Математические основы искусственного интеллекта
Системы случайных величин - II

Солодушкин Святослав Игоревич

Кафедра вычислительной математики и компьютерных наук,
УрФУ имени первого Президента России Б.Н. Ельцина

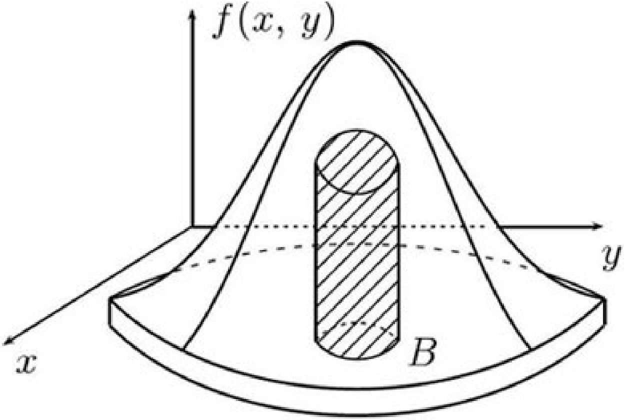
Ноябрь 2021

Пусть случайные величины $\xi_1, \ldots, \xi_n$ заданы на одном вероятностном пространстве $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$.

Определение

Функция $F_{\xi_1, \ldots, \xi_n}(x_1, \ldots, x_n) = \mathrm{P}(\xi_1 < x_1, \ldots, \xi_n < x_n)$ называется функцией распределения вектора $(\xi_1, \ldots, \xi_n)$ или функцией совместного распределения случайных величин $\xi_1, \ldots, \xi_n$.

Плотность распределения



Плотность распределения

Опрпеделение

Случайные величины ξ_1, ξ_2 имеют абсолютно непрерывное совместное распределение, если существует неотрицательная функция $f_{\xi_1,\,\xi_2}(x,\, y)$ такая, что для любого множества $B \in \mathfrak{B}(\mathbb{R}^2)$ имеет место равенство

$$\mathrm{P}((\xi_1,\, \xi_2) \in B) = \iint\limits_{B} f_{\xi_1,\,\xi_2}(x,\, y)\, dx\, dy.$$

Плотность распределения

Если случайные величины ξ_1, ξ_2 имеют абсолютно непрерывное совместное распределение, то для любых x_1, x_2 имеет место равенство

$$F_{\xi_1,\,\xi_2}(x_1,\, x_2) = \mathrm{P}(\xi_1 < x_1,\, \xi_2 < x_2) = \int\limits_{-\infty}^{x_1} \left(\int\limits_{-\infty}^{x_2} f_{\xi_1,\,\xi_2}(x,\, y)\, dy \right) dx.$$

Если совместное распределение абсолютно непрерывно, то по функции совместного распределения его плотность находится как смешанная частная производная:

$$f_{\xi_1,\,\xi_2}(x,\, y) = \frac{\partial^2}{\partial x \partial y} F_{\xi_1,\,\xi_2}(x,\, y).$$

Плотность распределения

Если случайные величины ξ_1 и ξ_2 имеют абсолютно непрерывное совместное распределение с плотностью $f_{\xi_1,\xi_2}(x, y)$, то ξ_1 и ξ_2 в отдельности также имеют абсолютно непрерывное распределение с плотностями:

$$f_{\xi_1}(x) = \int\limits_{-\infty}^{\infty} f_{\xi_1,\xi_2}(x, y)\, dy; \quad f_{\xi_2}(y) = \int\limits_{-\infty}^{\infty} f_{\xi_1,\xi_2}(x, y)\, dx.$$

Для $n > 2$ плотности случайных величин $\xi_1, \ldots, \xi_n$ по плотности их совместного распределения $f_{\xi_1,\ldots,\xi_n}(x_1, \ldots, x_n)$ находятся интегрированием функции f по всем «лишним» координатам.

