Подготовили: Безносиков А.Н.
Шестаков А.А.

Теория вероятностей «Случайная величина. Функция распределения и плотность распределения»

Случайная величина

Определение 1. Пусть $(X, \mathcal{F}_x)$ и $(Y, \mathcal{F}_Y)$ — два измеримых пространства. Отображение $T : X \to Y$ называется **измеримым**, если для любого измеримого множества $A \in \mathcal{F}_Y$ его прообраз

$$T^{-1}(A) \stackrel{\text{def}}{=} \{x \in X \mid T(x) \in A\}$$

тоже измерим, то есть $T^{-1}(A) \in \mathcal{F}_X$.

Определение 2. $\mathcal{B}(\mathbb{R})$ — **борелевская σ-алгебра**, т.е. *минимальная* σ-алгебра, содержащая все открытые множества в $\mathbb{R}$ (эквивалентно: все замкнутые множества; все интервалы (a, b); все лучи $(-\infty, a)$).

Также в $\mathcal{B}(\mathbb{R})$ входит $\mathbb{R}$, все точки $\{b\}$, все полуинтервалы $[a, b)$, $(a, b]$ и отрезки $[a, b]$.

Определение 3. Пусть дано вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$. **Случайной величиной** ξ будем называть измеримую (*борелевскую*) функцию $\xi : \Omega \to \mathbb{R}$ из $(\Omega, \mathcal{F})$ в $(\mathbb{R}, \mathcal{B})$, где $\mathcal{B}$ — борелевская σ-алгебра на $\mathbb{R}$.

Упражнение 1. Приведите пример пространства элементарных исходов Ω, функции $\xi : \Omega \to \mathbb{R}$ и двух таких σ-алгебр, что ξ является случайной величиной для одного измеримого пространства, но не является для другого.

Решение. Рассмотрим пространство исходов бросания кости $\Omega = \{1, 2, 3, 4, 5, 6\}$ и две σ-алгебры: $\mathcal{F}_1 = \{\Omega, \varnothing\}$ (тривиальная σ-алгебра) и $\mathcal{F}_2 = 2^\Omega$ (множество всех подмножеств). Рассмотрим функцию $\xi : \Omega \to \mathbb{R}$, заданную на элементах $\omega \in \Omega$ следующим образом: $\xi(\omega) = \omega \mod 2$.

Тогда относительно второй σ-алгебры ξ является случайной величиной. Докажем это так: рассмотрим произвольное множество B из $\mathcal{B}(\mathbb{R})$, нам важны четыре случая: 1)B не содержит ни 0, ни 1; 2)B содержит 0, но не 1; 3)B содержит 1, но не 0; 4) B содержит 0 и 1; (наличие остальных элементов нам в B не принципиально). Посмотрим на прообразы в каждом из случаев: 1) $\xi^{-1}(\{B\}) = \varnothing \in \mathcal{F}_2$; 2) $\xi^{-1}(\{B\}) = \{2, 4, 6\} \in \mathcal{F}_2$; 3) $\xi^{-1}(\{B\}) = \{1, 3, 5\} \in \mathcal{F}_2$; 5) $\xi^{-1}(\{B\}) = \{1, 2, 3, 4, 5, 6\} \in \mathcal{F}_2$. Заключаем, что ξ является случайной величиной относительно второй σ-алгебры.

Покажем, что ξ не является случайной величиной относительно первой σ-алгебры, т. к. прообраз борелевского множества $\{0\}$ равен $\xi^{-1}(\{0\}) = \{2, 4, 6\} \notin \mathcal{F}_1$. $\square$

Функция и плотность распределения

Определение 4. Функцией распределения случайной величины ξ называется такая функция $F_\xi : \mathbb{R} \to [0,1]$, что $F_\xi(x) = \mathbb{P}\{\xi < x\}$.

Важное замечание: можно встретить и довольно часто (особенно в зарубежной литературе) определение функции распределения с нестрогим знаком $\leq$.

