Лекция 12

Совместное распределение случайных величин

Пусть $\xi_1, \xi_2, \dots, \xi_n$ заданы на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, p)$

Def. Случайным вектором $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ называется упорядоченный набор случайных величин, заданных на одном вероятностном пространстве

Случайный вектор задает отображение $(\xi_1, \dots, \xi_n)(\omega) : \Omega \longrightarrow \mathbb{R}^n$

Поэтому случайный вектор еще называют многомерной случайной величиной, а соответствующее ей распределение многомерным распределением:

$\forall B \in \mathcal{B}(\mathbb{R}^n) \qquad P(B) = P(\omega \in \Omega \mid (\xi_1, \dots, \xi_n) \in B)$

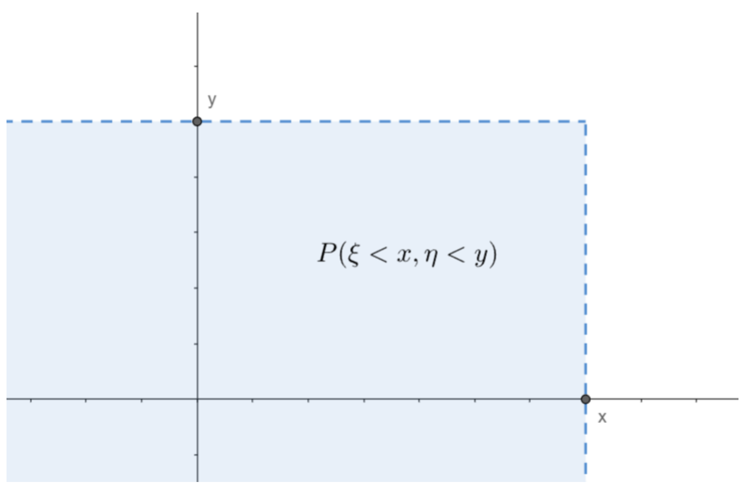
Таким образом, получили новое вероятностное пространство. В качестве элементарных исходов берем точки многомерного пространства, а σ-алгебра - многомерное Борелевская σ-алгебра $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), P(B))$

Функция распределения

Def. Функцией совместного распределения случайных величин $\xi_1, \xi_2, \dots, \xi_n$ называется функция $F_{\xi_1,\xi_2,\dots,\xi_n}(x_1, x_2, \dots, x_n) = P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n)$

Nota. Распределение полностью задается функцией распределения

Nota. В дальнейшем, в основном, будем рассматривать системы из 2 случайных величин. Функция распределения в данном случае $F_{\xi,\eta}(x, y) = P(\xi < x, \eta < y)$ - вероятность попадания в эту область.



Свойства функции распределения

1. $0 \le F_{\xi,\eta}(x, y) \le 1$

исходники найдутся тут:
https://github.com/pelmesh619/itmo_conspects

2. $F_{\xi,\eta}(x, y)$ - неубывающая по каждому аргументу
3. $\lim\limits_{x\to-\infty} F_{\xi,\eta}(x, y) = \lim\limits_{y\to-\infty} F_{\xi,\eta}(x, y) = 0$, $\lim\limits_{\substack{x\to\infty \\ y\to\infty}} F_{\xi,\eta}(x, y) = 1$
4. Восстановление маргинального (частного) распределения: $\lim\limits_{x\to\infty} F_{\xi,\eta}(x, y) = F_\eta(y)$, и наоборот - $\lim\limits_{y\to\infty} F_{\xi,\eta}(x, y) = F_\xi(x)$
5. $F_{\xi,\eta}(x, y)$ - непрерывна слева по каждому аргументу
6. $P(x_1 \le \xi < x_2, y_1 \le \eta < y_2) = F_{\xi,\eta}(x_2, y_2) - F_{\xi,\eta}(x_2, y_1) - F_{\xi,\eta}(x_1, y_2) + F_{\xi,\eta}(x_1, y_1)$

Независимость случайных величин

Def. Случайные величины $\xi_1, \dots, \xi_n$ независимы в совокупности, если для любого набора Борелевских множеств из $\mathcal{B}(\mathbb{R}^n)$, $B_1, B_2, \dots, B_n$
$p(\xi_1 \in B_1, \xi_2 \in B_2, \dots, \xi_n \in B_n) = p(\xi_1 \in B_1) \cdot p(\xi_2 \in B_2) \cdot \dots \cdot p(\xi_n \in B_n)$

