специальные вопросы ТВиМС

часть3

показатели распределения с.в.

лекция третья

УСРЕДНЕНИЕ ФУНКЦИЙ ОТ с.в.

Рассмотрим одномерную с.в. ξ с ф.р. $F_\xi(\cdot)$, и функцию от с.в. $g(\xi)$

Одной из важнейших операций над функциями от с.в. является усреднение.

<u>Def</u> Средним значением $g(\xi)$ по распределению с.в. ξ называется значение интеграла:

$$\mathrm{E}g(\xi) = \int g(x)\, d\,\mathrm{F}_\xi(x) = \begin{cases} \int\limits_{\mathrm{X}} g(x)\,\mathrm{f}_\xi(x)\,dx \\ \sum\limits_k g(k)\,\mathrm{P}(\xi = k) \end{cases}$$

Важно:

Это не мат.ожидание $g(\xi)$!!!

Мат.ожидание $g(\xi)$ – это усреднение по собственному распределению $g(\xi)$

ХАРАКТЕРИСТИКИ ОДНОМЕРНОЙ с.в.

Def Для степенной функции $g_k(\xi)=\xi^k$ с показателем *k* усреднение называется *начальным моментом* порядка *k*: $m_k = \mathrm{E}\,\xi^k$

Def Для центрированной степенной функции с показателем *k* $g_k(\xi)=(\xi-m_1)^k$ усреднение называется *центральным моментом* порядка k: $\overset{\circ}{m}_k = \mathrm{E}\left(\xi - \mathrm{E}\,\xi\right)^k$

$m_1 -$ мат.ожидание с.в. ξ, $\overset{\circ}{m}_2 -$ дисперсия с.в. ξ

$\dfrac{\overset{\circ}{m}_3}{\overset{\circ}{m}_2^{3/2}} -$ ассиметрия с.в. ξ, $\dfrac{\overset{\circ}{m}_4}{\overset{\circ}{m}_2^{2}} - 3 -$ эксцесс с.в. ξ

Выразите 2, 3 и 4 центральный моменты через начальные моменты младших порядков

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ одномерной с.в.

Группы показателей распределения одномерной с.в.:

- ✓ Показатели центра
- ✓ Показатели неоднородности
- ✓ Показатели формы

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ одномерной с.в.

Показатели центра распределения одномерной с.в.:

1. **<u>Медиана</u>** Me: $F_\xi(Me)=0.5$
2. **<u>Мода</u>** $Mo=\arg\max f_\xi(x)$
3. **<u>Средние степенные</u>** $c_k=(m_k)^{1/k}=(E\xi^k)^{1/k}$, $k\in\mathbb{R}$

$c_1=m_1=E\xi$ - мат.ожидание (ср. арифметическое)

$c_{-1}=1/m_{-1}=1/E(1/\xi)=H\xi$ - ср. гармоническое

$c_0=G\xi=\exp[E(\ln\xi)]=\lim(m_k)^{1/k}, k\to 0$ – ср. геометрическое

Предельные случаи:

$c_{+\infty}=\sup\xi=\lim(m_k)^{1/k}$, $k\to +\infty$

$c_{-\infty}=\inf\xi=\lim(m_k)^{1/k}$, $k\to -\infty$

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ одномерной с.в

Показатели центра распределения одномерной с.в.:

Упраженения

Докажите утверждения:

$c_k=(m_k)^{1/k} \nearrow$

$c_{+\infty}=\sup\xi=\lim(m_k)^{1/k}$, $k \rightarrow +\infty$

$c_{-\infty}=\inf\xi=\lim(m_k)^{1/k}$, $k \rightarrow -\infty$

$c_0=G\xi=\exp[E(\ln\xi)]=\lim (m_k)^{1/k}$, $k \rightarrow 0$

Для случая равновероятного распределения с.в. (дискретной и непрерывной)

Для случая конечной дискретной не равновероятно распределенной с.в.

