Представление распределений в виде смеси распределений Бернулли

А.А.Новоселов*

Лекция для студентов Института математики СФУ

Аннотация

В лекции рассматривается представление произвольного распределения с нулевым средним в виде смеси распределений Бернулли с нулевым средним. Разложения такого вида являются важным инструментом анализа, и могут использоваться для исследования неприятия риска [1].

Содержание

1 Введение

Представление сложных распределений в виде смеси более простых распределений представляет собой полезный инструмент для решения теоретико-вероятностных задач. В частности, представление произвольного распределения с нулевым средним в виде смеси распределений Бернулли с нулевым средним может использоваться при

*Сибирский Федеральный Университет, Свободный пр. 79, 660041, Красноярск, e-mail: arcady@novosyolov.com

вычислении неприятия риска в различных моделях предпочтений [1]. В настоящей лекции приводятся представления такого рода.

2 Основная теорема

Пусть W_z – функция распределения, вырожденного в точке $z \in \mathbf{R}$:

$$W_z(x) = \begin{cases} 0, & x < z, \\ 1, & x \geq z. \end{cases} \tag{1}$$

Тогда нетрудно вывести следующее представление.

Теорема 2.1 *Пусть F – произвольная вещественная функция распределения. Тогда справедливо представление*

$$F(x) = \int_{-\infty}^{\infty} W_z(x)\, dF(z), \quad x \in \mathbf{R}. \tag{2}$$

Замечание 2.1 *В теореме утверждается, что произвольная функция распределения F может быть представлена в виде смеси вырожденных функций распределения, причем в качестве смешивающей функции распределения используется сама F.*

Доказательство. Действительно, используя (1), получаем

$$\int_{-\infty}^{\infty} W_z(x)\, dF(z) = \int_{-\infty}^{x} dF(z) = F(x),$$

что и требовалось. $\diamond$

Пусть $L^1(\Omega, \mathcal{A}, \mathbf{P})$ – вероятностное пространство, $L^1(\Omega, \mathcal{A}, \mathbf{P})$ – пространство случайных величин с конечным математическим ожиданием, $\mathcal{F}$ – соответствующая совокупность функций распределения на $\mathbf{R}$ (то есть, функций распределения, для которых интеграл $\mathbf{E}F = \int x\, dF(x)$ конечен), и

$$\mathcal{F}_0 = \{F \in \mathcal{F} : \mathbf{E}F = 0\} -$$

совокупность функций распределения с нулевым средним. Представляет интерес аналогичное представление функций из $\mathcal{F}_0$ в виде смеси возможно более простых распределений также из $\mathcal{F}_0$. В настоящей работе мы получим такое представление в виде смеси распределений Бернулли

$$B_{a,b,p}(x) = \begin{cases} 0, & x < a, \\ 1 - p, & a \leq x < b, \\ 1, & b \leq x,. \end{cases} \tag{3}$$

где $a \leq b$ – атомы распределения, имеющие веса $1 - p$ и p, соответственно (при $a = b$ распределение Бернулли превращается в вырожденное распределение). Для $B_{a,b,p} \in \mathcal{F}_0$ имеем $\mathbf{E}B_{a,b,p} = a + (b - a)p = 0$, так что с необходимостью $a < 0 < b$, и параметр p однозначно определяется по значениям параметров a и b в виде $p = -a/(b - a)$. Поэтому распределения Бернулли из класса $\mathcal{F}_0$ будем обозначать $B_{a,b}$.

Основным результатом работы является следующая

Теорема 2.2 *Пусть $F \in \mathcal{F}_0$. Тогда найдется однопараметрическое семейство распределений Бернулли $B_u \in \mathcal{F}_0$, $u \in \mathcal{U} = [0, 1 - F(0+)]$ и функция распределения G на* $\mathbf{R}$ *такая, что*

$$F(x) = \int_{-\infty}^{\infty} B_u(x)\, dG(u), \quad x \in \mathbf{R}. \tag{4}$$

3 Доказательство теоремы 2.2

Пусть $F \in \mathcal{F}_0$, а X – некоторая случайная величина с функцией распределения F. Определяя, как обычно, случайные величины $X^+ = \max\{X, 0\}$, $X^- = \max\{-X, 0\}$, имеем $X = X^+ - X^-$ и, в частности, $\mathbf{E}X = \mathbf{E}X^+ - \mathbf{E}X^- = 0$.

