(\xi)$ в случае, когда ξ имеет плотность распределения p_ξ. Еще раз отметим, что слова «найти распределение случайной величины η» означают для нас, что ответ выдается в виде таблицы распределения, если η дискретна, в виде плотности распределения p_η, если η абсолютно непрерывна (это желательно), и в общем случае — в виде функции распределения F_η.

1. Если случайная величина η принимает конечное или счетное число значений t_i, то нужно решать уравнения $\varphi(z) = t_i$ и, получив ответ в виде некоторых подмножеств $z \in D_i \subset \mathbb{R}$, вычислить вероятности

$$p_i = \mathrm{P}(\xi \in D_i) = \int_{D_i} p_\xi(z)dz.$$

2. В остальных случаях считается функция распределения. Для этого нужно решить при всех x неравенство $\varphi(z) < x$, получить ответ в виде подмножеств $z \in B_x \subset \mathbb{R}$ и проинтегрировать плотность p_η по этим множествам:[42]

$$F_\eta(x) = \mathrm{P}(B_x) = \int_{B_x} p_\xi(z)dz.$$

Здесь могут помочь следующие соображения.

- Полезно определить сначала, какие значения принимает случайная величина η. Если, например, $a < \eta < b$, то автоматически $F_\eta(x) = 0$ при $x \le a$ и $F_\eta(x) = 1$ при $x > b$. Значит, решать неравенства $\varphi(z) < x$ при таких x не нужно.
- Решение неравенств $\varphi(z) < x$ можно контролировать графически: если найти решения уравнения $\varphi(z) = x$, то, используя график функции φ, несложно понять структуру множества B_x.

[42] В принципе, такие интегралы могут и не считаться в явном виде.

- После того, как функция распределения вычислена, следует проверить результат на правдоподобность. А именно, нужно проверить, удовлетворяет ли Ваш ответ основным свойствам функции распределения:
  a) лежат ли значения F_η между нулем и единицей?
  b) какие значения F_η принимает (в пределе) на $\pm\infty$?
  c) является ли F_η монотонно неубывающей?
  d) должны ли у этой функции быть скачки? Если «да», то правильно ли сосчитаны их значения и там ли они находятся, где нужно? Нет ли у Вашей функции распределения лишних скачков? Является ли Ваша функция распределения непрерывной именно слева в точках этих скачков?
  e) там, где случайная величина η не принимает значений, ее функция распределения должна быть постоянной. Выполнено ли у Вас это свойство?

  Такая проверка на правдоподобность почти гарантирует правильность ответа.
- Наконец, когда функция распределения найдена, нужно дифференцированием найти плотность распределения, если, конечно, она существует.

6.3 Примеры

Решим, используя эти соображения, пару простых примеров.

Задача 1. Пусть ξ имеет плотность распределения

$$p_\xi(x) = \begin{cases} 0 & \text{при } x \le 0, \\ 2x & \text{при } 0 < x \le 1, \\ 1 & \text{при } x > 1. \end{cases}$$

Нужно найти распределение случайной величины $\eta = \min(\xi, 1-\xi)$.

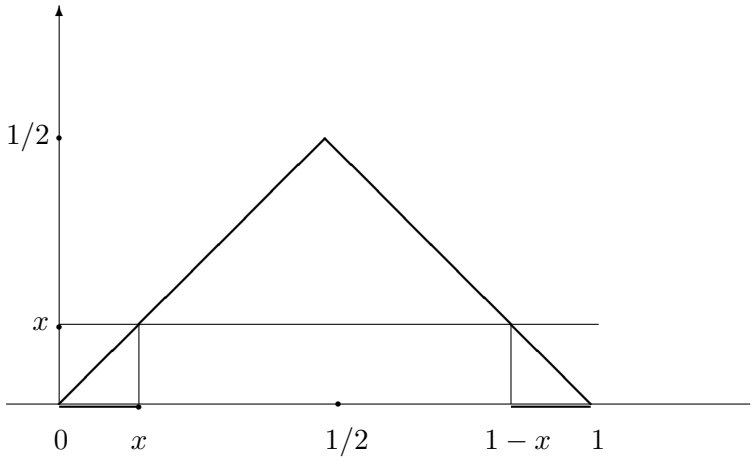

Рис. 1: График функции $\min(z, 1-z)$ при $z \in [0,1]$.

