Лекция 9. Условные законы распределения. Многомерное нормальное распределение

Велищанский Михаил Александрович

Московский Государственный Технический Университет имени Н.Э. Баумана

Москва, 2023

Условные законы распределения. Условный ряд распределения

Пусть (X, Y) — двумерная дискретная СВ, X и Y принимают значения x_i $i = \overline{1, n}$ и y_j $j = \overline{1, m}$ соответственно.
Тогда закон распределения можно задать набором вероятностей $p_{ij} = P\{X = x_i, Y = y_j\}$, $\forall i, j$.
Законы распределения каждой координаты имеют вид

$$p_{X_i} = P\{X = x_i\} = \sum_{j=1}^{m} p_{ij}, \quad p_{Y_j} = P\{Y = y_j\} = \sum_{i=1}^{n} p_{ij}.$$

Определение

Для двумерной дискретной случайной величины (X, Y) **условной вероятностью** π_{ij} $i = \overline{1, n}$, $j = \overline{1, m}$ того, что случайная величина X примет значение x_i при условии $Y = y_j$, называют условную вероятность события $\{X = x_i\}$ при условии $\{Y = y_j\}$:

$$\pi_{ij} = \mathsf{P}\{X = x_i | Y = y_j\} = \frac{\mathsf{P}\{X = x_i, Y = y_j\}}{\mathsf{P}\{Y = y_j\}} = \frac{p_{ij}}{p_{Y_j}}.$$

Условные законы распределения. Условный ряд распределения

Определение

Набор вероятностей π_{ij}, $i = \overline{1, n}$, характеризует **условное распределение** *дискретной случайной величины X при условии $Y = y_j$.*

Аналогично определяют условную вероятность π^*_{ij} того, что случайная величина Y примет значение y_j при условии $X = x_i$:

$$\pi^*_{ij} = P\{Y = y_j | X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{X_i}}.$$

Обычно условное распределение дискретной СВ X при условии, что дискретная СВ Y примет все возможные значения задают при помощи таблицы, аналогичной таблице для задания закона распределения двумерного дискретного случайного вектора.

Условная функция распределения

Ввести условную функцию распределения случайной величины X при условии $Y = y$ по формуле

$$F_X(x|Y = y) = \frac{P\{X < x, Y = y\}}{P\{Y = y\}}$$

не представляется возможным, т.к., к примеру, $P\{Y = y\} = 0$ для непрерывной случайной величины Y.

Поэтому для непрерывной СВ вместо события $Y = y$ рассматривают событие $y \leqslant Y < y + \Delta y$ при $\Delta y \to 0$.
Пусть *случайный вектор* (X, Y) имеет непрерывную *совместную плотность распределения* $f(x, y)$ и, следовательно, *маргинальные плотности распределения* СВ X и Y

$$f_X(x) = \int\limits_{-\infty}^{+\infty} f(x, y)dy \text{ и } f_Y(y) = \int\limits_{-\infty}^{+\infty} f(x, y)dx,$$

которые также будем считать непрерывными.

Определим условную вероятность $\{X < x\}$ при условии $\{y \leqslant Y < y + \Delta y\}$ как

$$\mathsf{P}\{X < x | y \leqslant Y < y + \Delta y\} = \frac{\mathsf{P}\{X < x, y \leqslant Y < y + \Delta y\}}{\mathsf{P}\{y \leqslant Y < y + \Delta y\}} =$$

$$= \frac{F(x, y + \Delta y) - F(x, y)}{F_Y(y + \Delta y) - F_Y(y)} = \frac{\int\limits_{y}^{y+\Delta y} dv \int\limits_{-\infty}^{x} f(u, v) du}{\int\limits_{y}^{y+\Delta y} f_Y(v) dv}$$

При сделанных предположений функция $\int\limits_{-\infty}^{x} f(u, v) du$ является непрерывной. Поэтому, согласно теореме о среднем значении,

$$\int\limits_{y}^{y+\Delta y} dv \int\limits_{-\infty}^{x} f(u, v) du = \Delta y \int\limits_{-\infty}^{x} f(u, \xi) du, \quad \int\limits_{y}^{y+\Delta y} f_Y(v) dv = f_Y(\eta) \Delta y$$

Условная функция распределения

и, следовательно,

$$\mathbf{P}\{X < x | y \leqslant Y < y + \Delta y\} = \frac{\int\limits_{-\infty}^{x} f(u, \xi) du}{f_Y(\eta)},$$