Определим интервал $\mathcal{V} = [0, F(0-)]$, и рассмотрим функции

$$g(u) = \int_{F(0+)}^{F(0+)+u} F^{-1}(t)\, dt,\ u \in \mathcal{U}, \quad h(v) = \int_{F(0-)}^{F(0-)-v} F^{-1}(t)\, dt,\ v \in \mathcal{V},$$

которые представляют собой площади заштрихованных фигур, показанных на рисунке 1. Ясно, что функции g и h являются строго возрастающими и непрерывными, $g(0) = h(0) = 0$, $g(1 - F(0+)) = h(F(0-)) = \mathbf{E}X^+ = \mathbf{E}X^-$.

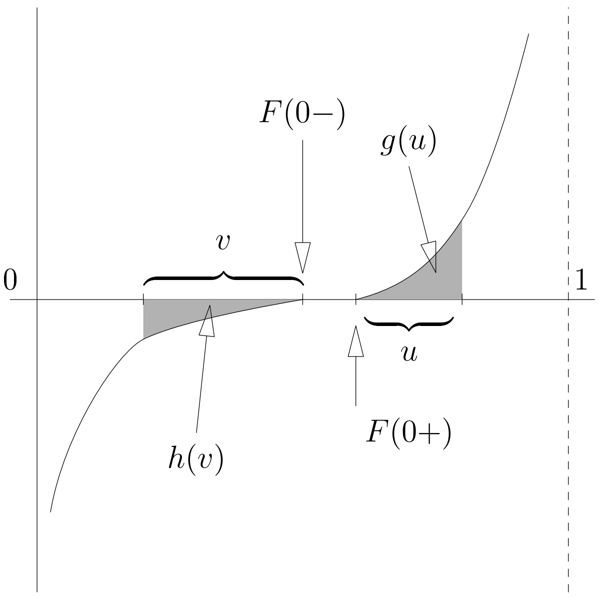

Рис. 1: Определение функций g, h

В частности, существует функция h^{-1}, обратная к h, и можно определить функцию $f(u) = h^{-1}(g(u))$, $u \in \mathcal{U}$. Функция f строго возрастает, непрерывна, $f(0) = 0$ и $f(1 - F(0+)) = F(0-)$. Более того, производные функций g, h, f, существующие почти всюду, вычисляются по формулам

$$g'(u) = F^{-1}(F(0+) + u), \quad h'(v) = -F^{-1}(F(0-) - v), \quad f'(u) = -\frac{F^{-1}(F(0+) + u)}{F^{-1}(F(0-) - f(u))}.$$

Определим еще функции

$$a(u) = -h'(f(u)), \quad b(u) = g'(u), \tag{5}$$

и отметим, что

$$f'(u) = \frac{g'(u)}{h'(f(u))} = -\frac{b(u)}{a(u)} \tag{6}$$

и

$$a(1 - F(0+)) = ess \inf X, \quad b(1 - F(0+)) = ess \sup X.$$

При $u \in (0, 1 - F(0+)]$ зададим распределение Бернулли B_u следующим образом:

$$B_u(x) = \begin{cases} 0, & x < a(u), \\ b(u)/(b(u) - a(u)), & a(u) \le x < b(u), \\ 1, & x \ge b(u). \end{cases} \tag{7}$$

В качестве B_0 будем использовать предельный случай: распределение, вырожденное в 0:

$$B_0(x) = \begin{cases} 0, & x < 0, \\ 1, & x \ge 0. \end{cases} \tag{8}$$

Ясно, что $\mathbf{E}B_u = 0$ при всех $u \in \mathcal{U}$, то есть, $B_u \in \mathcal{F}_0$, $u \in \mathcal{U}$.

