Лекция 13

Математическое ожидание и дисперсия случайного вектора

$\exists \vec{\xi} = (\xi_1, \dots, \xi_n)$ - случайный вектор, $\forall 1 \le i \le n \ \xi_i$ - случайная величина

Def. Математическим ожиданием случайного вектора называется вектор с координатами из математических ожиданий его компонент: $E\vec{\xi} = (E\xi_1, \dots, E\xi_n)$

Def. Дисперсией (или матрицей ковариаций) случайного вектора называется матрица $D\vec{\xi} = E(\vec{\xi} - E\vec{\xi})^T \cdot (\vec{\xi} - E\vec{\xi})$, состоящая из элементов $d_{i,j} = (\xi_i, \xi_j)$. В частности $d_{i,i} = (\xi_i, \xi_i) = D\xi_i$

Функции от двух случайных величин

Th. Пусть ξ_1, ξ_2 - случайные величины с общем плотностью $f_{\xi_1,\xi_2}(x, y)$, и есть функция $g(x, y) \ : \ \mathbb{R}^2 \to \mathbb{R}$. Тогда случайная величина $\eta = g(\xi_1, \xi_2)$ имеет функцию распределения $F_\eta(z) = \iint_{D_z} f(x, y)dxdy$, где $D_z = \{(x, y) \in \mathbb{R}^2 \mid g(x, y) < z\}$

$$F_\eta = p(\eta < z) = p(g(\xi_1, \xi_2) < z) = p((\xi_1, \xi_2) \in D_z) = \iint_{D_z} f(x, y)dxdy$$

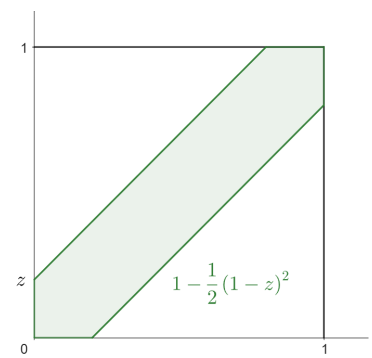
Ex. Задача о встрече. двое договорились встретится между 12:00 и 13:00. Случайная величина η - время ожидания. Найти функцию распределения

ξ_1 - время прихода первого, ξ_2 - второго; $\xi_1, \xi_2 \in U(0, 1)$, они независимы, $\forall x, y \in [0, 1] \ f_{\xi_1}(x) = 1, f_{\xi_2}(y) = 1$

Поэтому $f_{\xi_1,\xi_2}(x, y) = f_{\xi_1}(x)f_{\xi_2}(y) = 1, (x, y) \in [0, 1] \times [0, 1]$

$\eta = |\xi_1 - \xi_2| \Longrightarrow D_z = \{(x, y) \in \mathbb{R}^2 \mid |x - y| < z\}$

$F_\eta = \iint_{D_z} f_{\xi_1,\xi_2}(x, y)dxdy = \iint_{D_z} dxdy = 1 - 2 \cdot \frac{1}{2}(1 - z)^2 = 2z - z^2, \ z \in [0, 1]$



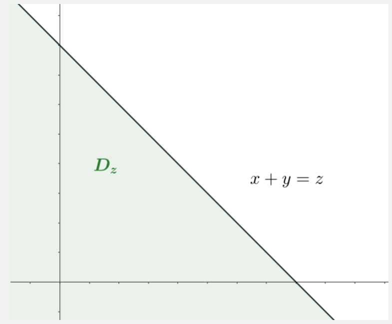
Th. $\exists \xi_1, \xi_2$ - независимые абсолютно непрерывные случайные величины с плотностями $f_{\xi_1}(x)$ и $f_{\xi_2}(y)$

Тогда плотность суммы $\xi_1 + \xi_2$ равна $f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} \underbrace{f_{\xi_1}(x)f_{\xi_2}(t - x)dx}_{\text{т. н. свертка}}$

Так как случайные величины ξ_1 и ξ_2 независимы, то $f_{\xi_1,\xi_2}(x,y) = f_{\xi_1}(x)f_{\xi_2}(y)$

И согласно предыдущей теореме $F_{\xi_1+\xi_2}(z) = \iint_{D_z} f_{\xi_1,\xi_2}(x,y)dxdy = \iint_{D_z} f_{\xi_1}(x)f_{\xi_2}(y)dxdy$,

где $D_z = \{(x,y) \in \mathbb{R}^2 \mid x+y<z\}$



$F_{\xi_1+\xi_2}(z) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{z-x} f_{\xi_1}(x)f_{\xi_2}(y)dy =$