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ одномерной с.в

Показатели неоднородности распределения с.в.:

1. Дисперсия $D\xi = \sigma_\xi^2 = \overset{\circ}{m}_2 = m_2 - m_1^2$

2. СКО, коэффициент вариации $\sqrt{D\xi} = \sigma_\xi$

3. Размах $R_\xi = \sup\xi - \inf\xi$

4. Относительная вариация $V_\xi = \dfrac{\sigma_\xi}{E\xi}$

5. Относительный размах $\bar{R}_\xi = \dfrac{R_\xi}{E\xi}$

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ одномерной с.в

Показатели неоднородности распределения с.в.:

6. Квантильное расстояние $Q_\xi = u_{1-q} - u_q$

7. Относительная квантильное расстояние $\overline{Q}_\xi = \dfrac{Q_\xi}{E\xi}$

8. Квантильный коэффициент дифференциации $K_d = \dfrac{u_{1-q}}{u_q}$

9. Коэффициент Джини G=1-2 $\cdot \int_0^1 q \cdot F\left(x_q\right) dq$

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ одномерной с.в

Показатели формы распределения с.в.:

1. Коэффициент ассиметрии $As\xi = \dfrac{\overset{\circ}{m}_3}{\left(\overset{\circ}{m}_2\right)^{3/2}}$

2. Коэффициент эксцесса $Ex\xi = \dfrac{\overset{\circ}{m}_4}{\left(\overset{\circ}{m}_2\right)^{2}} - 3$

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Показатель центра:

$\boldsymbol{\xi}=(\xi_1;\ldots;\xi_n)$

$\mathbf{M}_1=(m_{11};\ldots;m_{1n})=Eg_1(\boldsymbol{\xi})=(E\xi_1;\ldots;E\xi_n)$:

многомерный первый момент, мат.ожидание многомерной с.в.

<u>Пример</u>:

$\boldsymbol{\xi}=(\xi_1;\ldots;\xi_n)$ - оценки по предметам зимней сессии

$\mathbf{M}_1=(E\xi_1;\ldots;E\xi_n)$ – средняя успеваемость

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Свойства мат.ожидания многомерной с.в. **:**

$\xi=(\xi_1;\ldots;\xi_n)$, $\mathbf{a}=(a_1;\ldots;a_n)\in\mathbb{R}^n$, $\alpha\in\mathbb{R}$, $\mathbf{A}\in\mathbb{R}^{kxn}$

1. $E(\xi+\mathbf{a})=E\xi+\mathbf{a}$
2. $E(\alpha\cdot\xi)=\alpha\cdot E\xi$
3. $E(\mathbf{A}\cdot\xi^T)=\mathbf{A}\cdot E\xi^T$
4. $E(\xi\cdot\mathbf{a}^T)=E(\mathbf{a}\cdot\xi^T)=E(\Sigma\xi_i\cdot a_i)=\Sigma E(\xi_i\cdot a_i)=\Sigma(a_i\cdot E(\xi_i))$
5. Если компоненты ξ и η попарно независимы, то:

$E(\eta^T\cdot\xi)=(E\eta)^T\cdot E\xi$ и $E(\xi\cdot\eta^T)=E\xi\cdot(E\eta)^T$

Докажите эти свойства!

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Показатель вариации:

$$\mathbf{g}_2 = (\xi - E\xi)' \cdot (\xi - E\xi) = \begin{pmatrix} g_{11}^2 & \cdots & \cdots & g_{1n}g_{11} \\ \vdots & \ddots & g_{1i}g_{1j} & \vdots \\ \vdots & g_{1j}g_{1i} & \ddots & \vdots \\ g_{11}g_{1n} & \cdots & \cdots & g_{1n}^2 \end{pmatrix}$$

$$g_{1i}g_{1j} = (\xi_i - \mathrm{E}\,\xi_i) \cdot (\xi_j - \mathrm{E}\,\xi_j)$$

$$\mathbf{V} = E\mathbf{g}_2 = E(\xi - E\xi)' \cdot (\xi - E\xi) = \begin{pmatrix} Eg_{11}^2 & \cdots & \cdots & Eg_{1n}g_{11} \\ \vdots & \ddots & Eg_{1i}g_{1j} & \vdots \\ \vdots & Eg_{1j}g_{1i} & \ddots & \vdots \\ Eg_{11}g_{1n} & \cdots & \cdots & Eg_{1n}^2 \end{pmatrix}$$