Далее, определим функцию распределения G следующим образом:

$$G(u) = \begin{cases} 0, & u < 0, \\ F(0+) - F(0-) + u + f(u), & u \in \mathcal{U}, \\ 1, & u > 1 - F(0+). \end{cases} \tag{9}$$

Эта функция распределения имеет единственный скачок в точке $u = 0$, равный $F(0+) - F(0-)$ (скачок положителен, если исходная функция распределения имеет разрыв в 0), и непрерывна во всех остальных точках $\mathbf{R}$.

Параметрическое множество $\mathcal{U} = [0, 1 - F(0+)]$, семейство бернуллиевских распределений (7), (8) и функция распределения (9) определяют все элементы представления (4). Проверим это представление.

Прежде всего, заметим, что (4) преобразуется к виду

$$\begin{aligned} F(x) &= (F(0+) - F(0-))B_0(x) + \int_0^{1-F(0+)} B_u(x)\, dG(u) \\ &= (F(0+) - F(0-))B_0(x) + \int_0^{1-F(0+)} B_u(x)(1 + f'(u))\, du \\ &= (F(0+) - F(0-))B_0(x) + \int_0^{1-F(0+)} B_u(x)\, \frac{a(u) - b(u)}{a(u)}\, du \end{aligned}$$

Пусть сначала $x < 0$. Тогда

$$\begin{aligned} F(x) &= \int_{f^{-1}(F(0-)-F(x))}^{1-F(0+)} \frac{b(u)}{b(u) - a(u)}\, \frac{a(u) - b(u)}{a(u)}\, du \\ &= \int_{f^{-1}(F(0-)-F(x))}^{1-F(0+)} f'(u) du \\ &= f(1 - F(0+)) - f(f^{-1}(F(0-) - F(x))) \\ &= F(0-) - (F(0-) - F(x)) \\ &= F(x) \end{aligned}$$

Теперь рассмотрим случай $x \geq 0$. При этом

$$\begin{aligned}
F(x) &= (F(0+) - F(0-)) + \int_0^{F(x)-F(0-)} (1 + f'(u))\, du + \int_{F(x)-F(0-)}^{1-F(0+)} f'(u)\, du \\
&= (F(0+) - F(0-)) + \int_0^{F(x)-F(0-)} du + \int_0^{1-F(0+)} f'(u)\, du \\
&= (F(0+) - F(0-)) + (F(x) - F(0-)) - (F(0+) - F(0-)) \\
&\quad + f(1 - F(0+)) - f(0) \\
&= (F(0+) - F(0-)) + F(x) - F(0+) + F(0-) \\
&= F(x).
\end{aligned}$$

Теорема доказана.

4 Примеры

4.1 Симметричное непрерывное распределение

Пусть функция распределения F непрерывна и симметрична, то есть, $F(x) = 1 - F(-x)$, $x \in \mathbf{R}$. Тогда $F(0) = 1/2, \mathcal{U} = [0, 1/2]$,

$$h(u) \equiv g(u), \ \ f(u) = u, \ \ G(u) = 2u, \, u \in \mathcal{U},$$

$b(u) = -a(u) = F^{-1}(u)$, $u \in \mathcal{U}$ и

$$B_u(x) = \begin{cases} 0, & x < -F^{-1}(u), \\ 1/2, & -F^{-1}(u) \leq x < F^{-1}(u), \\ 1, & x \geq F^{-1}(u) \end{cases}$$

4.2 Дискретное распределение без атома в 0

Рассмотрим случайную величину X, принимающую значения -2, -1, 1 и 3 с вероятностями $1/4$, $3/8$, $1/8$ и $1/4$, соответственно. Функция распределения F этой случайной величины имеет вид

$$F(x) = \begin{cases} 0, & x < -2, \\ 1/4, & -2 \leq x < -1, \\ 5/8, & -1 \leq x < 1, \\ 3/4, & 1 \leq x < 3, \\ 1, & x \geq 3, \end{cases}$$

а ее график приведен на рисунке 2.