где ξ и η — некоторые числа, заключенные между y и $y + \Delta y$.
Тогда **условная функция распределения** имеет вид:

$$F_X(x | Y = y) = \lim_{\Delta y \to 0} \mathbf{P}\{X < x | y \leqslant Y < y + \Delta y\} =$$
$$= \lim_{\Delta y \to 0} \frac{\int\limits_{-\infty}^{x} f(u, \xi) du}{f_Y(\eta)}.$$

Таким образом, по определению, имеем

$$F_X(x | Y = y) = \frac{1}{f_Y(y)} \int\limits_{-\infty}^{x} f(u, y) du.$$

Условная функция плотности распределения

Определение

Условной плотностью распределения *случайной величины* X, *координаты случайного вектора* (X, Y), *при условии, что случайная величина* Y *приняла фиксированное значение* y, *называют функцию* $f_X(x|y)$, *определяемую соотношением*

$$f_X(x|y) = \frac{f(x, y)}{f_Y(y)}.$$

Определение

Условное распределение (для дискретной случайной величины), условная функция распределения и условная плотность распределения (для непрерывных случайных величин) называют **условными законами распределения**.

Пример

Пусть случайные величины X_1 и X_2 — координаты точки падения частицы, случайно брошенной в круг радиуса R с центром в начале координат.
Случайный вектор (X_1, X_2) имеет плотность распределения

$$f(x_1, x_2) = \begin{cases} 0, & x_1^2 + x_2^2 > R^2; \\ \dfrac{1}{\pi R^2}, & x_1^2 + x_2^2 \leqslant R^2. \end{cases}$$

Найдем условную плотность распределения СВ X_1 при условии $X_2 = x_2$.
Маргинальная плотность распределения $f_{X_2}(x_2)$ имеет вид

$$f_{X_2}(x_2) = \begin{cases} 0, & |x_2| > R; \\ \dfrac{2\sqrt{R^2 - x_2^2}}{\pi R^2}, & |x_2| \leqslant R. \end{cases}$$

Пример (продолжение)

Тогда, при $|x_2| \leqslant R$, имеем

$$f_{X_1}(x_1|x_2) = \frac{f(x_1, x_2)}{f_{X_2}(x_2)} = \begin{cases} 0, & |x_1| > \sqrt{R^2 - x_2^2}; \\ \dfrac{1}{2\sqrt{R^2 - x_2^2}}, & |x_1| \leqslant \sqrt{R^2 - x_2^2}. \end{cases}$$

Следовательно, случайная величина X_1 при условии $X_2 = x_2$ равномерно распределена на $\left[-\sqrt{R^2 - x_2^2}, \sqrt{R^2 - x_2^2}\right]$.

Если $|x_2| > R$ то условная плотность распределения $f_{X_1}(x_1|x_2)$ не определена, однако СВ X_2 не может быть больше R по модулю.

Пусть координаты X_1 и X_2 случайного вектора (X_1, X_2) являются случайными величинами, распределенными по нормальному закону, т.е.

$$f_{X_1}(x) = \varphi_{m_1,\sigma_1}(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-m_1)^2}{2\sigma_1^2}\right),$$

$$f_{X_2}(x) = \varphi_{m_2,\sigma_2}(x) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x-m_2)^2}{2\sigma_2^2}\right).$$

Если X_1 и X_2 — независимые случайные величины, то $f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$ и плотность **двумерного нормального распределения** имеет вид

$$f_{X_1 X_2}(x_1, x_2) = \frac{1}{(\sqrt{2\pi})^2 \sigma_1 \sigma_2} \exp\left(-\frac{(x_1-m_1)^2}{2\sigma_1^2} - \frac{(x_2-m_2)^2}{2\sigma_2^2}\right).$$

Двумерное нормальное распределение

Вектор $\overrightarrow{X} = (X_1, X_2)$ имеет (невырожденное) **двумерное нормальное распределение**, если его плотность распределения имеет вид

$$f_{X_1X_2}(x_1, x_2) = \frac{1}{(\sqrt{2\pi})^2\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2}Q(x_1-m_1,\, x_2-m_2)},$$

где

$$Q(y_1, y_2) = \frac{1}{1-\rho^2}\left(\frac{y_1^2}{\sigma_1^2} - \frac{2\rho y_1 y_2}{\sigma_1\sigma_2} + \frac{y_2^2}{\sigma_2^2}\right)$$

— положительно определенная квадратичная форма, т.е $Q(y_1, y_2) > 0\ \forall (y_1, y_2) \in \mathbb{R}^2 :\ (y_1, y_2) \neq (0, 0))$.
Можно показать:
$m_1 = MX_1$, $m_2 = MX_2$, $\sigma_1^2 = DX_1$, $\sigma_2^2 = DX_2$, $\rho = \dfrac{\operatorname{cov}(X_1, X_2)}{\sigma_1\sigma_2}$.

