§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин

ПАВЕЛ Е. РЯБОВ[1]

[1]Финансовый университет при Правительстве Российской Федерации

PERyabov@fa.ru

19 марта 2024 г.

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

§4. Случайные величины. Функция распределения и ее свойства. Числовые характеристики дискретных случайных величин.

План лекции

- Понятие случайной величины. Примеры.
- Функция распределения и ее свойства.
- Дискретные случайные величины. Независимость случайных величин.
- Числовые характеристики дискретных случайных величин: математическое ожидание
- Числовые характеристики дискретных случайных величин: дисперсия
- Ковариация случайных величин, коэффициент корреляции
- Ковариационная матрица и ее неотрицательная определенность

Случайные величины

Одним из важнейших понятий теории вероятностей является понятие случайной величины.

Примеры

1. Игральная кость бросается один раз. Пусть X — выпавшее число очков.
2. Покупается n лотерейных билетов. X — число выигрышей.

Определение

Пусть $(\Omega, \mathscr{F}, \mathbb{P})$ — вероятностное пространство. ***Случайной величиной*** *X называется однозначная действительная функция $X = X(\omega)$, определенная на пространстве элементарных исходов Ω, т.е. $X : \Omega \mapsto \mathbb{R}$, и такая, что для каждого действительного числа $x \in \mathbb{R}$ множество вида $X^{-1}((-\infty; x]) = \{X \leqslant x\} = \{\omega : X(\omega) \leqslant x\} \in \mathscr{F}$, т.е. является событием.*

Функция распределения случайной величины (Cumulative distribution function (cdf))

Из определения следует, что для каждого x определена вероятность события $\{X \leqslant x\}$, т.е. $\mathbb{P}(\omega : X(\omega) \leqslant x)$.

Определение

Функция

$$F_X(x) = \mathbb{P}\left(X \leqslant x\right) = \mathbb{P}\left(\omega : X(\omega) \leqslant x\right) \tag{1}$$

называется ***функцией распределения*** *случайной величины* X.

Реализация в *scipy.stats()*

$X \sim$ Закон распределения $\Leftrightarrow X =$ Закон распределения

cdf $\quad F_X(x) = \mathbb{P}(X \leqslant x) \Leftrightarrow X.cdf(x)$

Пример

$$\mathcal{Z} \sim \mathcal{N}(0;1) \Leftrightarrow \mathcal{Z} = norm(); \quad cdf \quad \Phi(x) = \mathbb{P}(\mathcal{Z} \leqslant x) \Leftrightarrow \mathcal{Z}.cdf(x).$$

Примеры

1. Пусть один раз бросают симметричную монету. Тогда $\Omega = \{\omega_1 = \text{«Г»}, \omega_2 = \text{«Р»}\}$, $\mathbb{P}(\omega_1) = \mathbb{P}(\omega_2) = \frac{1}{2}$. Рассмотрим

$$X(\omega) = \begin{cases} 1, \text{ если } & \omega = \omega_1, \\ -1, \text{ если } & \omega = \omega_2. \end{cases}$$

Тогда

$$\{X \leqslant x\} = \{\omega : X(\omega) \leqslant x\} = \begin{cases} \varnothing, x < -1, \\ \{\omega_2\}, -1 \leqslant x < 1, \\ \Omega, x \geqslant 1. \end{cases}$$

Поэтому $X(\omega)$ — случайная величина и её функция распределения $F_X(x)$ имеет вид

$$F_X(x) = \begin{cases} 0, \text{ если } & x < -1, \\ \frac{1}{2}, \text{ если } & -1 \leqslant x < 1, \\ 1, \text{ если } & x \geqslant 1. \end{cases}$$