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Показатель вариации:

$$Eg_{1i}g_{1j} = E\left(\xi_i - \mathrm{E}\,\xi_i\right)\cdot\left(\xi_j - \mathrm{E}\,\xi_j\right) = \mathrm{cov}\left(\xi_i;\xi_j\right)$$

$$Eg_{1i}^2 = E\left(\xi_i - \mathrm{E}\,\xi_i\right)\cdot\left(\xi_i - \mathrm{E}\,\xi_i\right) = \mathrm{cov}\left(\xi_i;\xi_i\right) = \mathrm{var}\left(\xi_i\right)$$

$$\mathbf{V} = \mathbf{Cov}\xi = \begin{pmatrix} \mathrm{cov}\left(\xi_1;\xi_1\right) & \dots & \dots & \mathrm{cov}\left(\xi_1;\xi_n\right) \\ \vdots & \ddots & \mathrm{cov}\left(\xi_i;\xi_j\right) & \vdots \\ \vdots & \mathrm{cov}\left(\xi_j;\xi_i\right) & \ddots & \vdots \\ \mathrm{cov}\left(\xi_n;\xi_1\right) & \dots & \dots & \mathrm{cov}\left(\xi_n;\xi_n\right) \end{pmatrix}$$

$\mathbf{V}=\mathbf{Cov}(\xi)=E[(\xi-E\xi)^T(\xi-E\xi)]$ – многомерный второй центральный момент, ковариационная матрица с.в. ξ

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Показатель вариации:

$$\xi = \left(\xi_1 \ldots \xi_n\right)$$

$$\Sigma = \left(\sigma_{ij}\right) = \begin{pmatrix} \sigma_{11} & \cdots & \cdots & \sigma_{n1} \\ \vdots & \ddots & \sigma_{ji} & \vdots \\ \vdots & \sigma_{ij} & \ddots & \vdots \\ \sigma_{1n} & \cdots & \cdots & \sigma_{nn} \end{pmatrix} \quad \begin{vmatrix} \sigma_{ii} = \sigma_i^2 = \mathrm{D}\,\xi_i \\ \sigma_{ij} = \sigma_{ji} = \operatorname{cov}\left(\xi_i, \xi_j\right) \end{matrix}$$

$$r_{ij} = \frac{\sigma_{ij}}{\left(\sigma_i^2 \sigma_i^2\right)^{1/2}}$$

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Показатель вариации:

$$R = \left(r_{ij}\right) = \begin{pmatrix} 1 & \dots & \dots & r_{n1} \\ \vdots & \ddots & r_{ji} & \vdots \\ \vdots & r_{ij} & \ddots & \vdots \\ r_{1n} & \dots & \dots & r_{nn} \end{pmatrix}$$ - корреляционная матрица

$$\Sigma = \left(\sigma_{ij}\right) = \begin{pmatrix} \sigma_1^2 & \dots & \dots & r_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & r_{ji}\sigma_j\sigma_i & \vdots \\ \vdots & r_{ij}\sigma_i\sigma_j & \ddots & \vdots \\ r_{n1}\sigma_n\sigma_1 & \dots & \dots & \sigma_n^2 \end{pmatrix}$$ - ковариационная матрица

Пример:

R – показывает взаимосвязь учебных курсов

Σ - показывает неоднородность оценивания

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Свойства ковариационной и корреляционных матриц:

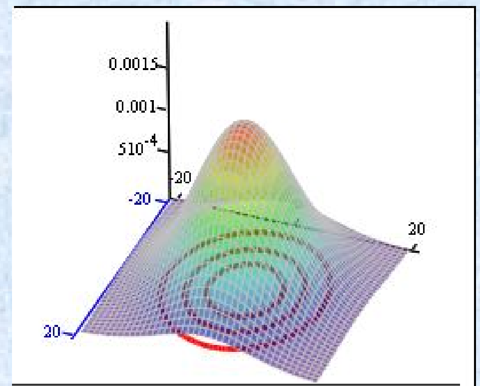
1. $\Sigma = \Sigma'$ - симметричная

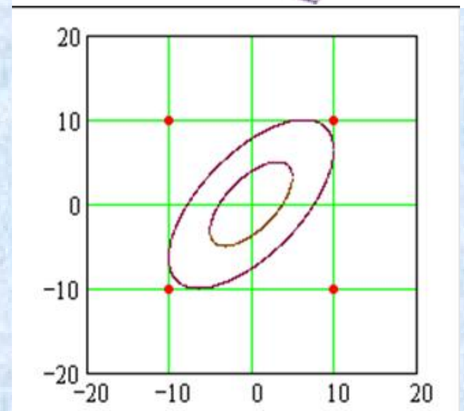
2. Σ - положительно определена

$\left(\textit{вспомнить} - \textit{что это означает} \right)$

$1., 2. \Rightarrow \det \Sigma \geq 0 \Rightarrow \exists\, \Sigma^{-1}, \Sigma^{-1/2}, \Sigma^{1/2}$