Независимость случайных величин

Определение

Случайные величины $\xi_1, \ldots, \xi_n$ независимы (в совокупности), если для любых $x_1, \ldots, x_n$ имеет место равенство

$$F_{\xi_1, \ldots, \xi_n}(x_1, \ldots, x_n) = F_{\xi_1}(x_1) \cdot \ldots \cdot F_{\xi_n}(x_n).$$

Условные законы распределения

Условная вероятность события B, если событие A наступило, вычисляется так:

$$P(B|A) = \frac{P(A\,B)}{P(A)}.$$

Рассмотрим двумерную дискретную случайную величину (ξ, η).

$\xi \mid \eta$	y_1	y_2	$\dots$	y_m
x_1	p_{11}	p_{12}	$\dots$	p_{1m}
x_2	p_{21}	p_{22}	$\dots$	p_{2m}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

Допустим, что в результате испытания величина η приняла значение y_1; при этим ξ принимает одно из значений: $x_1, x_2, \dots, x_n$. Обозначим условную вероятность того, что ξ примет значение x_1 (при условии, что $\eta = y_1$) через $p(x_1|y_1)$.

Условные законы распределения

Условная вероятность того, что ξ примет значение x_i (при условии, что $\eta = y_1$)

$$p(x_i|y_1) = \frac{p(x_i, y_1)}{p(y_1)}.$$

В общем случае условная вероятность того, что ξ примет значение x_i (при условии, что $\eta = y_j$)

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}.$$

Дано распределение дискретной случайной величины (ξ, η)

$\xi \mid \eta$	y_1	y_2	y_3
x_1	0.20	0.30	0.10
x_2	0.05	0.20	0.15

Найдем условное распределение η при условии, что $\xi = x_1$.

η при $\xi = x_1$	y_1	y_2	y_3
P	1/6	3/6	2/6

Найдем условное распределение ξ при условии, что $\eta = y_1$, а также ξ при условии, что $\eta = y_3$.

ξ при $\eta = y_1$	x_1	x_2
P	0.8	0.2

ξ при $\eta = y_3$	x_1	x_2
P	0.4	0.6

Условные законы распределения непрерывных случайных величин

Пусть (ξ, η) —непрерывная двумерная случайная величина. Условной плотностью $f_\xi(x|\eta = y)$ распределения составляющей ξ при условии $\eta = y$ называется функция

$$f_\xi(x|\eta = y) = \frac{f_{\xi,\eta}(x, y)}{f_\eta(y)}.$$

Если задана плотность $f_{\xi,\eta}(x, y)$ то условные плотности составляющих можно найти по формулам

$$f_\xi(x|\eta = y) = \frac{f_{\xi,\eta}(x, y)}{\int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x, y)\, dx}, \qquad f_\eta(y|\xi = x) = \frac{f_{\xi,\eta}(x, y)}{\int\limits_{-\infty}^{+\infty} f_{\xi,\eta}(x, y)\, dy}.$$

Условное математическое ожидание

Условное математическое ожидание случайной величины η при условии, что $\xi = x$,

$$E(\eta|\xi = x) = \int\limits_{-\infty}^{+\infty} y\, f_\eta(y|\xi = x)\, dy.$$

Условное математическое ожидание $E(\eta|\xi = x)$ — функция от x. Эту функцию называют *функцией регрессии*.

Математическое ожидание произведения

Теорема

Если ξ и η независимы и их математические ожидания существуют, то $\mathrm{E}\,(\xi\eta) = \mathrm{E}\,\xi\,\mathrm{E}\,\eta$.