Соответствие между вероятностными мерами и функциями распределения является взаимно-однозначным. Вообще говоря, F_ξ задаёт меру $\mathbb{P}_F$ на алгебре промежутков $\mathcal{A}$: $\mathbb{P}_F\{[a_1, b_1) \cup \ldots \cup [a_n, b_n)\} \overset{\text{def}}{=} \sum\limits_{i=1}^{n}(F(b_i) - F(a_i))$. Эта мера счётно-аддитивна, поэтому однозначно продолжается на борелевскую σ-алгебру по *теореме Каратеодори о продолжении меры* следующим образом: $\mathbb{P}_F\{A\} \overset{\text{def}}{=} \inf\{\sum\limits_i \mathbb{P}_F\{B_i\} \mid A \subseteq \bigcup\limits_i B_i, B_i \in \mathcal{A}\}$.

Упражнение 2. Будем подбрасывать "честную"монетку и обозначим ξ за число орлов. Требуется найти $F_\xi(x)$.

Решение. Нетрудно заметить, что $\mathbb{P}\{\xi = 0\} = \mathbb{P}\{\xi = 2\} = \frac{1}{4}, \mathbb{P}\{\xi = 1\} = \frac{1}{2}$. Тогда функци распределения ξ будет иметь следующий вид:

$$F_\xi(x) = \begin{cases} 0 & x \leq 0 \\ 1/4 & 0 < x \leq 1 \\ 3/4 & 1 < x \leq 2 \\ 1 & x > 2. \end{cases}$$

□

Перечислим и докажем **свойства функции распределения**.

1. $F_\xi(x)$ является неубывающей функцией.

   Действительно, в силу $\{\xi < x\} \subseteq \{\xi < y\}$ для $x < y$ и в силу монотонности вероятности получаем, что $F_\xi(x) \leqslant F_\xi(y)$.

2. $\lim\limits_{x \to -\infty} F_\xi(x) = 0$ и $\lim\limits_{x \to +\infty} F_\xi(x) = 1$.

   Докажем первое равенство. Рассмотрим произвольную последовательность $\{x_n\}_{n=1}^{\infty}$, монотонно стремящуюся к $-\infty$: $x_n \xrightarrow{n \to \infty} -\infty, x_n > x_{n+1}$. Рассмотрим события $B_n = \{\xi < x_n\}$. Тогда $B_i \supseteq B_{i+1}$. Заметим, что $\bigcap\limits_{n=1}^{\infty} B_n = \varnothing$. Применяя теорему непрерывности вероятности, получаем, что $0 = \mathbb{P}\{\bigcap\limits_{n=1}^{\infty} B_n\} = \lim\limits_{n \to \infty} \mathbb{P}\{B_n\} = \lim\limits_{n \to \infty} F_\xi(x_n)$. В силу произвольности выбора последовательности, получаем $\lim\limits_{x \to -\infty} F_\xi(x) = 0$. Второе равенство доказывается аналогично.

3. $F_\xi(x)$ непрерывна слева во всех точках $\mathbb{R}$: $\forall x \in \mathbb{R} \hookrightarrow \lim\limits_{t\to x-0} F_\xi(t) = F_\xi(x)$.

   Рассмотрим произвольную возрастающую последовательность t_n, стремящуюся к x, и рассмотрим события $B_n = \{t_n \leqslant \xi < x\}$. Тогда $B_n \supseteq B_{n+1}$ для любого n и $\bigcap\limits_{n=1}^{\infty} B_n = \varnothing$, а из теоремы непрерывности вероятности получаем, что $0 = \mathbb{P}\{\bigcap\limits_{n=1}^{\infty} B_n\} = \lim\limits_{n\to\infty} \mathbb{P}\{B_n\} = \lim\limits_{n\to\infty} \mathbb{P}\{t_n \leqslant \xi < x\}$. В силу произвольности выбора возрастающей последовательности t_n получаем отсюда, что $0 = \lim\limits_{t\to x-0} \mathbb{P}\{t \leqslant \xi < x\} = \lim\limits_{t\to x-0}(F_\xi(x) - F_\xi(t)) = F_\xi(x) - \lim\limits_{t\to x-0} F_\xi(t)$.

   Важное замечание: Если определить функцию распределения с нестрогим знаком, то будет непрерывность справа.

4. $F_\xi(x)$ имеет предел справа (этот предел)

   Рассмотрим произвольный x. Пусть x_n убывает к x, тогда $F_\xi(x_n) = \mathbb{P}\{\xi < x_n\} \to \mathbb{P}\{\xi \leq x\} = F_\xi(x+0)$

   Важное замечание: Если определить функцию распределения с нестрогим знаком, то будет иметь предел слева.