Матрица коррляций обладает теми же свойствами





$D_{общ} = \det \Sigma$ - обобщенная дисперсия

$\mathrm{tr}\Sigma = \sum_i \sigma_i^2 \geq 0$ - след матрицы

ξ_i, ξ_j - независимы $\Rightarrow \mathrm{E}\,\xi_i \xi_j = \mathrm{E}\,\xi_i\, \mathrm{E}\,\xi_i \Rightarrow$

$\Rightarrow r_{ij} = 0 \Rightarrow \Sigma -$ диагональная, $\mathrm{R} = \mathrm{I} -$ единичная

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Свойства матрицы ковариаций многомерной с.в. **:**

$\xi=(\xi_1;\ldots;\xi_n)$, $\mathbf{a}=(a_1;\ldots;a_n)\in\mathbb{R}^n$, $\mathbf{A}\in\mathbb{R}^{k\mathrm{x}n}$, $\alpha\in\mathbb{R}$

1. $V(\xi+\mathbf{a})=V\xi$
2. $V(\alpha\cdot\xi)=\alpha^2\cdot V\xi$
3. $V(\mathbf{A}\cdot\xi^T)=\mathbf{A}\cdot V\xi\cdot\mathbf{A}^T$
4. Если компоненты ξ **попарно независимы**, то:

$V(\xi\cdot\mathbf{a}^T)=V(\mathbf{a}\cdot\xi^T)=V(\Sigma(\xi_i\cdot a_i))=\Sigma V(\xi_i\cdot a_i)=\Sigma(a^2_i\cdot V(\xi_i))$

Докажите эти свойства!

ПОКАЗАТЕЛИ РАСПРЕДЕЛЕНИЯ многомерной с.в

Свойства ковариационной матрицы

Упражнение

Для двумерного случая:

1. Проверьте положительную определенность Σ
2. Найдите $\det\Sigma$, $\mathrm{tr}\Sigma$
3. Найдите Σ^{-1}, $\Sigma^{-1/2}$, $\Sigma^{1/2}$

ПОКАЗАТЕЛИ УСЛОВНОГО РАСПРЕДЕЛЕНИЯ с.в

Условное мат.ожидание и условная дисперсия

$$E(\xi_1 \mid \xi_2 = y) = \int x_1 f_{\xi_1|\xi_2=y}(x_1) dx_1$$ -

условное мат. ожидание ξ_1 при $\xi_2 = y$

$$D(\xi_1 \mid \xi_2 = y) = \int \left[x_1 - E(\xi_1 \mid \xi_2 = y)\right]^2 f_{\xi_1|\xi_2=y}(x_1) dx_1$$ -

условная дисперсия ξ_1 при $\xi_2 = y$

<u>Пример</u>:

$E(\xi_1|\xi_2)$ – средний доход в определенной группе респондентов

$D(\xi_1|\xi_2)$ – вариация дохода в данной группе респондентов

УСЛОВНОЕ МАТ.ОЖИДАНИЕ с.в

Свойства условного мат.ожидания

1. $E_{\xi_2}(\xi_1|\xi_2)=\xi_1|\xi_2$:

ξ_1 по отношению к ξ_2 является константой - нечего усреднять

2. $E_{\xi_1}(g(x_2)\cdot\xi_1|\xi_2)= g(x_2)\cdot E_{\xi_1}(\xi_1|\xi_2)$:

$g(x_2)$ по отношению к ξ_1 является константой – можно вынести за знак мат.ожидания

3. $E_{(\xi_1;\xi_2)}(h(x_1;x_2))= E_{\xi_2}[E_{\xi_1}(h(x_1;x_2)|\xi_2)]$:

усреднять можно последовательно: сперва по одной компоненте при фиксированной второй, а затем - усреднить полученный результат по второй компоненте