Здесь $F(0+) = F(0-) = F(0) = 5/8$, множество значений параметра $\mathcal{U} = [0, 3/8]$, а функции g, h имеют вид

$$g(u) = \begin{cases} u, & u \in [0, 1/8], \\ 3u - 1/4, & u \in [1/8, 3/8], \end{cases} \quad h(v) = \begin{cases} v, & v \in [0, 3/8], \\ 2v - 3/8, & v \in [3/8, 5/8]. \end{cases} \tag{10}$$

Отсюда

$$f(u) = h^{-1}(g(u)) = \begin{cases} u, & u \in [0, 1/8], \\ 3u - 1/4, & u \in [1/8, 5/24], \\ 3u/2 + 1/16, & u \in [5/24, 3/8]. \end{cases} \tag{11}$$

Графики функций g, h, f показаны на рисунке 3.

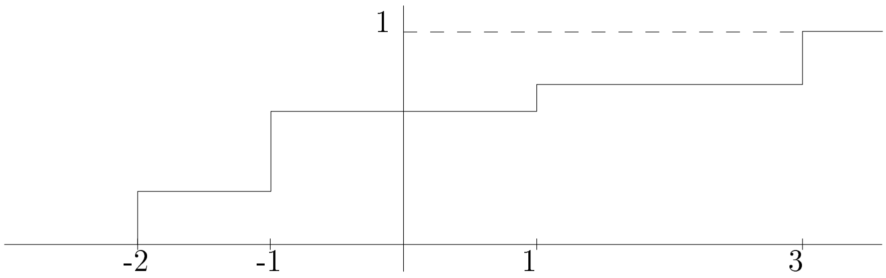

Рис. 2: График функции распределения F

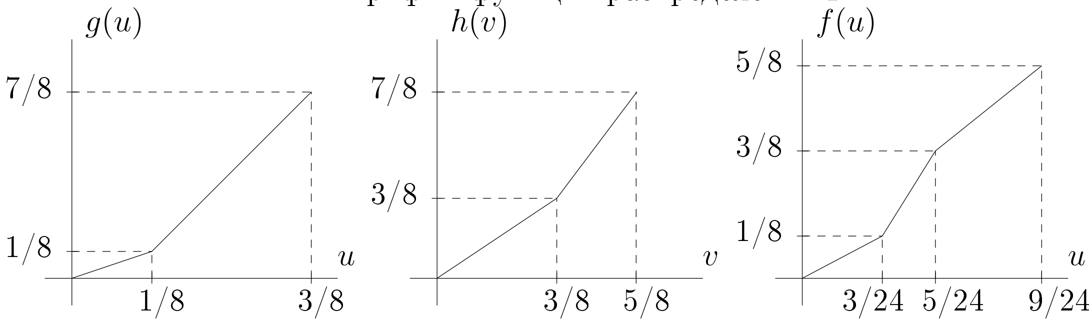

Рис. 3: Графики функций g, h, f

Функции a, b и G имеют вид

$$a(u) = \begin{cases} -1, & u \in [0, 5/24], \\ -2, & u \in (5/24, 3/8], \end{cases} \quad b(u) = \begin{cases} 1, & u \in [0, 1/8], \\ 3, & u \in (1/8, 3/8], \end{cases}$$

$$G(u) = u + f(u) = \begin{cases} 2u, & u \in [0, 1/8], \\ 4u - 1/4, & u \in (1/8, 5/24], \\ 5u + 1/16, & u \in [5/24, 3/8]. \end{cases}$$

Другими словами представление можно описать так. Обозначим $F^{(1)} = B_{-1,1}$, $F^{(2)} = B_{-1,3}$, $F^{(3)} = B_{-2,3}$ и

$$\lambda_1 = \int_0^{1/8} dG(u) = 1/4, \quad \lambda_2 = \int_{1/8}^{5/24} dG(u) = 1/3, \quad \lambda_3 = \int_{5/24}^{3/8} dG(u) = 5/12;$$

отметим, что $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Тогда исходная функция распределения имеет вид смеси бернуллиевских распределений

$$F(x) = \lambda_1 F^{(1)}(x) + \lambda_2 F^{(2)}(x) + \lambda_3 F^{(3)}(x), \quad x \in \mathbf{R}.$$

4.3 Пример с атомом в 0

Пусть теперь случайная величина X принимает значения $-3, -1, 0$ и 2 с вероятностью $1/6$ каждое, а значение 1 – с вероятностью $1/3$. Тогда ее функция распределения

F имеет вид

$$F(x) = \begin{cases} 0, & x < -3, \\ 1/6, & -3 \le x < -1, \\ 1/3, & -1 \le x < 0, \\ 1/2, & 0 \le x < 1, \\ 5/6, & 1 \le x < 2, \\ 1, & x \ge 2, \end{cases}$$

а ее график приведен на рисунке 4.

