

Лекция 6

Функция распределения системы случайных величин. Независимость случайных величин.

Пусть на вероятностном пространстве $\{\Omega, \mathcal{F}, \mathbb{P}\}$ определены n случайных величин $X_1 = X_1(\omega), \dots, X_n = X_n(\omega)$.

Для любых чисел $x_1, x_2, \dots, x_n \in (-\infty; +\infty)$ определена вероятность

$$\mathbb{P}\{X_1 < x_1, X_2 < x_2, \dots, X_n < x_n\},$$

которая называется *функцией распределения системы случайных величин* (X_1, X_2, \dots, X_n) и обозначается $F(x_1, x_2, \dots, x_n)$:

$$F(x_1, x_2, \dots, x_n) = \mathbb{P}\{X_1 < x_1, X_2 < x_2, \dots, X_n < x_n\}.$$

Пример 1

Система (X_1, X_2, \dots, X_n) называется *равномерно распределённой в параллелепипеде* $a_i \leq x_i \leq b_i$ ($1 \leq i \leq n$), если вероятность попадания точки (x_1, x_2, \dots, x_n) в любую внутреннюю область этого параллелепипеда пропорциональна её объёму и вероятность попадания внутрь параллелепипеда есть достоверное событие.

Функция распределения такой системы имеет вид:

$$F(x_1, x_2, \dots, x_n) = \begin{cases} 0, & \text{если } x_i \leq a_i \text{ хотя бы при одном } i, \\ \prod_{i=1}^n \frac{c_i - a_i}{b_i - a_i}, & \text{где } c_i = \begin{cases} x_i, & \text{если } a_i < x_i \leq b_i, \\ b_i, & \text{если } x_i > b_i \end{cases} \end{cases}$$

Пример 2

Система (X_1, X_2) имеет (*двумерное*) *нормальное распределение*, если $F(x_1, x_2) =$

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \exp \left\{ -\frac{1}{2(1-r^2)} \left[\frac{(u_1-a_1)^2}{\sigma_1^2} - 2r \frac{(u_1-a_1)(u_2-a_2)}{\sigma_1\sigma_2} + \frac{(u_2-a_2)^2}{\sigma_2^2} \right] \right\} du_1 du_2,$$

где $\sigma_1 > 0$, $\sigma_2 > 0$, $|r| < 1$, a_1 , a_2 — произвольные действительные числа.

Функция $F(x_1, x_2, \dots, x_n)$ удовлетворяет следующим свойствам:

1. F есть неубывающая функция по каждому аргументу;
2. F непрерывна слева по каждому аргументу;
3. $\lim_{\substack{x_1 \rightarrow +\infty \\ \vdots \\ x_n \rightarrow +\infty}} F(x_1, x_2, \dots, x_n) = 1;$
4. $\forall k (1 \leq k \leq n) \quad \lim_{x_k \rightarrow -\infty} F(x_1, \dots, x_k, \dots, x_n) = 0.$

Плотность распределения

Если существует функция $f(x_1, x_2, \dots, x_n)$ такая, что $\forall x_1, x_2, \dots, x_n$

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n,$$

то эта функция называется *плотностью распределения системы случайных величин* (X_1, X_2, \dots, X_n) .

Её свойства:

1. $f(x_1, x_2, \dots, x_n) \geq 0;$
2. вероятность попадания случайной n -мерной точки (X_1, X_2, \dots, X_n) в какую-либо область $G \subset \mathbb{R}^n$ равна

$$\int_G \cdots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n;$$

3. в частности, если f непрерывна в точке (x_1, x_2, \dots, x_n) , то при малых $\Delta x_1, \Delta x_2, \dots, \Delta x_n$

$$P \left\{ \begin{array}{l} x_1 \leq X_1 < x_1 + \Delta x_1, \\ x_2 \leq X_2 < x_2 + \Delta x_2, \\ \dots \\ x_n \leq X_n < x_n + \Delta x_n \end{array} \right\} \approx f(x_1, x_2, \dots, x_n) \Delta x_1 \Delta x_2 \dots \Delta x_n$$

с точностью до малых высшего порядка.