УСЛОВНОЕ МАТ.ОЖИДАНИЕ с.в

3. $E_{(\xi_1;\xi_2)}(h(x_1;x_2)) = E_{\xi_2}[E_{\xi_1}(h(x_1;x_2)|\xi_2)]$

<u>Доказательство</u>:

$$\int\limits_{X_2}\int\limits_{X_1} \mathrm{h}(z_1,z_2)\,\mathrm{f}_{\xi_1,\xi_2}(z_1,z_2)\,dx_1 dx_2 =$$

$$= \int\limits_{X_2}\left(\int\limits_{X_1} \mathrm{h}(z_1,z_2)\frac{\mathrm{f}_{\xi_1,\xi_2}(z_1,z_2)}{\mathrm{f}_{\xi_2}(z_2)}dx_1\right)\mathrm{f}_{\xi_2}(z_2)\,dx_2 =$$

УСЛОВНОЕ МАТ.ОЖИДАНИЕ с.в

3. $E_{(\xi_1;\xi_2)}(h(x_1;x_2)) = E_{\xi_2}[E_{\xi_1}(h(x_1;x_2)|\xi_2)]$

Доказательство:

$$\int_{X_2}\int_{X_1} \mathrm{h}(z_1, z_2)\,\mathrm{f}_{\xi_1,\xi_2}(z_1, z_2)\,dx_1 dx_2 =$$

$$= \int_{X_2}\left(\int_{X_1} \mathrm{h}(z_1, z_2)\frac{\mathrm{f}_{\xi_1,\xi_2}(z_1, z_2)}{\mathrm{f}_{\xi_2}(z_2)}dx_1\right)\mathrm{f}_{\xi_2}(z_2)\,dx_2 =$$

$$= \int_{X_2} \mathrm{E}_{\xi_1}\left[\mathrm{h}(x_1, z_2) \mid \xi_2\right]\mathrm{f}_{\xi_2}(z_2)\,dx_2 = \mathrm{E}_{\xi_2}\;\mathrm{E}_{\xi_1}\left[\mathrm{h}(x_1, x_2) \mid \xi_2\right]$$

УСЛОВНОЕ МАТ.ОЖИДАНИЕ с.в

Пример:

$\xi_{1;}\ \xi_2$ – независимы и одинаково распределены.

Найти $M=E_{\xi_1}(\xi_1 |\ \xi_1+\xi_2=z)$

Решение:

$M=E_{\xi_1}(\xi_1|\xi_1+\xi_2=z)=E_{\xi_1}(\xi_1|\xi_1=z-\xi_2)=E_{\xi_1}(z-\xi_2)=z-\xi_2$

$M= z-\xi_1$ в силу одинаковой распределенности

$2M=2z-\xi_1-\xi_2=2z-z=z$

$M=z/2$

УСЛОВНОЕ МАТ.ОЖИДАНИЕ с.в

Упражнение:

Дискретная двумерная с.в. $\xi=(\xi_x; \xi_y)$

ξ	-2	-1	1	2
-1	0,06	0,03	0,01	0,00
0	0,05	0,25	0,35	0,05
1	0,01	0,02	0,07	0,10

Найти $E_{\xi_x}\left(\xi_x \mid \xi_x \neq 0; \xi_y = -2 \vee \xi_y = 1\right)$

УСЛОВНАЯ ДИСПЕРСИЯ с.в

Свойства условной дисперсии

1. $V_{\xi_2}(\xi_1|\xi_2)=0$:

ξ_1 по отношению к ξ_2 является константой – вариация равна нулю

2. $V_{\xi_1}(g(x_2)\cdot\xi_1|\xi_2)= g(x_2)\cdot V_{\xi_1}(\xi_1|\xi_2)\cdot g^T(x_2)$:

$g(x_2)$ по отношению к ξ_1 является константой – можно вынести за знак дисперсии с квадратом

3. $V(y)= V_{\xi_2}[E_{\xi_1}(\xi_1|\xi_2)]+ E_{\xi_2}[V_{\xi_1}(\xi_1|\xi_2)]$:

формула разложения дисперсии на межгрупповую и внутригрупповую.