Решение 1. Прежде всего,[43]

$$F_\xi(x) = \begin{cases} 0 & \text{при } x \le 0, \\ x^2 & \text{при } 0 < x \le 1, \\ 1 & \text{при } x > 1. \end{cases} \tag{6.4}$$

Далее, видно, что случайная величина ξ принимает значения от 0 до 1. Поэтому η принимает значения между нулем и 1/2 (см. график функции $\varphi(z) = \min(z, 1-z)$ при $z \in [0,1]$ на рис. 1). Значит (здесь используется непрерывность функции распределения слева), $F_\eta(x) = 0$ при $x \le 0$ и $F_\eta(x) = 1$ при $x > 1/2$. Поэтому остается рассмотреть случай $0 < x \le 1/2$.

При таких x нужно найти вероятности

$$F_\eta(x) = \mathrm{P}\big(\min(\xi, 1-\xi) < x\big) = \mathrm{P}(\xi \in A_x) = \int_{A_x} p_\xi(t)dt,$$

где $A_x = \{t \in \mathbb{R} : \min(t, 1-t) < x\}$. Поскольку $p_\xi(t) = 0$ при $t \notin [0,1]$, то на самом деле нам нужно сосчитать вероятность

$$F_\eta(x) = \int_{B_x} p_\xi(t)dt,$$

где $B_x = \{t \in [0,1] : \min(t, 1-t) < x\}$.

[43] Проверьте и заодно убедитесь, что при задании $p_\xi(x)$ нет ошибок!

Найдем теперь множество B_x в явном виде. Нетрудно понять, что B_x является объединением двух непересекающихся интервалов (эти интервалы отмечены жирным на оси абсцисс рис. 1):

$$B_x = [0, x) \cup (1 - x, 1].$$

Поэтому при $0 < x \leq 1/2$ (см. формулу (6.4))

$$F_\eta(x) = \int_0^x p_\xi(t)dt + \int_{1-x}^1 p_\xi(t)dt = F_\xi(x) - F_\xi(0) + F_\xi(1) - F_\xi(1-x) = x^2 + 1 - (1-x)^2 = 2x.$$

Тем самым

$$F_\eta(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ 2x & \text{при } 0 < x \leq 1/2, \\ 1 & \text{при } x > 1/2. \end{cases} \tag{6.5}$$

Поскольку (см. Предложение 6.1) у случайной величины η существует плотность распределения, то нужно ее найти. Получим (дифференцируя $F_\eta(x)$), что

$$p_\eta(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ 2 & \text{при } 0 < x \leq 1/2, \\ 0 & \text{при } x > 1/2. \end{cases}$$

Тем самым $\eta \in \mathrm{U}(0, 1/2)$. □

Задача 2. Пусть $\xi \in \mathrm{U}(0,1)$. Требуется найти распределение случайной величины $\eta = \varphi(\xi)$, где

$$\varphi(z) = \begin{cases} z & \text{при } z \le 1/2, \\ 1/4 & \text{при } z > 1/2. \end{cases} \tag{6.6}$$

Решение 2. Поскольку $\xi \in [0,1]$ с вероятностью 1 (то есть $F_\xi(z) = 0$ при $z \le 0$ и $F_\xi(z) = 1$ при $z > 1$), то достаточно рассматривать функцию φ только на отрезке $[0,1]$. Поэтому $\eta \in [0,1/2]$ и при вычислении $F_\eta(x)$ достаточно рассматривать $x \in (0,1/2]$. Как и раньше,

$$F_\eta(x) = \int_{B_x} p_\xi(t)dt,$$

где $B_x = \{t \in [0,1] : \varphi(t) < x\}$, а φ определена в (6.6). Тем самым нам снова нужно явно выписать множество B_x.

Оказывается, что в зависимости от величины x множество B_x имеет различную структуру. А именно, $B_x = [0,x)$ при $0 < x \le 1/4$ (см. рис. 2, где снова множество B_x выделено жирным на оси абсцисс).