Пример 3

Пусть (X_1, \dots, X_n) равномерно распределена в n -мерной области G . Если V — n -мерный объём области G , то

$$f(x_1, \dots, x_n) = \begin{cases} \frac{1}{V}, & \text{если } (x_1, \dots, x_n) \in G, \\ 0, & \text{если } (x_1, \dots, x_n) \notin G. \end{cases}$$

Пример 4

Плотность двумерного нормального распределения:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)} \left[\frac{(x_1-a_1)^2}{\sigma_1^2} - 2r\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2} \right] \right\}.$$

Независимость случайных величин

Случайные величины X_1, X_2, \dots, X_n называются *независимыми*, если для любой группы $X_{i_1}, X_{i_2}, \dots, X_{i_k}$ этих величин

$$\mathsf{P}\{X_{i_1} < x_{i_1}, X_{i_2} < x_{i_2}, \dots, X_{i_k} < x_{i_k}\} = \prod_{j=1}^k \mathsf{P}\{X_{i_j} < x_{i_j}\}$$

для произвольных $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ и любом k ($1 \leq k \leq n$).

В частности, при $k = n$ в терминах функций распределения это равенство принимает вид:

$$F(x_1, x_2, \dots, x_n) = F_1(x_1)F_2(x_2) \dots F_n(x_n),$$

где $F_k(x_k) = \mathsf{P}\{X_k < x_k\}$ — функция распределения случайных величин X_k .

Из определения плотности $f(x_1, x_2, \dots, x_n)$ следует, что в точках непрерывности функции f

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}.$$

В дальнейшем пусть для определенности $n = 2$.

Итак, если X_1 и X_2 — независимые случайные величины, то

$$F(x_1, x_2) = F_1(x_1)F_2(x_2).$$

Если при этом существуют плотности распределения $f_1(x_1)$ и $f_2(x_2)$ случайных величин X_1 и X_2 , то

$$f(x_1, x_2) = f_1(x_1)f_2(x_2).$$

Одно свойство нормального распределения

Пусть X_1 и X_2 — независимые нормально распределённые случайные величины, т.е.

$$f_1(x_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2}}, \quad f_2(x_2) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x_2-a_2)^2}{2\sigma_2^2}}.$$

Тогда плотность совместного распределения системы (X_1, X_2) будет иметь вид

$$f(x_1, x_2) = f_1(x_1)f_2(x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left[-\frac{(x_1-a_1)^2}{2\sigma_1^2} - \frac{(x_2-a_2)^2}{2\sigma_2^2} \right].$$

Полученная функция $f(x_1, x_2)$ в соответствии с примером 4 определяет плотность двумерного нормального распределения при $r = 0$.

Итак, если каждая компонента системы (X_1, X_2) имеет нормальное распределение и компоненты независимы, то совместное распределение системы (X_1, X_2) тоже будет нормальным.

Поставим вопрос: а если не требовать независимости компонент X_1 и X_2 , можно ли утверждать, что совместное распределение системы (X_1, X_2) будет нормальным? Оказывается, нет. Это иллюстрирует следующий пример.

Пример 5

Пусть плотность $f(x_1, x_2)$ совместного распределения системы X_1 и X_2 имеет вид (не является нормальным):

$$f(x_1, x_2) = \begin{cases} \frac{1}{\pi\sigma^2} e^{-\frac{x_1^2+x_2^2}{2\sigma^2}}, & \text{при } x_1x_2 \geq 0, \\ 0, & \text{при } x_1x_2 < 0. \end{cases}$$