НАИЛУЧШИЙ ПРЕДИКТОР

$$\text{MSPE (BP)}: \quad \min_{g(\mathbf{X})} \mathrm{E}\left(Y - g\left(\mathbf{X}\right)\right)^2$$

$$g\left(\mathbf{X}\right) = \mathrm{E}\left(Y \mid \mathbf{X}\right)$$

$$\mathrm{E}\left(Y - g\left(\mathbf{X}\right)\right)^2 = \mathrm{E}\left(Y - \mathrm{E}\left(Y \mid \mathbf{X}\right) + \mathrm{E}\left(Y \mid \mathbf{X}\right) - g\left(\mathbf{X}\right)\right)^2 =$$

НАИЛУЧШИЙ ПРЕДИКТОР

$$\text{MSPE (BP)}: \quad \min_{g(\mathbf{X})} \mathrm{E}\left(Y - g\left(\mathbf{X}\right)\right)^2$$

$$g\left(\mathbf{X}\right) = \mathrm{E}\left(Y \mid \mathbf{X}\right)$$

$$\mathrm{E}\left(Y - g\left(\mathbf{X}\right)\right)^2 = \mathrm{E}\left(Y - \mathrm{E}\left(Y \mid \mathbf{X}\right) + \mathrm{E}\left(Y \mid \mathbf{X}\right) - g\left(\mathbf{X}\right)\right)^2 =$$

$$= \mathrm{E}\left(e\left(\mathbf{X}\right) + h\left(\mathbf{X}\right)\right)^2 =$$

$$= \mathrm{E}\left(e\left(\mathbf{X}\right)\right)^2 + \mathrm{E}\left(h\left(\mathbf{X}\right)\right)^2 + 2 \cdot \mathrm{E}\left(e\left(\mathbf{X}\right) \cdot h\left(\mathbf{X}\right)\right) =$$

НАИЛУЧШИЙ ПРЕДИКТОР

$$\text{MSPE (BP)}: \quad \min_{g(\mathbf{X})} \mathrm{E}\left(Y - g\left(\mathbf{X}\right)\right)^2$$

$$g\left(\mathbf{X}\right) = \mathrm{E}\left(Y \mid \mathbf{X}\right)$$

$$\mathrm{E}\left(Y - g\left(\mathbf{X}\right)\right)^2 = \mathrm{E}\left(Y - \mathrm{E}\left(Y \mid \mathbf{X}\right) + \mathrm{E}\left(Y \mid \mathbf{X}\right) - g\left(\mathbf{X}\right)\right)^2 =$$

$$= \mathrm{E}\left(e\left(\mathbf{X}\right) + h\left(\mathbf{X}\right)\right)^2 =$$

$$= \mathrm{E}\left(e\left(\mathbf{X}\right)\right)^2 + \mathrm{E}\left(h\left(\mathbf{X}\right)\right)^2 + 2 \cdot \mathrm{E}\left(e\left(\mathbf{X}\right) \cdot h\left(\mathbf{X}\right)\right) =$$

$$= \mathrm{E}\left(e\left(\mathbf{X}\right)\right)^2 + \mathrm{E}\left(h\left(\mathbf{X}\right)\right)^2 \geq \mathrm{E}\left(e\left(\mathbf{X}\right)\right)^2 = \mathrm{E}\left(Y - \mathrm{E}\left(Y \mid \mathbf{X}\right)\right)^2$$

НАИЛУЧШИЙ ПРЕДИКТОР

Ошибка наилучшего предиктора:

$$e\left(\mathbf{X}\right) = Y - \mathrm{E}\left(Y \mid \mathbf{X}\right)$$

$$\mathrm{E}\left(e\left(\mathbf{X}\right) \mid \mathbf{X}\right) = \mathrm{E}\left(Y - \mathrm{E}\left(Y \mid \mathbf{X}\right) \mid \mathbf{X}\right) =$$

НАИЛУЧШИЙ ПРЕДИКТОР

Ошибка наилучшего предиктора:

$$e(\mathbf{X}) = Y - \mathrm{E}(Y \mid \mathbf{X})$$

$$\mathrm{E}\big(e(\mathbf{X}) \mid \mathbf{X}\big) = \mathrm{E}\big(Y - \mathrm{E}(Y \mid \mathbf{X}) \mid \mathbf{X}\big) =$$

$$= \mathrm{E}(Y \mid \mathbf{X}) - \mathrm{E}\big(\mathrm{E}(Y \mid \mathbf{X}) \mid \mathbf{X}\big) =$$