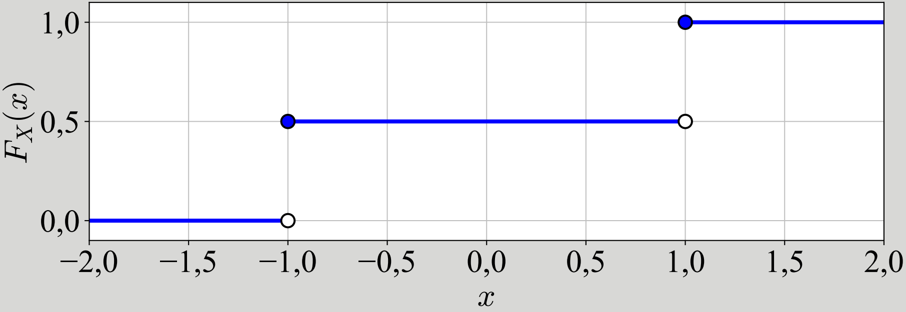

Рис. 1. *Функция распределения $F_X(x)$ для примера 1.*

2. **Равномерное распределение на отрезке $[a,b]$.** На отрезке $[a,b]$ наугад бросают точку, причём считают, что все положения точки «одинаково возможны». Рассмотрим вероятностное пространство $(\Omega=[a,b],\mathscr{F},\mathbb{P})$, где $\mathscr{F}$ — σ-алгебра борелевских подмножеств из $[a,b]$, $\mathbb{P}([\alpha,\beta])=\frac{\beta-\alpha}{b-a}$, если $[\alpha,\beta]\subset[a,b]$. Определим функцию $X(\omega)=\omega$, т.е. ω — координата полученной точки. Тогда

$$\{X\leqslant x\}=\{\omega : X(\omega)\leqslant x\}=\begin{cases}\varnothing, x<a,\\ [a,x], a\leqslant x<b,\\ \Omega=[a,b], x\geqslant b\end{cases}$$

Поэтому при каждом x множество $\{X\leqslant x\}\in\mathscr{F}$, т.е. является событием, следовательно, $X(\omega)$ является случайной величиной. Функция распределения $F_X(x)$ случайной величины $X(\omega)$:

$$F_X(x)=\begin{cases}0, \text{ если } & x<a,\\ \frac{x-a}{b-a}, \text{ если } & a\leqslant x<b,\\ 1, \text{ если } & x\geqslant b.\end{cases}$$

Эта функция определяет ***равномерное распределение на*** $[a,b]$.

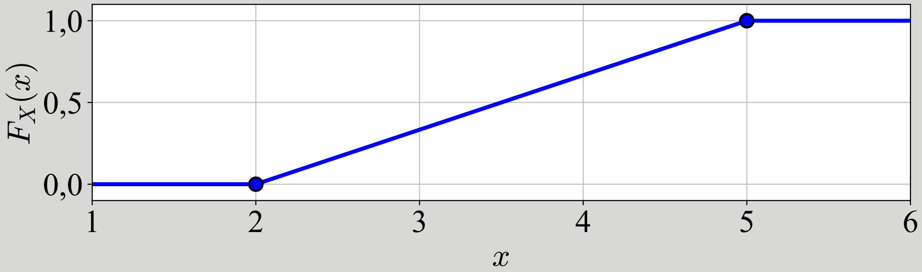

Рис. 2. *Функция распределения (cdf) $F_X(x)$ для примера 2.*

$$a = 2, b = 5$$
$$U \sim U([a, b]) \Leftrightarrow U = uniform(loc = a, scale = b - a)$$

3. Пусть $(\Omega, \mathscr{F}, \mathbb{P})$ — произвольное вероятностное пространство. Для любого события $A \in \mathscr{F}$ определим функцию

$$I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A. \end{cases}$$

Тогда

$$\{I_A \leqslant x\} = \{\omega : I_A(\omega) \leqslant x\} = \begin{cases} \varnothing, x < 0, \\ \overline{A}, 0 \leqslant x < 1, \\ \Omega, x \geqslant 1, \end{cases}$$

следовательно, множество $\{\omega : I_A(\omega) \leqslant x\} \in \mathscr{F}$ и поэтому $I_A(\omega)$ является случайной величиной, которая называется ***индикатором случайного события*** A. Функция распределения $F_I(x)$ случайной величины I_A имеет вид

$$F_I(x) = \begin{cases} 0, \text{ если } \quad x < 0, \\ \mathbb{P}\left(\overline{A}\right), \text{ если } \quad 0 \leqslant x < 1, \\ 1, \text{ если } \quad x \geqslant 1. \end{cases}$$

Упражнение

Привести пример вероятностного пространства и функции на нём, которая не является случайной величиной.