$\begin{bmatrix} y = t - x; & dy = dt; & t = y + x \\ t(-\infty) = -\infty; & t(z-x) = z & \end{bmatrix} =$

$\int_{-\infty}^{\infty} f_{\xi_1}(x)dx \int_{-\infty}^{z} f_{\xi_2}(t-x)dt =$

$\int_{-\infty}^{z}\left(\int_{-\infty}^{\infty} f_{\xi_1}(x)f_{\xi_2}(t-x)dx\right)dt \Longrightarrow f_{\xi_1+\xi_2}(t) = \int_{-\infty}^{\infty} f_{\xi_1}(x)f_{\xi_2}(t-x)dx$

Следствие: сумма двух независимых абсолютно непрерывных случайных величин также имеет абсолютно непрерывное распределение

Nota. Условие независимости существенно, контр-пример: $\xi_1; \xi_2 = -\xi_1$, тогда $\xi_1 + \xi_2 \equiv 0$

Сумма стандартных распределений. Устойчивость относительно суммирования

Def. Если сумма двух независимых случайных величин одного типа распределения также будет этого же типа, то говорят, что распределение устойчиво относительно суммирования

Ex. 1. $\xi \in B_{n,p}; \eta \in B_{m,p}$. Тогда ясно, что $\xi + \eta \in B_{n+m,p}$ (по определению биномиального распределения $B_{n,p}$ - число успехов из n испытаний, где p - вероятность успеха)

Ex. 2. $\xi \in \Pi_\lambda, \eta \in \Pi_\mu$, они независимы. Тогда $\xi + \eta \in \Pi_{\lambda+\mu}$

$\xi + \eta = 0, 1, 2, 3, \ldots \quad \exists k \geq 0$. Тогда $p(\xi+\eta=k) = \sum_{i=0}^{k} P(\xi=i, \eta=k-i) = \sum_{i=0}^{k} P(\xi=i)P(\eta=k-i) = \sum_{i=0}^{k} \frac{\lambda^i}{i!}e^{-\lambda}\frac{\mu^{k-i}}{(k-i)!}e^{-\mu} = e^{-\lambda-\mu}\sum_{i=0}^{k}\frac{\lambda^i\mu^{k-i}}{i!(k-i)!} = e^{-\lambda-\mu}\frac{1}{k!}\sum_{i=0}^{k}\frac{\lambda^i\mu^{k-i}k!}{i!(k-i)!} = e^{-\lambda-\mu}\frac{1}{k!}\sum_{i=0}^{k}\lambda^i\mu^{k-i}C_k^i = e^{-\lambda-\mu}\frac{(\lambda+\mu)^k}{k!} \Longrightarrow \xi+\eta \in \Pi_{\lambda+\mu}$

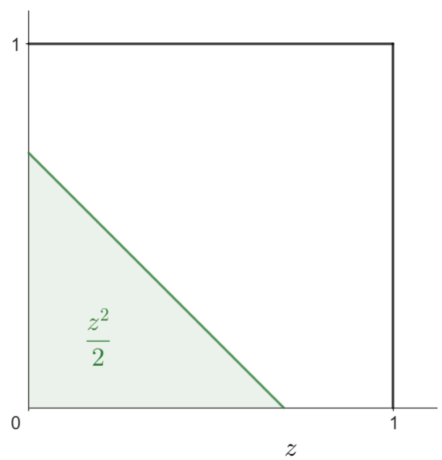
Ex. 3. $\xi, \eta \in N(0,1)$ и независимы. Тогда $\xi + \eta \in N(0,2)$

$f_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}; f_\eta(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$

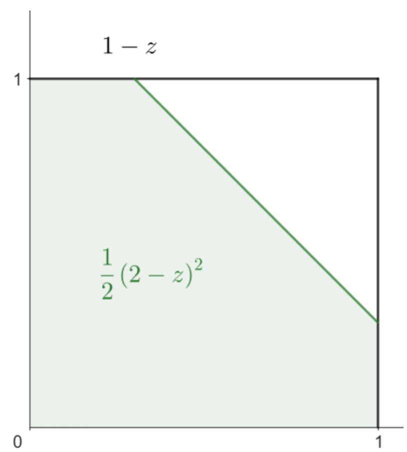
По формуле свертки $f_{\xi+\eta}(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-x)^2}{2}} = \frac{1}{2\pi} \int e^{-(x^2-tx+\frac{t^2}{2})} = \frac{1}{2\pi} e^{-\frac{t^2}{4}} \int_{-\infty}^{\infty} e^{-(x^2-tx+\frac{t^2}{4})} dx = \frac{1}{2\pi} e^{-\frac{t^2}{4}} \int_{-\infty}^{\infty} e^{-(x-\frac{t}{2})^2} d(x-\frac{t}{2}) = \frac{1}{2\pi} e^{-\frac{t^2}{4}} \sqrt{\pi} = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{t^2}{2(\sqrt{2})^2}} \Longrightarrow$ $\xi+\eta \in N(0,2)$