5. $F_\xi(x)$ имеет не более чем счетное число разрывов.

   Пусть x_0 точка разрыва $F_\xi(x)$, значит между $F_\xi(x_0)$ и $F_\xi(x_0+0)$ есть скачок, т.е. $F_\xi(x_0) \neq F_\xi(x_0+0)$. Между двумя любыми числами есть рациональное число. Т.е. мы каждой точке разрыва можем поставить в соответствие уникальное (в силу монотонности) рациональное число. Рациональных чисел счетное число, значит точек разрыва не более чем счетно.

Следующую теорему рассмотрим без доказательства.

Теорема 1. Если функция $F : \mathbb{R} \to [0,1]$ удовлетворяет свойствам 1-3, то она является функцией распределения некоторой случайно величины.

С помощью этой теоремы можно проверять является ли данная функция функцией распределения. С помощью отдельных свойств можно доказывать обратное (что не является).

Вероятностная мера отрезков и интервалов выражается через функцию распределения следующим образом:

1) $\mathbb{P}\{\xi \leqslant a\} = \lim\limits_{t\to+0} \mathbb{P}\{\xi < a+t\} = F_\xi(a+0)$;

2) $\mathbb{P}\{\xi \geqslant a\} = 1 - \mathbb{P}\{\xi < a\} = 1 - F_\xi(a)$;

3) $\mathbb{P}\{\xi > a\} = 1 - \mathbb{P}\{\xi \leqslant a\} = 1 - F_\xi(a+0)$;

4) $\mathbb{P}\{a < \xi < b\} = \mathbb{P}\{\xi < b\} - \mathbb{P}\{\xi \leqslant a\} = F_\xi(b) - F_\xi(a+0)$;

5) $\mathbb{P}\{a \leqslant \xi \leqslant b\} = \mathbb{P}\{\xi \leqslant b\} - \mathbb{P}\{\xi < a\} = F_\xi(b+0) - F_\xi(a)$;

6) $\mathbb{P}\{a \leqslant \xi < b\} = \mathbb{P}\{\xi < b\} - \mathbb{P}\{\xi < a\} = F_\xi(b) - F_\xi(a)$;

7) $\mathbb{P}\{a < \xi \leqslant b\} = \mathbb{P}\{\xi \leqslant b\} - \mathbb{P}\{\xi \leqslant a\} = F_\xi(b+0) - F_\xi(a+0)$.

Определение 5. Распределение случайно величины ξ называется **дискретным**, если ξ может принимать конечное или счётное число значений $x_1, x_2, \ldots$ таких, что

$$p_k = \mathbb{P}\{\xi = x_k\} > 0, \quad \sum_k p_k = 1.$$

Заметим, что распределение *дискретной* случайной величины ξ однозначно задаётся её **функцией вероятности**, т. е. значениями $\mathbb{P}\{X = x_i\}$.

Пример 1. Примеры дискретных распределений:

1) распределение Бернулли: $\xi \sim \mathrm{Be}(p) \Leftrightarrow \mathbb{P}\{\xi = k\} = p^k(1-p)^{1-k}, k \in \{0, 1\}$;

   Смысл: Рассмотрим некоторое событие $A \in \mathcal{F}$ и случайную величину $X = \mathbb{I}_A$ — индикатор этого события:

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

   Очевидно, $\mathbb{P}\{X = 1\} = \mathbb{P}\{A\} = p$, $\mathbb{P}\{X = 0\} = 1 - \mathbb{P}\{A\} = 1 - p$.

   Пример: бросок несимметричной монеты, у которой вероятность выпадения орла p.

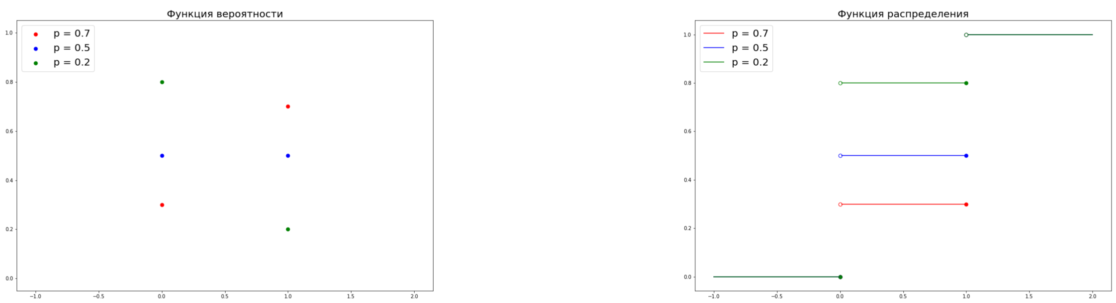

Рис. 1: распределение Бернулли

2) биномиальное распределение: $\xi \sim \mathrm{Binom}(n, p) \Leftrightarrow \mathbb{P}(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k \in \{0, 1, \ldots, n\}$;