Def. Случайные величины $\xi_1, \xi_2, \dots, \xi_n$ попарно независимы, если независимы любые две из них
Nota. Из независимости в совокупности следует попарная независимость:
$\xi_1, \dots, \xi_n$ независимы в совокупности, тогда покажем $\forall i, j$ ξ_i и ξ_j - независимы
Возьмем набор $B_i, B_j \in \mathcal{B}(\mathbb{R}^n)$, при $k \ne i, j$ $B_k = \mathbb{R}$ $\quad P(\xi_k \in B_k) = 1$
Тогда $p(\xi_1 \in B_1, \dots, \xi_n \in B_n) = P(\xi_i \in B_i, \xi_j \in B_j) = P(\xi_i \in B_i) \cdot P(\xi_j \in B_j)$

Nota. Из попарной независимости не следует независимость в совокупности, как видно из примера Бернштейна
Под независимыми величинами будем понимать независимые в совокупности

Дискретная система двух случайных величин

Def. Случайные величины ξ, η имеют совместное дискретное распределение, если случайный вектор (ξ, η) принимает не более, чем счетное число значений, то есть существует конечный или счетный набор пар чисел (x_i, y_i), таких что $P(\xi = x_i, \eta = y_i) > 0$, $\sum\limits_{i,j} P(\xi = x_i, \eta = y_i) = 1$
Таким образом двумерная дискретная случайная величина задается законом распределения - таблице вероятностей

$\xi \backslash \eta$	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}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\dots$	p_{nm}

Условие нормировки: $\sum\limits_{i,j} p_{i,j} = 1$

Зная общий закон распределения, можно восстановить частное (маргинальное) распределение по формулам:

$p_i = \sum\limits_{j=1}^{m} p_{i,j} \qquad q_j = \sum\limits_{i=1}^{n} p_{i,j}$

Def. Дискретные случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ независимы, если для любых $x_1, x_2, \ldots, x_n$

$p(\xi_1 = x_1, \xi_2 = x_2, \ldots, \xi_n = x_n) = p(\xi_1 = x_1) \cdot p(\xi_2 = x_2) \cdot \cdots \cdot p(\xi_n = x_n)$

При $n = 2$: $p_{i,j} = p_i \cdot q_j \ \forall i, j$

Ex.

$\xi \backslash \eta$	-1	0	1	p_i
-1	0.1	0.2	0.1	0.4
2	0.2	0.3	0.1	0.6
q_j	0.3	0.5	0.2	$\Sigma = 1$

Найти маргинальное распределение и проверить независимость случайных величин

ξ	-1	2
p_i	0.4	0.6

η	-1	0	1
q_j	0.3	0.5	0.2

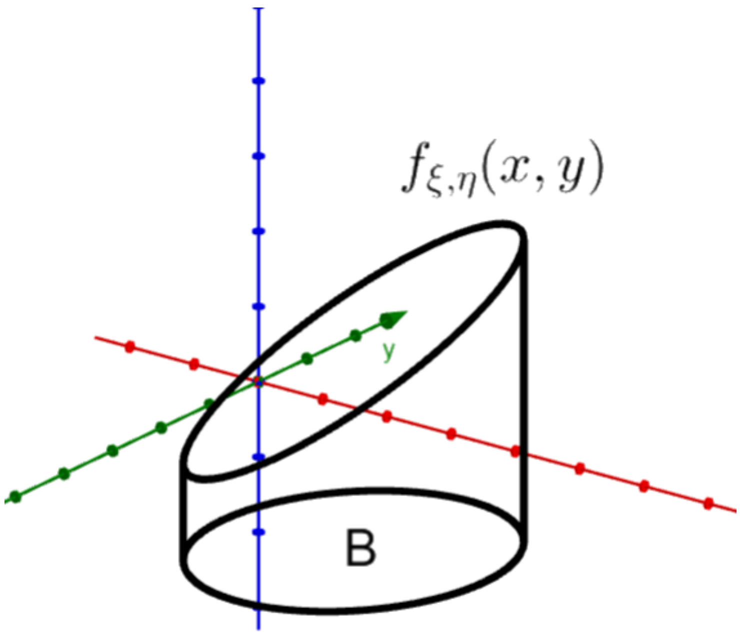
$p_{11} = 0.1 \neq 0.12 = p_1 \cdot q_1 \qquad \Longrightarrow \xi, \eta$ - зависимы