Теорема (Общие свойства функции распределения)

Функция распределения $F_X(x) = \mathbb{P}(X \leqslant x)$ удовлетворяет следующим свойствам:

1. $0 \leqslant F(x) \leqslant 1$.
2. $F(x)$ — неубывающая функция для всех $x \in \mathbb{R}^1$, т.е. $F(x_1) \leqslant F(x_2)$ для $\forall x_1 < x_2$.
3. $\mathbb{P}(x_1 < X \leqslant x_2) = F(x_2) - F(x_1)$,
4. $F(-\infty) = 0$, $F(+\infty) = 1$.
5. $F(x)$ — непрерывная справа функция, т.е. $F(x) = F(x+0)$, где $F(x+0) = \lim_{x_n \to x+0} F(x_n)$.
6. $F(x-0) = \mathbb{P}(X < x)$, где $F(x-0) = \lim_{x_n \to x-0} F(x_n)$.
7. Если $x_1 \leqslant x_2$, тогда
   а) $\mathbb{P}(x_1 \leqslant X \leqslant x_2) = F(x_2) - F(x_1 - 0)$,
   б) $\mathbb{P}(X = x) = F(x) - F(x-0)$,
   в) $\mathbb{P}(x_1 \leqslant X < x_2) = F(x_2 - 0) - F(x_1 - 0)$,
   г) $\mathbb{P}(x_1 < X < x_2) = F(x_2 - 0) - F(x_1)$.

Таким образом, каждая функция распределения не убывает, непрерывна справа и удовлетворяет свойству (4). Верно и обратное утверждение.

Теорема

Пусть $F(x)$ обладает следующими свойствами

1. $F(x)$ не убывает,
2. $F(x)$ непрерывна справа,
3. $F(-\infty) = 0$, $F(+\infty) = 1$.

Тогда существует вероятностное пространство $(\Omega, \mathscr{F}, \mathbb{P})$ и случайная величина X на нём такая, что $F_X(x) = F(x)$.

Пример

.

Упражнение

Является ли функцией распределения: а) $e^{-e^{-x}}$; б) $\frac{2}{\pi} arctg\, x$.

Дискретные случайные величины

Определение

Случайная величина X называется ***дискретной***, *если множество её значений не более чем счётно, т.е. конечно или счётно.*

Заметим, что для каждого действительного x множество $\{X = x\} = \{\omega : X(\omega) = x\} \in \mathscr{F}$, т.е. является событием. Поэтому, если $X(\omega)$ — дискретная случайная величина на вероятностном пространстве $(\Omega, \mathscr{F}, \mathbb{P})$ принимает значения $x_1, x_2, \dots, x_n, \dots$, то для каждого n определена вероятность

$$p_n = \mathbb{P}(X = x_n) = \mathbb{P}(\omega : X(\omega) = x_n). \tag{2}$$

Определение

Набор вероятностей (2) называется ***распределением случайной величины*** X.

Заметим, что $p_n \geqslant 0$, $\sum_{n=1}^{\infty} p_n = 1$. Закон распределения дискретной случайной величины удобно представлять в виде таблицы:

Значения $X(\omega)$	x_1	x_2	$\dots$	x_n	$\dots$
Вероятность $p_k = \mathbb{P}(X = x_k)$	p_1	p_2	$\dots$	p_n	$\dots$

Упражнение

Показать, что функция распределения $F_X(x)$ дискретной случайной величины X определяется по формуле:

$$F_X(x) = \sum_{k:\, x_k \leqslant x} p_k = \begin{cases} 0, & x < x_1 \\ p_1, & x_1 \leqslant x < x_2 \\ p_1 + p_2, & x_2 \leqslant x < x_3 \\ \vdots & \\ 1, & x \geqslant x_n \end{cases} \tag{3}$$

Таким образом, функция распределения любой дискретной случайной величины разрывна, возрастает скачками в точках $x = x_k$, величина скачка равна

$$F(x_k) - F(x_k - 0) = p_k = \mathbb{P}(X = x_k). \tag{4}$$

Упражнение

Пусть (cdf) имеет вид (см. Рис.3). Найдите

а) Закон распределения случайной величины X;

б) $\mathbb{P}(2 \leqslant X < 5)$;

в) $\mathbb{P}(2 < X < 5)$;

г) $\mathbb{P}(2 < X \leqslant 5)$;

д) $\mathbb{P}(X > 4)$.