   Смысл: распределение количество "успехов" в последовательности из n независимых случайных экспериментов, таких, что вероятность "успеха" равна p.

   Пример: n бросков несимметричной монеты, у которой вероятность выпадения орла p.

3) геометрическое распределение: $\xi \sim \mathrm{Geom}(p) \Leftrightarrow \mathbb{P}(\xi = k) = p(1-p)^{k-1}$, $k \in \mathbb{N}$;

   Смысл: распределение дискретной случайной величины, равной количеству испытаний случайного эксперимента до наблюдения первого "успеха" (сам успешный бросок при этом считается, иногда его не считают и тогда распределение начинается с 0).

   Пример: бросаем пока не выпадет орел на несимметричной монете, у которой вероятность выпадения орла p.

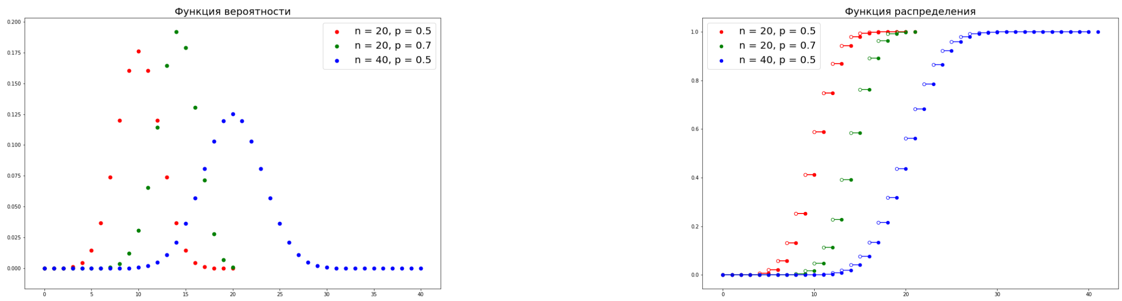

Рис. 2: биномиальное распределение

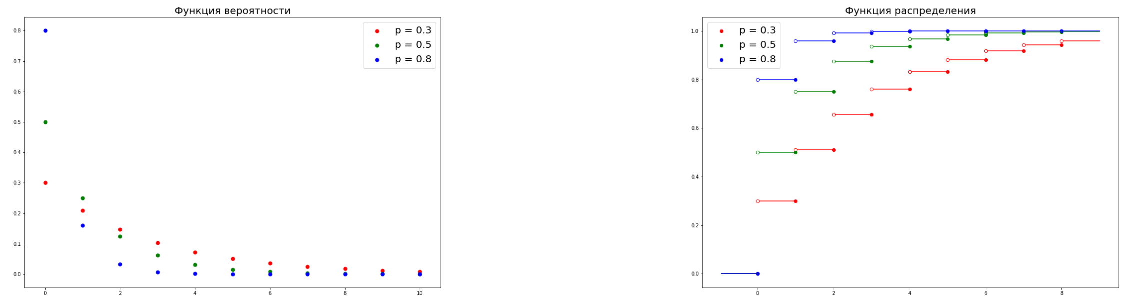

Рис. 3: геометрическое распределение

4) распределение Пуассона: $\xi \sim \text{Poisson}(\lambda) \Leftrightarrow \mathbb{P}(\xi = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k \in \mathbb{N} \cup \{0\}$;

   Смысл: рассмотрим некоторый поток событий, в котором события наступают независимо друг от друга и с некоторой фиксированной средней интенсивностью λ (событий в единицу времени). Тогда случайная величина X, равная числу событий k, произошедших за фиксированное время, имеет распределение Пуассона.

   Пример: покупатели заходящие в магазин, машины, проезжающие отметку на дороге и множество других примеров.