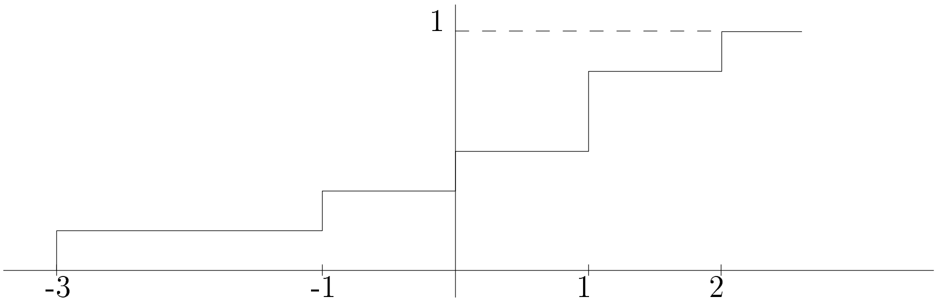

Рис. 4: График функции распределения F с атомом в 0

Здесь $F(0+) = 1/2,\ F(0-) = 1/3$, множество значений параметра $\mathcal{U} = [0, 1/2]$, а функции g, h имеют вид

$$g(u) = \begin{cases} u, & u \in [0, 1/3], \\ 2u - 1/3, & u \in [1/3, 1/2], \end{cases} \quad h(v) = \begin{cases} v, & v \in [0, 1/6], \\ 3v - 1/3, & v \in [1/6, 1/3]. \end{cases} \tag{12}$$

Отсюда

$$f(u) = h^{-1}(g(u)) = \begin{cases} u, & u \in [0, 1/6], \\ u/3 + 1/9, & u \in [1/6, 1/3], \\ 2u/3, & u \in [1/3, 1/2]. \end{cases} \tag{13}$$

Графики функций g, h, f показаны на рисунке 5.

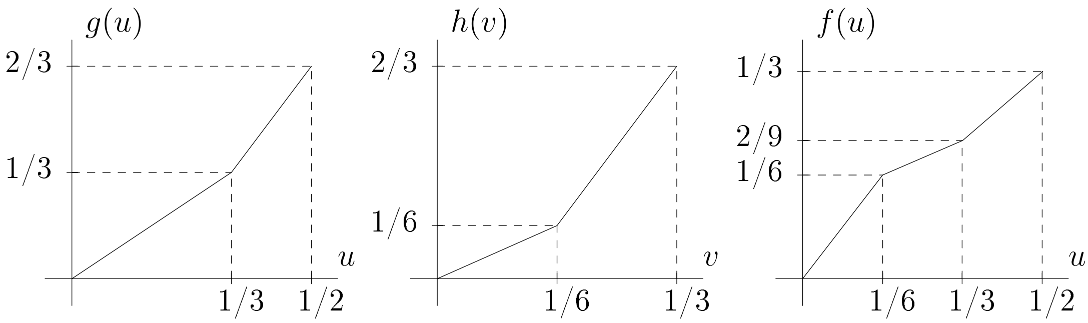

Рис. 5: Графики функций g, h, f при наличии атома в 0

Функции a, b имеют вид

$$a(u) = \begin{cases} -1, & u \in [0, 1/6], \\ -3, & u \in (1/6, 1/2], \end{cases} \quad b(u) = \begin{cases} 1, & u \in [0, 1/3], \\ 2, & u \in (1/3, 1/2]. \end{cases}$$