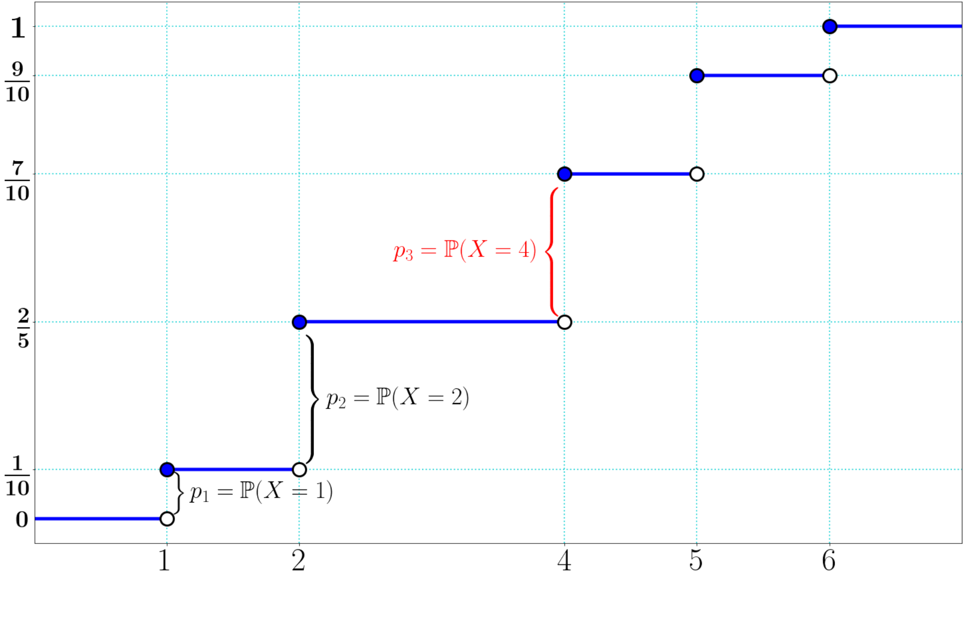

Рис.3. *Функция распределения (cdf) $F_X(x)$ для упражнения.*

Упражнение

Пусть (cdf) имеет вид (см. Рис.4). Найдите

а) Закон распределения случайной величины X;

б) $\mathbb{P}(3 \leqslant X < 9)$;

в) $\mathbb{P}(3 < X < 9)$;

г) $\mathbb{P}(3 < X \leqslant 9)$;

д) $\mathbb{P}(X > 5)$;

е) $\mathbb{P}(X = 7)$.