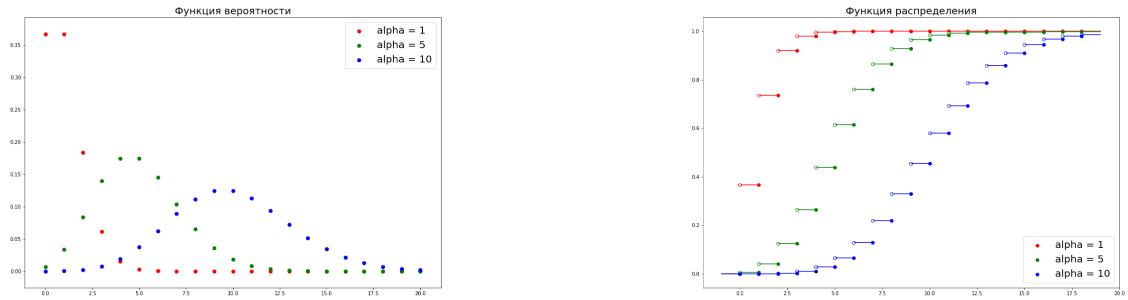

Рис. 4: распределение Пуассона

5) отрицательное биномиальное распределение: $\xi \sim \mathrm{NB}(n,p)$: $\mathbb{P}(\xi = k) = C_{k+n-1}^{k} p^n (1-p)^k$, $k \in \mathbb{Z}_+$.

Упражнение 3. Найти распределение дискретной случайной величины, равной количеству произошедших неудач в последовательности испытаний Бернулли с вероятностью успеха p, проводимой до r-го успеха. Т.е. мы имеем последовательность независимых с.в. $\xi_1, \xi_2, \ldots \sim \mathrm{Be}(p)$.

Решение. Обозначим искомую с.в. через Y . Событие $\{Y = k\}$ означает, что в ходе $(k+r)$ испытаний произошло ровно r успехов, причем последнее испытание было успешным. Следовательно,

$$\mathbb{P}\{Y = k\} = p\binom{k+r-1}{r-1} p^{r-1}(1-p)^k = \binom{k+r-1}{k} p^r (1-p)^k$$

Y имеет отрицательное биномиальное распределение с параметрами r, p. □

Определение 6. Распределение $\mathbb{P}$ случайной величины ξ называется **абсолютно непрерывным** (относительно меры Лебега), если существует такая интегрируемая функция $f(x)$, что для любого борелевского множества B

$$\mathbb{P}\{\xi \in B\} = \int_B f(x)dx \text{ (интеграл Лебега).}$$

Функция $f(x)$ называется **плотностью распределения**.

Теорема 2. Если функция $f : \mathbb{R} \to \mathbb{R}$ удовлетворяет следующим условиям:

1) $\forall x \quad f(x) \geq 0$

2) $\int\limits_{-\infty}^{\infty} f(x)dx = 1$

то существует распределение $F(x)$ такое, что $f(x)$ является его плотностью.

Из математического анализа известно, что случайная величина абсолютно непрерывная, если её функция распределения абсолютно непрерывная

Если дискретные распределения задавались функцией вероятности (которую можно отразить на графике или представить в виде таблицы), то непрерывные задаются плотностью распределения.

Пример 2. Примеры непрерывных распределений:

1) равномерное распределение на отрезке: $\xi \sim \mathcal{U}[a,b] \Leftrightarrow f(x) = \frac{1}{b-a}$, $a \leqslant x \leqslant b$;

Рис. 5: равномерное распределение на отрезке $[-1;1]$

2) нормальное распределение: $\xi \sim \mathcal{N}\left(m, \sigma^2\right) \Leftrightarrow f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$, $x \in \mathbb{R}$;

Рис. 6: нормальное распределение

3) экспоненциальное распределение: $\xi \sim \mathrm{Exp}(\lambda) \Leftrightarrow f(x) = \lambda e^{-\lambda x}$, $x \geqslant 0$;