Функция G имеет в 0 скачок размера 1/6, и при неотрицательных значениях аргумента задается выражениями

$$G(u) = u + f(u) + 1/6 = \begin{cases} 2u + 1/6, & u \in [0, 1/6], \\ 4u/3 + 5/18, & u \in [1/6, 1/3], \\ 5u/3 + 1/6, & u \in [1/3, 1/2]. \end{cases}$$

Другими словами представление можно описать так. Обозначим $F^{(0)} = W_0$, $F^{(1)} = B_{-1,1}$, $F^{(2)} = B_{-3,1}$, $F^{(3)} = B_{-3,2}$ и

$$\lambda_0 = \int_{0-}^{0} dG(u) = 1/6, \ \ \lambda_1 = \int_0^{1/6} dG(u) = 1/3,$$

$$\lambda_2 = \int_{1/6}^{1/3} dG(u) = 2/9, \ \ \lambda_3 = \int_{1/3}^{1/2} dG(u) = 5/10;$$

отметим, что $\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 = 1$. Тогда исходная функция распределения имеет вид смеси бернуллиевских распределений и вырожденного распределения

$$F(x) = \lambda_0 F^{(0)}(x) + \lambda_1 F^{(1)}(x) + \lambda_2 F^{(2)}(x) + \lambda_3 F^{(3)}(x), \ \ x \in \mathbf{R}.$$

4.4 Непрерывный пример

Пусть теперь случайная величина X имеет непрерывное распределение с функцией распределения F, заданной посредством

$$F(x) = \begin{cases} (2/3)(x+1), & x \in [-1, 0], \\ (1/6)(x+4), & x \in [0, 2]. \end{cases}$$

График этой функции распределения приведен на рисунке.

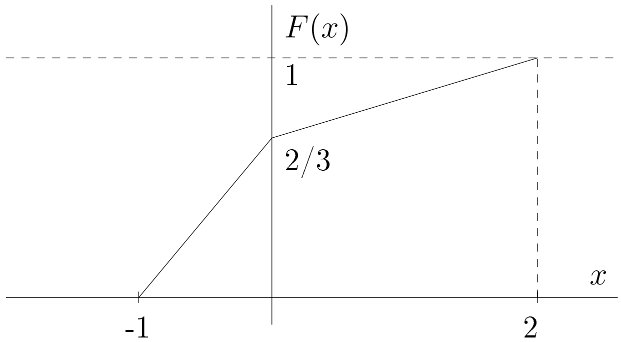

Рис. 6: График непрерывной функции распределения

Здесь $\mathcal{U} = [0, 1/3]$, а выражения для функций g, h, f имеют вид:

$$g(u) = 3u^2,\ u \in [0, 1/3], \ \ h(v) = 3v^2/4,\ v \in [0, 2/3],$$

$$f(u) = h^{-1}(g(u)) = 2u,\ u \in [0, 1/3].$$

Ясно также, что $a(u) = -3u$, $b(u) = 6u$, $G(u) = u + f(u) = 3u$, $\ u \in \mathcal{U}$. Тривиальная проверка подтверждает справедливость равенства

$$F(x) = \int_0^{1/3} B_u(x)\, dG(u), \ \ x \in \mathbf{R}$$

и в этом примере.

5 Непрерывный вариант

Если исходная функция распределения непрерывна, то возможно и другое представление в виде смеси, которое приведено в следующей теореме.

Теорема 5.1 *Пусть $F \in \mathcal{F}_0$ непрерывна. Тогда найдется однопараметрическое семейство распределений Бернулли $B_z \in \mathcal{F}_0$, $z \geq 0$ и функция распределения G такая, что $G(z) = 0$ при $z < 0$ и*

$$F(x) = \int_0^\infty B_z(x)\, dG(z), \quad x \in \mathbf{R}. \tag{14}$$

Доказательство этой теоремы предоставляется читателю в качестве упражнения 6.1.

6 Упражнения

Упражнение 6.1 *Доказать теорему 5.1.*

Упражнение 6.2 *Почему теорема 5.1 неприменима к дискретным распределениям?*

Список литературы

[1] Новоселов А.А. (2001) *Математическое моделирование финансовых рисков: теория измерения.* Наука: Новосибирск, 102 с.