Типовой расчет по Теории Вероятностей

Зенкова Дарья, M3336

Вариант 2

1 Функции случайных величин

1.1 Задача

Функция распределения $F_X(t)$ случайной величины X имеет вид:

$$F_X(t) = \begin{cases} 1 - e^{-t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Случайные величины $Y = X^2$ и $Z = -3X + 2$ являются функциями от случайной величины X. Найти:

а) плотность распределения $f_y(v)$ случайной величины Y;

б) моменты $\mathbf{E}Z$, $\mathbf{D}Z$, $\mathbf{cov}(X, Z)$.

1.2 Решение

а) Пусть D_v - множество значений случайной величины X, для которых $y = g(x) = x^2 < v$: $D_v = \{x : y = x^2 < v\}$.
При $v \leq 0$ $D_v = \emptyset \implies F_Y(v) = \mathbf{P}\{Y \in D_v\} = 0 \implies f_Y(v) = 0$ при $v \leq 0.77$ При $v > 0$ D_v имеет вид:

$$D_v = \{x : x^2 < v\} = \{x : -\sqrt{v} < x < \sqrt{v}\}$$

Тогда функция распределения Y:

$$F_Y(v) = \mathbf{P}\{-\sqrt{v} < X < \sqrt{v}\} = F_X(\sqrt{v}) - F_X(-\sqrt{v}) = 1 - e^{\sqrt{v}}$$

Дифференцируя по v, получим плотность распределения Y:

$$f_Y(v) = (1 - e^{\sqrt{v}})' = -e^{\sqrt{v}} \cdot \frac{1}{2\sqrt{v}}$$

Таким образом,

$$f_Y(v) = \begin{cases} -\frac{e^{\sqrt{v}}}{2\sqrt{v}}, & v > 0 \\ 0, & v \leq 0 \end{cases}$$

б) Пусть $Z = -3X + 2$. Для начала найдем плотность распределения X:

$$f_X(t) = \begin{cases} e^{-t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Математическое ожидание случайной величины X:

$$\mathbf{E}X = \int\limits_{-\infty}^{\infty} t f_X(t) dt = \int\limits_{0}^{\infty} t e^{-t} dt = 1$$

Математическое ожидание X^2:

$$\mathbf{E}(X^2) = \int\limits_0^\infty t^2 e^{-t} dt = 2$$

Дисперсия случайной величины X:

$$\mathbf{D}X = \mathbf{E}(X^2) - (\mathbf{E}X)^2 = 2 - 1 = 1$$

Воспользуемся линейностью математического ожидания, чтобы найти $\mathbf{E}Z$:

$$\mathbf{E}Z = \mathbf{E}(-3X + 2) = -3\mathbf{E}X + 2 = -1$$

. Дисперсия Z:

$$\mathbf{D}Z = \mathbf{D}(-3X + 2) = 9\mathbf{D}X = 9$$

Воспользуемся следующим свойством ковариации:

$$\mathbf{cov}(aX + b, cY + d) = ab\mathbf{cov}(X, Y)$$

Тогда ковариация случайных величин X и Z:

$$\mathbf{cov}(X, Z) = \mathbf{cov}(X, -3X + 2) = -3\mathbf{cov}(X, X) = -3\mathbf{D}X = -3$$

1.3 Ответ

а)

$$f_Y(v) = \begin{cases} -\frac{e^{\sqrt{v}}}{2\sqrt{v}}, & v > 0 \\ 0, & v \leq 0 \end{cases}$$

б) $\mathbf{E}Z = -1$, $\mathbf{D}Z = 9$, $\mathbf{cov}(X, Z) = -3$

2 Регрессия. Случай дискретных случайных величин

2.1 Задача

Случайный вектор (X, Y) задан матрией распределения:

Y	X			
	-2	0	2	4
-2	0.05	0.05	0.2	0.1
2	0.25	0.15	0.1	0.1

Найти условные ожидания $\mathbf{E}(X|Y)$ и $\mathbf{E}(Y|X)$, проверить формулу полного математического ожидания. Построить линейную регрессию X на Y и Y на X и вычислить значения этих функций в точках x_i и y_j. Геометрически сравнить значения регрессии и линейной регрессии.

2.2 Решение

$\mathbf{E}(Y|X) = \mathbf{E}(Y|T = \{X^{-1}(x_j)\}) = \sum_j \mathbf{E}(Y|X^{-1}(x_j))I_{X^{-1}(x_j)} = \sum_j \mathbf{E}(Y|X = x_j)I_{C_j}$,

где $T = \{C_j\}_j$, $C_j = X^{-1}(x_j) = \{\omega : X(\omega) = x_j\}$, $x_i \neq x_j \ \forall i \neq j$. Найдем условное математическое ожидание $\mathbf{E}(Y|X) = m_{Y|X}(x)$:

$$\mathbf{E}(Y|X = x_i) = y_1 \cdot \frac{\mathbf{P}(Y = y_1|X = x_i)}{\mathbf{P}(X = x_i)} + y_2 \cdot \frac{\mathbf{P}(Y = y_2|X = x_i)}{\mathbf{P}(X = x_i)}$$

$$\mathbf{E}(Y|X = -2) = -2 \cdot \frac{0.05}{0.05 + 0.25} + 2 \cdot \frac{0.25}{0.05 + 0.25} = \frac{4}{3}$$

$$\mathbf{E}(Y|X = 0) = 1$$

$$\mathbf{E}(Y|X = 2) = -\frac{2}{3}$$

$$\mathbf{E}(Y|X = 4) = 0$$

$$\mathbf{E}(Y|X) = \frac{4}{3}I_{X=-2} + I_{X=0} - \frac{2}{3}I_{X=2}$$

Найдем условное математическое ожидание $\mathbf{E}(X|Y) = m_{X|Y}(y)$:

$$\mathbf{E}(X|Y = -2) = \frac{7}{4}$$

$$\mathbf{E}(X|Y = 2) = \frac{1}{6}$$

$$\mathbf{E}(X|Y) = \frac{7}{4}I_{Y=-2} + \frac{1}{6}I_{Y=2}$$

Def $g^*(x)$ - функция линейной регрессии Y на X, если:

$$\inf_{g(x)=ax+b} \mathbf{E}(Y - g(X))^2 = \mathbf{E}(Y - g^*(X))^2$$

$$g^*(x) = a^*x + b^*$$

$$a^* = \frac{\mathbf{cov}(X, Y)}{\mathbf{D}X},\ b^* = \mathbf{E}Y - a^*\mathbf{E}X$$

Пусть $g(x) = ax + b$ - линейная регрессия Y на X, $h(y) = cy + d$ - линейная регрессия X на Y. Для начала найдем все необходимые моменты:

$$\mathbf{cov}(X, Y) = \mathbf{E}((X - \mathbf{E}X)(Y - \mathbf{E}Y)) = \mathbf{E}(XY) - \mathbf{E}X \cdot \mathbf{E}Y$$

$$\mathbf{E}(XY) = \sum_{i=1}^{4}\sum_{j=1}^{2} x_i y_j \mathbf{P}(X = x_i, Y = y_j) = -1.2$$

$$\mathbf{E}X = \sum_{i=1}^{4} x_i \mathbf{P}(X = x_i) = 0.8$$

$$\mathbf{E}Y = \sum_{j=1}^{2} y_j \mathbf{P}(Y = y_j) = 0.4$$

$$\mathbf{cov}(X, Y) = -1.2 - 0.8 \cdot 0.4 = -1.52$$

$$\mathbf{E}(X^2) = 5.6$$

$$\mathbf{E}(Y^2) = 4$$

$$\mathbf{D}X = \mathbf{E}(X^2) - (\mathbf{E}X)^2 = 5.6 - 0.64 = 4.96$$

$$\mathbf{D}Y = \mathbf{E}(Y^2) - (\mathbf{E}Y)^2 = 4 - 0.16 = 3.84$$

Теперь можем найти все коэффициенты:

$$a = \frac{\mathbf{cov}(X, Y)}{\mathbf{D}X} = \frac{-1.52}{4.96} = -\frac{19}{62}$$

$$b = \mathbf{E}Y - a\mathbf{E}X = 0.4 + \frac{19}{62} \cdot 0.8 = \frac{20}{31}$$

$$c = \frac{\mathbf{cov}(X, Y)}{\mathbf{D}Y} = \frac{-1.52}{3.84} = -\frac{19}{48}$$

$$d = \mathbf{E}X - c\mathbf{E}Y = 0.8 + \frac{19}{48} \cdot 0.4 = \frac{23}{24}$$

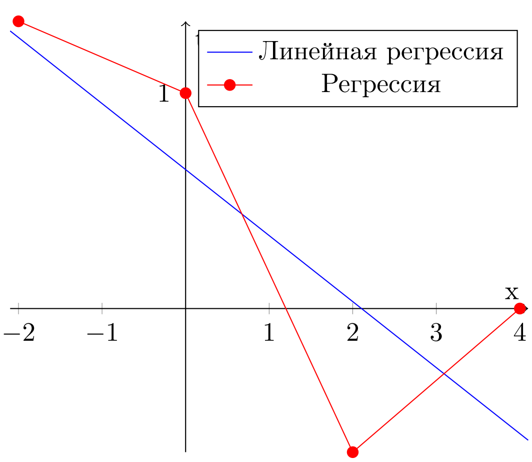
Таким образом, искомые линейные регрессии имеют следующий вид:

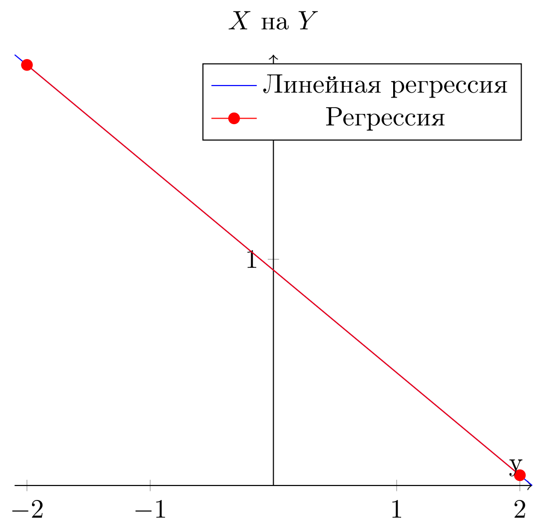
$$g(x) = -\frac{19}{62}x + \frac{20}{31}$$

$$h(y) = -\frac{19}{48}y + \frac{23}{24}$$

Сравним получившиеся регрессии и линейные регрессии.

Y на X





2.3 Ответ

$$E(Y|X) = m_{Y|X}(x) = \frac{4}{3}I_{X=-2} + I_{X=0} - \frac{2}{3}I_{X=2}$$

$$E(X|Y) = m_{X|Y}(y) = \frac{7}{4}I_{Y=-2} + \frac{1}{6}I_{Y=2}$$

$$g(x) = -\frac{19}{62}x + \frac{20}{31}$$

$$h(y) = -\frac{19}{48}y + \frac{23}{24}$$