Доказательство. В дискретном случае

$$\mathrm{E}\,(\xi\eta) = \sum_{k,\,n}(x_k\,y_n)\,\mathrm{P}(\xi = x_k,\,\eta = y_n) =$$

$$= \sum_{k} x_k\,\mathrm{P}(\xi = x_k)\sum_{n} y_n\,\mathrm{P}(\eta = y_n) = \mathrm{E}\,\xi\,\mathrm{E}\,\eta.$$

Человек идет 1 км на работу пешком с постоянной скоростью. Его скорость зависит от настроения и описывается случайной величиной ξ:

ξ, км/ч	1	2	4	5
P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Найти мат. ожидание скорости. Сколько в среднем минут он потратит на дорогу?

$$\mathsf{E}\,\xi = \frac{1}{4}(1+2+4+5) = 3 \text{ км/ч}.$$

В 1 часе — 60 минут.
Найдем время в пути: 1км / 3 км/ч * 60 мин = 20 мин.

Мат. ожидание функции от случайной величины

Решение, представленное на предыдущем слайде, неверное!

$$\mathsf{E}\, f(\xi) \neq f(\mathsf{E}\,\xi).$$

Мат. ожидание от функции не равно функция от мат. ожидания.

ξ, км/ч	1	2	4	5
$1/\xi$, ч	1	0.5	0.25	0.2
P	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

$$\mathsf{E}\left(\frac{1}{\xi}\right) = \frac{1}{4}(1 + 0.5 + 0.25 + 0.2) = 39/20 \text{ ч}.$$

Это 29 минут 15 секунд.

1. Найти вероятность попадания случайной точки (ξ, η) в прямоугольник $[\pi/6,\ \pi/2] \times [\pi/4,\ \pi/3]$, если известна функция распределения

$$F_{\xi,\eta}(x, y) = \sin x \sin y, \quad (x, y) \in [0,\ \pi/2] \times [0,\ \pi/2].$$

В условии задачи функция $F_{\xi,\eta}(x, y)$ задана не на всем множестве. А как надо формально продлить ее на $\mathbb{R}^2$?
Найти плотность распределения $f_{\xi,\eta}(x, y)$.

2. Плотность распределения двумерной случайной величины задана следующим образом

$$f_{\xi,\eta}(x,y) = \begin{cases} C\cos x \cos y, & \text{если } (x,y) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \\ 0, & \text{иначе }. \end{cases} \tag{1}$$

Найти напамерт C, вероятность попадания в прямоугольник $\left[0, \frac{\pi}{6}\right] \times \left[0, \frac{\pi}{3}\right]$, функцию распределения $F_{\xi,\eta}(x,y)$. Найти плотности распределений $f_\xi(x)$ и $f_\eta(y)$.

3. Плотность распределения двумерной случайной величины задана следующим образом

$$f_{\xi,\eta}(x,y) = \begin{cases} C, & \text{если } 2x + y \le 4,\ 0 \le x \le 3,\ 0 \le y. \\ 0, & \text{иначе }. \end{cases} \tag{2}$$

Найти параметр C.
Найти условное распределение $f_\xi(x|\eta = y)$, найти математическое условное ожидание $E(x|\eta = y)$.

Содержательная задача про слепого Пью

Как-то раз слепому Пью показалось, что он окружен солдатами. Прижавшись к стене, он стал стрелять во все стороны, равномерно выбирая угол от 90° влево до 90° вправо. Напротив Пью стоял забор, в который и попадали пули.

Каково распределение пуль, попавших в доски забора? Сколько пуль в среднем попало в доску прямо напротив Пью? Сколько в доску в 10 ярдах справа от нее? Каково среднее этого распределения?

Случайная величина ξ имеет распределение Коши, если плотность распределения $f_\xi(x)$ имеющет вид

$$f_X(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x-x_0}{\gamma}\right)^2\right]} = \frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right],$$

где
$x_0 \in \mathbb{R}$ — параметр сдвига;
$\gamma > 0$ — параметр масштаба.

Если $x_0 = 0$ и $\gamma = 1$, то такое распределение называется стандартным распределением Коши:

$$f_X(x) = \frac{1}{\pi}\frac{1}{x^2+1}.$$

График плотности распределения Коши