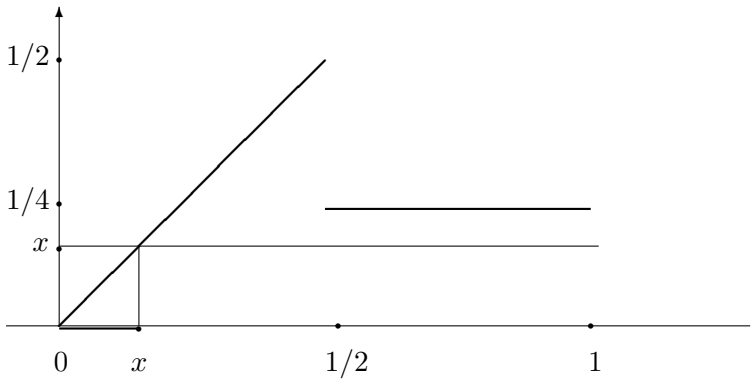

Рис. 2: График функции (6.6) при $z \in [0,1]$ и множество B_x при $x \le 1/4$.

При $x > 1/4$ множество B_x выглядит по-другому: $B_x = [0,x) \cup [1/2,1]$, это хорошо видно на рис. 3.

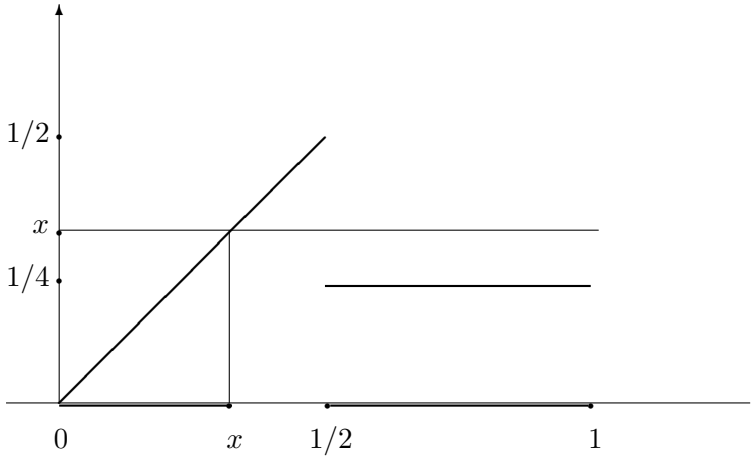

Рис. 3: График функции (6.6) при $z \in [0,1]$ и множество B_x при $x > 1/4$.

Следовательно,[44] при $0 < x \le 1/4$

$$F_\eta(x) = \int_0^x p_\xi(t)dt = x,$$

а при $1/4 < x \le 1/2$

$$F_\eta(x) = \int_0^x p_\xi(t)dt + \int_{1/2}^1 p_\xi(t)dt = x + 1/2.$$

Окончательно

$$F_\eta(x) = \begin{cases} 0 & \text{при } x \le 0, \\ x & \text{при } 0 < x \le 1/4, \\ x + 1/2 & \text{при } 1/4 < x \le 1/2, \\ 1 & \text{при } x > 1/2. \end{cases} \tag{6.7}$$

Этот результат имеет смысл проанализировать. Конечно, F_η удовлетворяет всем стандартным свойствам функций распределения.[45] Но она имеет скачок в точке $1/4$ и величина этого скачка равна $1/2$. Похоже ли это на истину?

Их (6.6) видно, что $\eta = 1/4$ при $\xi \in (1/2, 1]$. Так как $\xi \in \mathrm{U}(0, 1)$, то отсюда следует,[46] что $\mathrm{P}(\eta = 1/4) = \mathrm{P}(\xi \in (1/2, 1]) = 1/2$, что в точности соответствует равенству (4.2).

Так как полученная функция распределения имеет скачок, то плотности распределения у случайной величины η не существует, и поэтому равенство (6.7) является в этой задаче окончательным ответом. □

[44] Так как $p_\xi(t) = 1$ при $t \in [0, 1]$

[45] То есть она монотонно не убывает, непрерывна слева и пр.

[46] $\eta = 1/4$ еще и при $\xi = 1/4$, но такое событие происходит с нулевой вероятностью.