Абсолютно непрерывная система двух случайных величин

Def. Случайные величины ξ и η имеют абсолютно непрерывное совместное распределение, если $\exists f_{\xi,\eta}(x, y)$, такая что $\forall B \in \mathcal{B}(\mathbb{R}^2)\ P((\xi, \eta) \in B) = \iint\limits_B f_{\xi,\eta}(x, y) dx dy$

Функцию $f_{\xi,\eta}(x, y)$ будем называть функцией плотности совместного распределения случайных величин ξ и η

Геометрический смысл плотности:



Свойства плотности:

1. $f_{\xi,\eta}(x, y) \geq 0$
2. Условие нормировки: $\iint_{\mathbb{R}^2} f_{\xi,\eta}(x, y)dxdy = 1$
3. $F_{\xi,\eta} = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\xi,\eta}(x, y)dydx$
4. $f_{\xi,\eta}(x, y) = \dfrac{\partial^2 F_{\xi,\eta}(x, y)}{\partial x \partial y}$
5. Если случайные величины ξ, η имеют абсолютно непрерывное совместное распределение с плотностью $f(x, y)$, то маргинальное распределение величин ξ, η также имеют абсолютно непрерывное распределение с плотностями $f_\xi(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x, y)dy, f_\eta(y) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x, y)dx$

   $F_\xi(x) = \lim_{y\to\infty} F_{\xi,\eta}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(x, y)dydx$

   Из этого $\int_{-\infty}^{\infty} f(x, y)dy = f_\xi(x)$

6. Так как вероятность попадания в Борелевские множества полностью задается функцией распределения, то условие независимости случайных величин эквивалентно следующему: $\xi_1, \xi_2, \ldots, \xi_n$ независимы, если функция общего распределения распадается в произведение отдельных функцию распределения

   $F_{\xi_1,\xi_2,\ldots,\xi_n}(x_1, x_2, \ldots, x_n) = F_{\xi_1}(x_1) \cdot F_{\xi_2}(x_2) \cdot \cdots \cdot F_{\xi_n}(x_n)$
7. *Равносильное определение*: абсолютно непрерывные случайные величины $\xi_1, \ldots, \xi_n$ неза-

висимы в совокупности тогда и только тогда, когда плотность совместного распределения $f_{\xi_1,\xi_2,\dots,\xi_n}(x_1,x_2,\dots,x_n) = f_{\xi_1}(x_1)\cdot f_{\xi_2}(x_2)\cdot\dots\cdot f_{\xi_n}(x_n)$

При $n = 2$ случайные величины ξ и η независимы $\iff F_{\xi,\eta}(x,y) = F_\xi(x)\cdot F_\eta(y) = \int_{-\infty}^{x} f_\xi(x)dx \cdot \int_{-\infty}^{y} f_\eta(y)dy = \int_{-\infty}^{x}\int_{-\infty}^{y} f_\xi(x)\cdot f_\eta(y)dxdy \Longrightarrow f_{\xi,\eta}(x,y) = f_\xi(x)f_\eta(y)$
Аналогично для высших размерностей

Nota. Совместное распределение абсолютно непрерывных случайных величин не обязано быть абсолютно непрерывным, оно может быть сингулярным
Ex. Бросаем точку на отрезок прямой $y = x$ $(0 \le x, y \le 1)$, ξ - абсцисса точки, η - ордината точки
Случайный вектор (ξ, η) имеют сингулярное распределение (непрерывное с нулевой областью) - так как число элементарных исходов несчетно, но мера Лебега в $\mathbb{R}^2$ отрезка равна 0

Nota. Совместное распределение $\xi_1, \xi_2, \dots, \xi_n$ будет сингулярным, если одна из координат является функцией других (наблюдается функциональная зависимость)

Многомерное равномерное распределение

Def. $\exists D \subset \mathbb{R}^n$ - Борелевское множество в $\mathbb{R}^n$ с конечной мерой Лебега $(0 < \lambda(D) < \infty)$, случайный вектор $(\xi_1, \dots, \xi_n)$ имеет равномерное распределение, если плотность совместного распределения постоянна в данной области и равна нулю вне данной области

$$f_{\xi_1,\dots,\xi_n}(x_1,\dots,_n) = \begin{cases} \frac{1}{\lambda(D)}, & \text{если } (x_1,\dots,x_n) \in D \\ 0, & \text{если } (x_1,\dots,x_n) \notin D \end{cases}$$