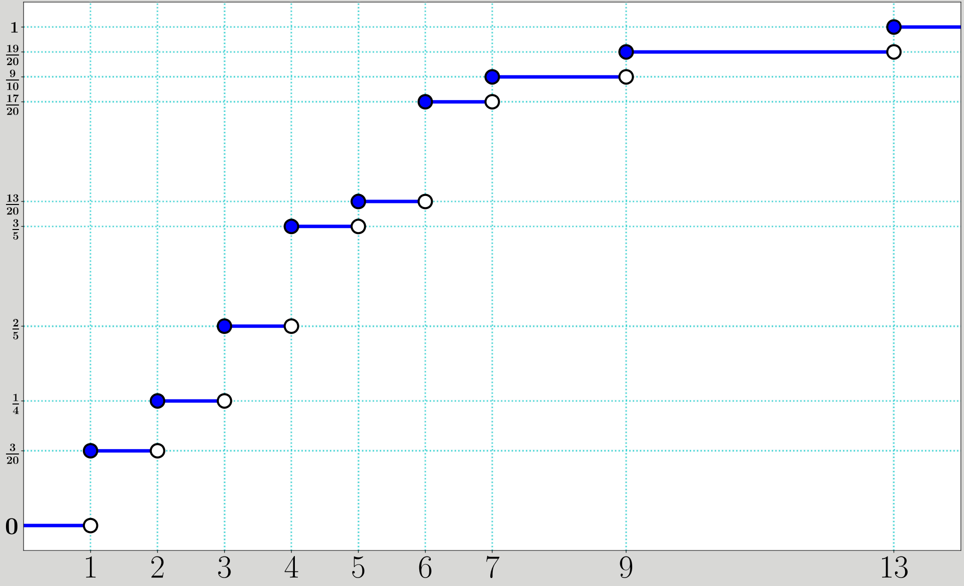

Рис.4. *Функция распределения (cdf) $F_X(x)$ для упражнения.*

Ответ:

а) Закон распределения случайной величины X имеет вид:

x_k	1	2	3	4	5	6	7	9	13
$\mathbb{P}(X = x_k)$	0,15	0,10	0,15	0,20	0,05	0,20	0,05	0,05	0,05

б) $\mathbb{P}(3 \leqslant X < 9) = 0{,}65$;

в) $\mathbb{P}(3 < X < 9) = 0{,}5$;

г) $\mathbb{P}(3 < X \leqslant 9) = 0{,}55$;

д) $\mathbb{P}(X > 5) = 0{,}35$;

е) $\mathbb{P}(X = 7) = 0{,}05$.

Независимость случайных величин

Определение

*Дискретные случайные величины $X_1, X_2, \ldots, X_n$ называются **независимыми в совокупности**, если для любых $x_1, x_2, \ldots, x_n$ события $(X_1 = x_1)$, $(X_2 = x_2)$, $\ldots$, $(X_n = x_n)$ независимы в совокупности, т.е.*

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \mathbb{P}(X_1 = x_1)\cdot\mathbb{P}(X_2 = x_2)\cdot\ldots\cdot\mathbb{P}(X_n = x_n). \quad (5)$$

В частности, дискретные случайные величины X и Y ***независимы***, если

$$p_{kl} = \mathbb{P}(X = x_k, Y = y_l) = \mathbb{P}(X = x_k) \cdot \mathbb{P}(Y = y_l) = p_k \cdot \tilde{p}_l, \qquad (6)$$

где x_k и y_l — возможные значения дискретных случайных величин X и Y соответственно. В противном случае, т.е. если найдутся такие значения x_k и y_l для которых равенство (6) не выполняется, случайные величины X и Y называются ***зависимыми***.

Пример

Пусть X — выигрыш по первому билету, Y — выигрыш по второму билету в некотором тираже, причём $a = 100$ д.е. — размер возможного выигрыша. Тогда ясно, что $\mathbb{P}(Y = a|X = a) < \mathbb{P}(Y = a)$, так как $\mathbb{P}(Y = a|X = a) = \frac{\mathbb{P}(X=a,Y=a)}{\mathbb{P}(X=a)}$ и тогда получаем, что

$$\mathbb{P}(X = a, Y = a) < \mathbb{P}(X = a) \cdot \mathbb{P}(Y = a).$$

Теорема

Пусть X и Y — независимые случайные величины, а $f(x)$ и $g(y)$ — непрерывные функции. Тогда случайные величины $f(X)$ и $g(Y)$ также будут независимыми.

Пример

Пусть X и Y — независимые случайные величины. Тогда независимыми будут 3^X и 3^Y.

Теорема

Пусть $X_1, X_2, \ldots, X_n$, $Y_1, Y_2, \ldots, Y_m$ — независимые случайные величины, $f(x_1, x_2, \ldots, x_n)$, $g(y_1, y_2, \ldots, y_m)$ — непрерывные функции. Тогда $f(X_1, X_2, \ldots, X_n)$ и $g(Y_1, Y_2, \ldots, Y_m)$ также независимы.

Пример

Если X, Y, Z — независимые случайные величины, то независимыми будут $X + Y$ и Z^2. Однако, что нельзя сказать, вообще говоря, о независимости $X + Y$ и $Z - Y$.

Упражнение

Пусть I_A, I_B, I_C — индикаторы трёх независимых событий. Показать, что I_A, I_B, I_C — независимые случайные величины.