Двумерное нормальное распределение

Представим $Q(y_1, y_2)$ в матричной форме $Q(\vec{y}) = \vec{y}\,\Sigma^{-1}\,\vec{y}^{\,T}$, где

$\vec{y} = (y_1, y_2)$, $\Sigma^{-1} = \dfrac{1}{1-\rho^2}\begin{pmatrix} \dfrac{1}{\sigma_1^2} & -\dfrac{\rho}{\sigma_1\sigma_2} \\ -\dfrac{\rho}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_2^2} \end{pmatrix}$ — обратная

матрица к ковариационной матрице $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$.

Поскольку $\det\Sigma = \sigma_1^2\sigma_2^2(1-\rho^2)$ то окончательно получаем:

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^2(\det\Sigma)^{\frac{1}{2}}}e^{-\frac{1}{2}(\vec{x}-\vec{m})\Sigma^{-1}(\vec{x}-\vec{m})^T},$$

где $\vec{x} = (x_1, x_2)$, $\vec{m} = (m_1, m_2)$.

Определение

Случайный вектор $\vec{X} = (X_1, \ldots, X_n)$ *имеет* **многомерное нормальное распределение** *с вектором математических ожиданий* $\vec{m} = (m_1, \ldots, m_n)$ *и ковариационной матрицей* $\Sigma = (\sigma_{ij})$ $i, j = \overline{1, n}$, *если его функция плотности имеет вид*

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(\sqrt{2\pi})^n (\det \Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-\vec{m})\Sigma^{-1}(\vec{x}-\vec{m})^T}.$$

Если $\Sigma = \Sigma^{-1} = E$ *и* $\vec{m} = (0, \ldots, 0)$, *то*

$$f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}(x_1^2 + \ldots + x_n^2)}.$$

Такую функцию плотности называют **плотностью стандартного многомерного нормального распределения.**

Многомерное нормальное распределение. Геометрическая интерпретация плотности нормального распределения

В 2-мерном случае функция $f_{X_1X_2}(x_1, x_2)$ задает некоторую поверхность в $\mathbb{R}^3$.

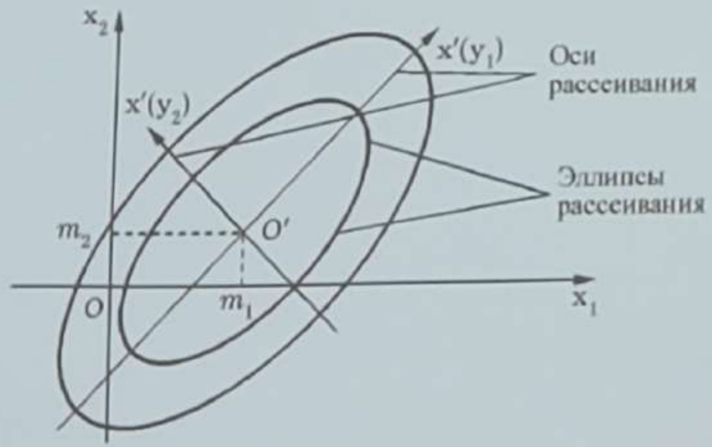

Рис. 1 : Линии уровня $f_{X_1X_2}(x_1, x_2) = a$

Линии уровня данной поверхности имеют вид: $f_{X_1X_2}(x_1, x_2) = a$ или $Q(x_1 - m_1, x_2 - m_2) = b$, где $b = -2\ln\{2\pi a(\det\Sigma)^{\frac{1}{2}}\}$.

Уравнение $Q(x_1 - m_1, x_2 - m_2) = b$ представляют собой семейство эллипсов при разных значениях b.

Оси симметрии $O'x_1'$ и $O'x_2'$ проходят через т. $O'(m_1, m_2)$, их направления совпадают с направлениями собственных векторов e_i $i = 1, 2$ матрицы Σ^{-1}, соответствующие собственным значениям λ_i.