НАИЛУЧШИЙ ПРЕДИКТОР

Ошибка наилучшего предиктора:

$$e(\mathbf{X}) = Y - \mathrm{E}(Y \mid \mathbf{X})$$

$$\mathrm{E}\big(e(\mathbf{X}) \mid \mathbf{X}\big) = \mathrm{E}\big(Y - \mathrm{E}(Y \mid \mathbf{X}) \mid \mathbf{X}\big) =$$

$$= \mathrm{E}(Y \mid \mathbf{X}) - \mathrm{E}\big(\mathrm{E}(Y \mid \mathbf{X}) \mid \mathbf{X}\big) =$$

$$= \mathrm{E}(Y \mid \mathbf{X}) - \mathrm{E}(Y \mid \mathbf{X}) \cdot \mathrm{E}(1 \mid \mathbf{X}) = 0$$

НАИЛУЧШИЙ ПРЕДИКТОР

Ошибка наилучшего предиктора:

$$e(\mathbf{X}) = Y - \mathrm{E}(Y \mid \mathbf{X})$$

$$\mathrm{E}(e(\mathbf{X}) \mid \mathbf{X}) = 0$$

$$\mathrm{E}(e(\mathbf{X})) = \mathrm{E}(\mathrm{E}(e(\mathbf{X}) \mid \mathbf{X})) = 0$$

НАИЛУЧШИЙ ПРЕДИКТОР

Ошибка наилучшего предиктора:

$$e(\mathbf{X}) = Y - \mathrm{E}(Y \mid \mathbf{X})$$

$$\mathrm{E}\big(e(\mathbf{X}) \mid \mathbf{X}\big) = 0$$

$$\mathrm{E}\big(e(\mathbf{X}) \cdot h(\mathbf{X})\big) = \mathrm{E}\Big(\mathrm{E}\big(e(\mathbf{X}) \cdot h(\mathbf{X}) \mid \mathbf{X}\big)\Big) =$$

НАИЛУЧШИЙ ПРЕДИКТОР

Ошибка наилучшего предиктора:

$$e(\mathbf{X}) = Y - \mathrm{E}(Y \mid \mathbf{X})$$

$$\mathrm{E}\left(e(\mathbf{X}) \mid \mathbf{X}\right) = 0$$

$$\mathrm{E}\left(e(\mathbf{X}) \cdot h(\mathbf{X})\right) = \mathrm{E}\left(\mathrm{E}\left(e(\mathbf{X}) \cdot h(\mathbf{X}) \mid \mathbf{X}\right)\right) =$$

$$= \mathrm{E}\left(h(\mathbf{X}) \cdot \mathrm{E}\left(e(\mathbf{X}) \mid \mathbf{X}\right)\right) = \mathrm{E}\left(h(\mathbf{X}) \cdot 0\right) = 0$$

НАИЛУЧШИЙ ЛИНЕЙНЫЙ ПРЕДИКТОР

$$\mathrm{MSPE}\ \ (\mathrm{BLP}): \quad \min_{\alpha,\beta} \mathrm{E}\left(Y - \alpha - \mathbf{X}\beta\right)^2$$

FOC (УСЛОВИЯ ПЕРВОГО ПОРЯДКА):

$$-2\,\mathrm{E}\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0; \quad -2\,\mathrm{E}\,\mathbf{X}'\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0$$

НАИЛУЧШИЙ ЛИНЕЙНЫЙ ПРЕДИКТОР

$$\text{MSPE (BLP)}: \quad \min_{\alpha,\beta} \mathrm{E}\left(Y - \alpha - \mathbf{X}\beta\right)^2$$

FOC (УСЛОВИЯ ПЕРВОГО ПОРЯДКА):

$$-2\,\mathrm{E}\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0; \quad -2\,\mathrm{E}\,\mathbf{X}'\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0$$

SOC (УСЛОВИЯ ВТОРОГО ПОРЯДКА):

$$\mathbf{V} = \begin{pmatrix} -2(-1) & 2\,\mathrm{E}\,\mathbf{X} \\ 2\,\mathrm{E}\,\mathbf{X}' & 2\,\mathrm{E}\left(\mathbf{X}'\mathbf{X}\right) \end{pmatrix} = \begin{pmatrix} 2 & 2\,\mathrm{E}\,\mathbf{X} \\ 2\,\mathrm{E}\,\mathbf{X}' & 2\,\mathrm{E}\left(\mathbf{X}'\mathbf{X}\right) \end{pmatrix} \geq 0$$

$$\left.\begin{array}{l} \left(\mathbf{X}'\mathbf{X}\right)' = \mathbf{X}'\mathbf{X} \text{ симметрична} \\ \det \mathrm{V} = 2 \cdot \left(\mathrm{E}\left(\mathbf{X}'\mathbf{X}\right) - \mathrm{E}\,\mathbf{X} \cdot \mathrm{E}\,\mathbf{X}'\right) = 2 \cdot \det\left(\operatorname{cov} \mathbf{X}\right) \geq 0 \end{array}\right\} \Rightarrow \mathbf{V} \text{ пол. определена}$$