Ex. 4. В общности для независимых $\xi \in N(a_1, \sigma_1^2), \eta \in N(a_2, \sigma_2^2)$ $\xi+\eta \in N(a_1+a_2, \sigma_1^2+\sigma_2^2)$

Ex. 5. Равномерное распределение неустойчиво относительно суммирования, контрпример:
$\xi, \eta \in U(0,1)$ - независимы
$\forall x, y \in [0,1]$ $f_\xi(x) = 1, f_\eta(y) = 1$ и $f_{\xi,\eta}(x,y) = 1$
По первой теореме $F_{\xi,\eta}(x,y) = \iint_{D_z} f_{\xi,\eta}(x,y)dxdy = \iint_{D_z} dxdy = S_{D_z}$, где $D_z = \{(x,y) \mid x+y<z\}$



а) $0 < z \le 1$



б) $1 < z \le 2$

$$S_{D_z} = \begin{cases} 0, & z < 0 \\ \frac{z^2}{2}, & 0 \le z \le 2 \\ 1 - \frac{1}{2}(2-z)^2 & 1 \le z \le 2 \\ 0, & z > 2 \end{cases}$$

$$f_{\xi+\eta}(z) = \begin{cases} 0, & z < 0 \\ z, & 0 \le z \le 2 \\ 2-z & 1 \le z \le 2 \\ 0, & z > 2 \end{cases} \not\equiv C \Longrightarrow \xi+\eta \text{ не имеют равномерное распределение}$$

Nota. FUN FACT: сумма нескольких величин с равномерным распределением приближается к нормальному распределению

Условное распределение

Def. Условным распределением случайной величины из системы случайных величин (ξ, η) называется ее распределение, найденное при условии, что другая случайная величина приняла определенное значение. Обозначается $\xi | \eta = y$

Def. A.: Условным математическим ожиданием (обозначается $E(\xi|\eta = y)$) называется математическим ожиданием случайной величины ξ при соответствующем условном распределении

I. Условное распределение в дискретной системе двух случайных величин

Пусть (ξ, η) задана законом распределения:

$\xi \backslash \eta$	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}

Формула условной вероятности: $P(A|B) = \dfrac{P(AB)}{P(B)}$

Вероятности условных распределений считаем по формулам:

$\xi|\eta = y_j$: $p_i = p(\xi = x_i \mid \eta = y_j) = \dfrac{p(\xi = x_i, \eta = y_j)}{p(\eta = y_j)} = \dfrac{p_{ij}}{q_j} = \dfrac{p_{ij}}{\sum_i p_{ij}}$

$\eta|\xi = x_i$: $q_j = p(\eta = y_j \mid \xi = x_i) = \dfrac{p(\xi = x_i, \eta = y_j)}{p(\xi = x_i)} = \dfrac{q_{ij}}{p_i} = \dfrac{q_{ij}}{\sum_j p_{ij}}$

То есть вероятность в соответствующем столбце/строке делим на суммарную вероятность по строке или столбцу, в зависимости от того, какое условие мы рассматриваем

II. Условное распределение в непрерывной системе двух случайных величин

Пусть (ξ, η) задана плотностью $f_{\xi,\eta}(x, y)$ совместного распределения, тогда плотность условного распределения $\xi|\eta = y$:

$f(x|y) = \dfrac{f_{\xi,\eta}(x, y)}{\int_{\mathbb{R}} f_{\xi,\eta}(x, y)dx} = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$

Def. Функция $f(x|y) = \dfrac{f_{\xi,\eta}(x, y)}{f_\eta(y)}$ называется условной плотностью

Def. Условное математические ожидание вычисляется по формуле $E(\xi|\eta = y) = \displaystyle\int_{-\infty}^{\infty} x f(x|y)dx$

Аналогично $E(\eta|\xi = x) = \int_{-\infty}^{\infty} yf(y|x)dy$

Nota. При фиксированном значении x $f(y|x)$ зависит только от y, а $E(\eta|\xi = x) \in \mathbb{R}$. Если рассматривать x как переменную, то условное математическое ожидание $E(\eta|\xi = x)$ является функцией от x и называется функцией регрессии η на ξ. График такой функции называют линией регрессии

Nota. Так как значение x - значение случайной величины ξ, то условное матожидание $E(\eta|\xi = x)$ можно рассматривать как случайную величину