Многомерное нормальное распределение. Геометрическая интерпретация плотности нормального распределения

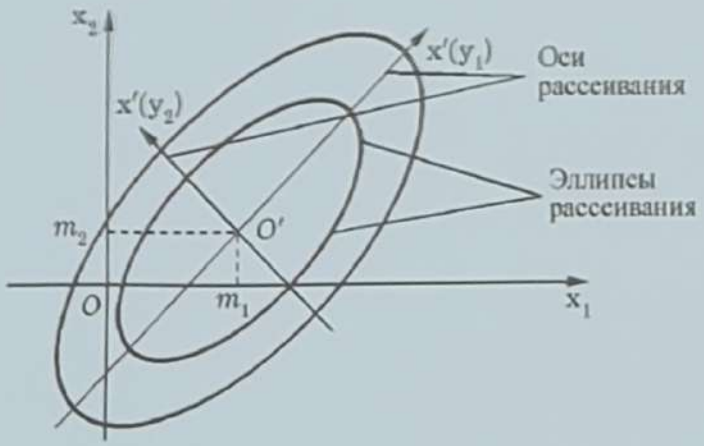

Рис. 2 : Линии уровня $f_{X_1X_2}(x_1, x_2) = a$

В системе координат $O'x_1'x_2'$ случайный вектор (X_1', X_2') имеет нормальное распределение с $\vec{m}' = (0, 0)$ и $\Sigma' = \begin{pmatrix} \sigma_1'^2 & 0 \\ 0 & \sigma_2'^2 \end{pmatrix}$,

где $\sigma_1' = \dfrac{1}{\sqrt{\lambda_1}}$, $\sigma_2' = \dfrac{1}{\sqrt{\lambda_2}}$.

Если ввести новую систему координат $x_i'' = \sigma_i' x_i'$ $i = 1, 2$, то в в $O'x_1''x_2''$ случайный вектор (X_1'', X_2'') будет иметь двумерное стандартное нормальное распределение.

В случае $n > 2$ $f_{\vec{X}}(\vec{x}) = a$ в силу положительной определенности матрица Σ задает семейство n-мерных эллипсоидов, называемых **эллипсоидами рассеивания**, с осями симметрии называемыми **осями рассеивания**.

Свойства плотности многомерного нормального распределения

Теорема

1. *Закон распределения координаты X_i $i = \overline{1, n}$ случайного вектора $\vec{X}$, имеющего n-мерное нормальное распределение с $\vec{m} = (m_1, \dots, m_n)$ и $\Sigma = (\sigma_{ij})$, является нормальным с параметрами m_i и σ_i.*
2. *Если ковариационная матрица Σ случайно вектора $\vec{X}$, распределенного по невырожденному нормальному закону, является диагональной, то координаты вектора $X_1, \dots, X_n$ являются независимыми случайными величинами.*
3. *Если вектор $\vec{X} = (X_1, \dots, X_n)$ имеет нормальный закон распределения с $\vec{m} = (m_1, \dots, m_n)$ и матрицей ковариаций Σ, то $\vec{X}' = (X_1, \dots, X_{n-1})$ также распределен по нормальному закону с $\vec{m}' = (m_1, \dots, m_{n-1})$ и Σ', полученной из Σ вычеркиванием последних строки и столбца.*

Пример

Пусть двумерный случайный вектор (X, Y) *имеет нормальное распределение с* (m_1, m_2) *и*
$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \qquad (\sigma_1,\ \sigma_2 > 0, \quad -1 < \rho < 1).$
Найдем условную плотность распределения случайной величины X *при условии* $Y = y$.

Так как

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-m_1)^2}{\sigma_1^2} - \right.\right.$$

$$\left.\left. - \frac{2\rho(x-m_1)(y-m_2)}{\sigma_1\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2}\right)\right\},$$

$$f_Y(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(y-m_2)^2/(2\sigma_2^2)},$$

Пример (продолжение)

то

$$f_X(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{1}{\sigma_1\sqrt{2\pi(1-\rho^2)}} \exp\left\{-\frac{1}{2\sigma_1^2(1-\rho^2)}\left[x- \right.\right.$$

$$\left.\left. - \left(m_1 + \frac{\rho\sigma_1(y-m_2)}{\sigma_2}\right)\right]^2\right\}.$$

Таким образом, условное распределение X при условии $Y = y$ также является нормальным распределением с математическим ожиданием $g(y) = m_1 + \rho\dfrac{\sigma_1}{\sigma_2}(y - m_2)$ и средним квадратичным отклонением $\sigma_{X|y} = \sigma_1\sqrt{1-\rho^2}$. Аналогично условное распределение Y при условии $X = x$ является нормальным с математическим ожиданием $h(x) = m_2 + \rho\dfrac{\sigma_2}{\sigma_1}(x - m_1)$ и средним квадратичным отклонением $\sigma_{Y|x} = \sigma_2\sqrt{1-\rho^2}$.

Определение

Функцию $g(y) = M(X|y)$ *называют* **функцией регрессии**, *или просто* **регрессией**, *случайной величины* X *на случайную величину* Y, *а ее график — линией регрессии случайной величины* X *на случайную величину* Y, *или* X *на* Y.