Проверим условие нормировки:

$$\begin{aligned}
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x_1, x_2) dx_1 dx_2 &= \int_0^{+\infty} \int_0^{+\infty} f(x_1, x_2) dx_1 dx_2 + \int_{-\infty}^0 \int_{-\infty}^0 f(x_1, x_2) dx_1 dx_2 = \\
&= 2 \int_0^{+\infty} \int_0^{+\infty} f(x_1, x_2) dx_1 dx_2 = \frac{2}{\pi\sigma^2} \int_0^{+\infty} \int_0^{+\infty} e^{-\frac{x_1^2+x_2^2}{2\sigma^2}} dx_1 dx_2 = \\
&= \frac{2}{\pi\sigma^2} \left(\int_0^{+\infty} e^{-\frac{x_1^2}{2\sigma^2}} dx_1 \right)^2 = \frac{2}{\pi\sigma^2} \left(\sigma \int_0^{+\infty} e^{-\frac{t^2}{2}} dt \right)^2 = \\
&= \frac{2}{\pi} \left(\int_0^{+\infty} e^{-\frac{t^2}{2}} dt \right)^2 = \frac{2}{\pi} \left(\sqrt{2\pi} \cdot \frac{1}{2} \right)^2 = \frac{2}{\pi} \cdot \frac{\pi}{2} = 1.
\end{aligned}$$

Воспользуемся следующими общими формулами, связывающими двумерную и одномерные плотности:

$$f_1(x_1) = \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2, \quad f_2(x_2) = \int_{-\infty}^{+\infty} f(x_1, x_2) dx_1. \quad (1)$$

Действительно,

$$\begin{aligned}
f_1(x_1) &= \frac{d}{dx_1} F_1(x_1) = \frac{d}{dx_1} \lim_{x_2 \rightarrow +\infty} F(x_1, x_2) = \\
&= \frac{d}{dx_1} \lim_{x_2 \rightarrow +\infty} \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(t_1, t_2) dt_1 dt_2 = \frac{d}{dx_1} \int_{-\infty}^{x_1} \left[\int_{-\infty}^{+\infty} f(t_1, t_2) dt_2 \right] dt_1 = \\
&= \int_{-\infty}^{+\infty} f(x_1, t_2) dt_2 \quad — получили первую формулу (1);
\end{aligned}$$

вторая получается аналогично.

Вернёмся к примеру и найдём $f_1(x_1)$:

$$\begin{aligned}
f_1(x_1) &= \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 = \\
&= \begin{cases} e^{-\frac{x_1^2}{2\sigma^2}} \cdot \frac{1}{\pi\sigma^2} \int_0^{+\infty} e^{-\frac{x_2^2}{2\sigma^2}} dx_2, & \text{при } x_1 \geq 0, \\ e^{-\frac{x_1^2}{2\sigma^2}} \cdot \frac{1}{\pi\sigma^2} \int_{-\infty}^0 e^{-\frac{x_2^2}{2\sigma^2}} dx_2, & \text{при } x_1 < 0 \end{cases} \\
&= e^{-\frac{x_1^2}{2\sigma^2}} \cdot \frac{1}{\pi\sigma^2} \int_0^{+\infty} e^{-\frac{x_2^2}{2\sigma^2}} dx_2 = e^{-\frac{x_1^2}{2\sigma^2}} \cdot \frac{1}{\pi\sigma^2} \int_0^{+\infty} e^{-\frac{t^2}{2}} \sigma dt = \\
&= e^{-\frac{x_1^2}{2\sigma^2}} \cdot \frac{1}{\pi\sigma^2} \cdot \sigma \cdot \sqrt{2\pi} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x_1^2}{2\sigma^2}} \quad \forall x_1 \in (-\infty; +\infty).
\end{aligned}$$

Аналогично устанавливается, что

$$f_2(x_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2\sigma^2}} \quad \forall x_2 \in (-\infty; +\infty).$$

Таким образом, мы видим, что каждая случайная величина X_1 и X_2 имеет по отдельности нормальное распределение, но совместное распределение системы (X_1, X_2) не является нормальным.