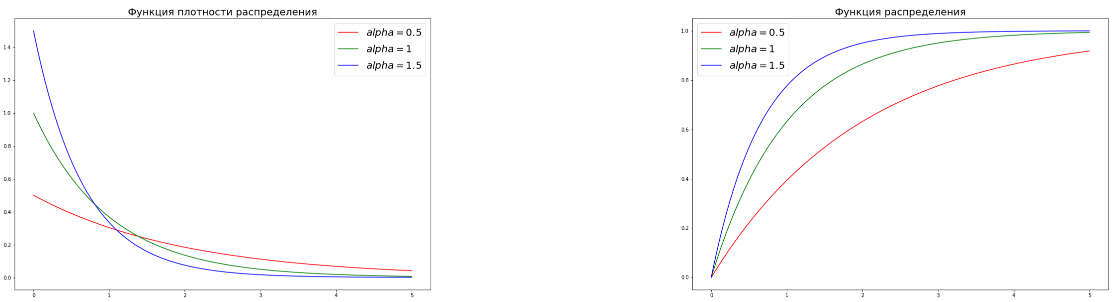

Рис. 7: экспоненциальное распределение

4) распределение Коши: $\xi \sim \mathrm{Ca}(m, \gamma) \Leftrightarrow f(x) = \frac{1}{\pi} \frac{\gamma}{(x-m)^2+\gamma^2}$, $x \in \mathbb{R}$;

5) гамма-распределение: $\xi \sim \mathrm{Gamma}(\alpha, \lambda) \Leftrightarrow f(x) = \frac{\lambda^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}$, $x \geqslant 0$;

6) хи-квадрат с n степенями свободы $\xi \sim \chi^2(n) = \mathrm{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right) \Leftrightarrow f(x) = \frac{\frac{1}{2}^{\frac{n}{2}} x^{\frac{n}{2}-1}}{\Gamma\left(\frac{n}{2}\right)} e^{-\frac{x}{2}}$, $x \geqslant 0$;

7) бета-распределение: $\xi \sim \mathrm{Beta}(\alpha, \beta) \Leftrightarrow f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $0 \leqslant x \leqslant 1$;

8) логнормальное распределение: $\xi \sim \log \mathcal{N}(m, \sigma^2) \Leftrightarrow f(x) = \frac{1}{x\sqrt{2\pi}\sigma} e^{-\frac{(\ln x - m)^2}{2\sigma^2}}$, $x \geqslant 0$.

Для нормального распределения не приведен график функции распределения, ее можно найти например здесь. Отсюда возникает вопрос, как по плотности находить само распределение и наоборот:

Упражнение 4. Найти функцию распределения, зная, что функция плотности распределения есть $f(x) = \lambda e^{-\lambda x}$, $x \geqslant 0$.

Решение. Формула очень простая:

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt = \int\limits_{0}^{x} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$$

□

Эта формула справедлива для непрерывных случайных величин. Чтобы найти плотность нужно просто взять производную $F(x)$ по x.

Определение 7. Случайные величины $\xi_1, \xi_2, \dots \xi_n$ называются **независимыми в совокупности**, если

$$\mathbb{P}\{\xi_1 \in B_1, \xi_2 \in B_2, \dots, \xi_n \in B_n\} = \mathbb{P}\{\xi_1 \in B_1\}\mathbb{P}\{\xi_2 \in B_2\} \dots \mathbb{P}\{\xi_n \in B_n\}$$

Теорема 3. Случайные величины $\xi_1, \dots, \xi_n$ независимы в совокупности тогда и только тогда, когда для любых $x_1, \dots, x_n \in \mathbb{R}$ выполняется равенство

$$\mathbb{P}\{\xi_1 < x_1, \dots, \xi_n < x_n\} = \prod_{i=1}^{n} F_{\xi_i}(x_i).$$

Упражнение 5. Предположим, что $X_1, \dots, X_n$ независимые одинаково распределенные случайные величины (в русской литературе НОРСВ, в английской i.i.d. - independent and identically-distributed) с функцией распределения $F_X(x)$. Найти функцию распределения случайной величины $Y = \max\limits_i X_i$.

Решение. Пойдем по определению: $F_Y(x) = \mathbb{P}\{\max\limits_i X_i < x\}$. Если максимум случайных величин меньше x, то и каждая из них тоже будет меньше x. С другой стороны, очевидно, что если каждая не превосходит x, то и максимум соответственно. Получается, из равенства множеств имеем $\mathbb{P}\{\max\limits_i X_i < x\} = \mathbb{P}\{X_1 < x, \dots, X_n < x\}$. Воспользуемся независимостью и получим $\mathbb{P}\{X_1 < x\} \cdots \mathbb{P}\{X_n < x\} = F_X(x)^n$. □