НАИЛУЧШИЙ ЛИНЕЙНЫЙ ПРЕДИКТОР

$$\text{MSPE (BLP)}: \quad \min_{\alpha,\beta} \mathrm{E}\left(Y - \alpha - \mathbf{X}\beta\right)^2$$

FOC (УСЛОВИЯ ПЕРВОГО ПОРЯДКА):

$$-2\,\mathrm{E}\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0; \quad -2\,\mathrm{E}\,\mathbf{X}'\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0$$

РЕШЕНИЕ:

$$\alpha = \mathrm{E}\,Y - \mathrm{E}\,\mathbf{X}\cdot\beta;$$

НАИЛУЧШИЙ ЛИНЕЙНЫЙ ПРЕДИКТОР

$$\text{MSPE (BLP)}: \quad \min_{\alpha,\beta} \mathrm{E}\left(Y - \alpha - \mathbf{X}\beta\right)^2$$

FOC (УСЛОВИЯ ПЕРВОГО ПОРЯДКА):

$$-2\,\mathrm{E}\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0; \quad -2\,\mathrm{E}\,\mathbf{X}'\left(Y - \alpha - \mathbf{X}\boldsymbol{\beta}\right) = 0$$

РЕШЕНИЕ:

$$\alpha = \mathrm{E}\,Y - \mathrm{E}\,\mathbf{X} \cdot \beta$$

$$\mathrm{E}\left(\mathbf{X}'Y\right) - \mathrm{E}\,\mathbf{X}'\,\mathrm{E}\,Y + \mathrm{E}\,\mathbf{X}'\,\mathrm{E}\left(\mathbf{X}\beta\right) - \mathrm{E}\left(\mathbf{X}'\mathbf{X}\beta\right) = 0 \Leftrightarrow$$

$$\Leftrightarrow \operatorname{cov}\left(\mathbf{X}, Y\right) = \operatorname{Var}(\mathbf{X}) \cdot \beta$$

НАИЛУЧШИЙ ЛИНЕЙНЫЙ ПРЕДИКТОР

$$\text{MSPE (BLP)}: \quad \min_{\alpha,\beta} \mathrm{E}\left(Y - \alpha - \mathbf{X}\beta\right)^2$$

РЕШЕНИЕ:

$\alpha = \mathrm{E}\,Y - \mathrm{E}\,\mathbf{X} \cdot \beta;$

$\beta = \mathrm{Var}^{-1}(\mathbf{X}) \cdot \mathrm{cov}\left(\mathbf{X}, Y\right)$

Выборочные оценки:

$$\widehat{\beta} = \widehat{\mathrm{Var}}^{-1}(\mathbf{X}) \cdot \widehat{\mathrm{cov}}\left(\mathbf{X}, Y\right); \quad \widehat{\alpha} = \overline{Y} - \overline{\mathbf{X}} \cdot \widehat{\mathrm{Var}}^{-1}(\mathbf{X}) \cdot \widehat{\mathrm{cov}}\left(\mathbf{X}, Y\right)$$

НАИЛУЧШАЯ ЛИНЕЙНАЯ АПРОКСИМАЦИЯ УСЛОВНОГО СРЕДНЕГО

$$\text{MSAE (BLP)}: \quad \min_{\alpha,\beta} \mathrm{E}\left(\mathrm{E}\left(Y \mid \mathbf{X}\right) - \alpha - \mathbf{X}\beta\right)^2$$

$$\alpha = \mathrm{E}\,Y - \mathrm{E}\,\mathbf{X} \cdot \beta$$

$$\mathrm{cov}\left(\mathbf{X}, Y\right) = \mathrm{Var}(\mathbf{X}) \cdot \beta$$

Упражнение:
докажите аналогично предыдущему

КОНЕЦ ЛЕКЦИИ