# Теоретические основы информатики (концептуальные модели и математические основы)

## Лекция № 2. Неравенства концентрации меры

А.С. Шундеев

# Содержание

# Содержание

# Введение

Феномен концентрации меры состоит в следующем.

# Введение

Феномен концентрации меры состоит в следующем.

Если рассмотреть измеримую функцию $f(x_1, \ldots, x_n)$, не очень «чувствительную» к небольшим изменениям значений своих аргументов, и набор независимых случайных элементов $\xi_1, \ldots, \xi_n$, то оказывается, что случайная величина $f(\xi_1, \ldots, \xi_n)$ будет не сильно отклоняться от своего математического ожидания $\mathbf{E}\big[f(\xi_1, \ldots, \xi_n)\big]$. Величина подобных отклонений формально оценивается с помощью так называемых неравенств концентрации меры.

# Введение

Феномен концентрации меры состоит в следующем.

Если рассмотреть измеримую функцию $f(x_1, \ldots, x_n)$, не очень «чувствительную» к небольшим изменениям значений своих аргументов, и набор независимых случайных элементов $\xi_1, \ldots, \xi_n$, то оказывается, что случайная величина $f(\xi_1, \ldots, \xi_n)$ будет не сильно отклоняться от своего математического ожидания $\mathbf{E}\big[f(\xi_1, \ldots, \xi_n)\big]$. Величина подобных отклонений формально оценивается с помощью так называемых неравенств концентрации меры.

Нами будут рассмотрены неравенства Хёффдинга, Азумы-Хёффдинга и МакДиармида, играющие значительную роль при построении и исследовании свойств формальных моделей обучения.

# Введение

Феномен концентрации меры состоит в следующем.

Если рассмотреть измеримую функцию $f(x_1, \dots, x_n)$, не очень «чувствительную» к небольшим изменениям значений своих аргументов, и набор независимых случайных элементов $\xi_1, \dots, \xi_n$, то оказывается, что случайная величина $f(\xi_1, \dots, \xi_n)$ будет не сильно отклоняться от своего математического ожидания $\mathbf{E}\big[f(\xi_1, \dots, \xi_n)\big]$. Величина подобных отклонений формально оценивается с помощью так называемых неравенств концентрации меры.

Нами будут рассмотрены неравенства Хёффдинга, Азумы-Хёффдинга и МакДиармида, играющие значительную роль при построении и исследовании свойств формальных моделей обучения.

Далее, будем считать, что задано некоторое вероятностное пространство $(\Omega, \mathcal{G}, \mathsf{P})$.

# Содержание

# Метод Чернова получения вероятностных оценок

**Утверждение 2.1 (Неравенство Маркова).**

Пусть $\xi$ – неотрицательная случайная величина и $\varepsilon > 0$. Тогда

$$\mathsf{P}\{\xi \geqslant \varepsilon\} \leqslant \frac{\mathbf{E}\xi}{\varepsilon}.$$

# Метод Чернова получения вероятностных оценок

**Утверждение 2.1 (Неравенство Маркова).**

Пусть $\xi$ – неотрицательная случайная величина и $\varepsilon > 0$. Тогда

$$\mathsf{P}\{\xi \geqslant \varepsilon\} \leqslant \frac{\mathbf{E}\xi}{\varepsilon}.$$

◀ Действительно,

$$\mathsf{P}\{\xi \geqslant \varepsilon\} = \mathbf{E}\,\mathbf{1}_{\{\xi \geqslant \varepsilon\}} \leqslant \frac{1}{\varepsilon}\,\mathbf{E}\left[\xi\,\mathbf{1}_{\{\xi \geqslant \varepsilon\}}\right] \leqslant \frac{\mathbf{E}\xi}{\varepsilon}.$$

■

# Метод Чернова получения вероятностных оценок

## Следствие 2.1.

Предположим, что случайная величина $\xi$ принимает свои значения из промежутка $[0,1]$ и число $\varepsilon \in (0,1)$. Тогда

$$\mathsf{P}\{\xi > 1-\varepsilon\} \geqslant \frac{\mathbf{E}\xi + \varepsilon - 1}{\varepsilon}.$$

◀ Запишем для случайной величины $\eta := 1-\xi$ неравенство Маркова

$$\mathsf{P}\{\xi \leqslant 1-\varepsilon\} = \mathsf{P}\{\eta \geqslant \varepsilon\} \leqslant \frac{\mathbf{E}\eta}{\varepsilon} = \frac{1-\mathbf{E}\xi}{\varepsilon},$$

# Метод Чернова получения вероятностных оценок

## Следствие 2.1.

Предположим, что случайная величина $\xi$ принимает свои значения из промежутка $[0,1]$ и число $\varepsilon \in (0,1)$. Тогда

$$\mathsf{P}\{\xi > 1-\varepsilon\} \geqslant \frac{\mathbf{E}\xi + \varepsilon - 1}{\varepsilon}.$$

◀ Запишем для случайной величины $\eta := 1-\xi$ неравенство Маркова

$$\mathsf{P}\{\xi \leqslant 1-\varepsilon\} = \mathsf{P}\{\eta \geqslant \varepsilon\} \leqslant \frac{\mathbf{E}\eta}{\varepsilon} = \frac{1-\mathbf{E}\xi}{\varepsilon},$$

но тогда

$$\mathsf{P}\{\xi > 1-\varepsilon\} \geqslant 1 - \frac{1-\mathbf{E}\xi}{\varepsilon} = \frac{\mathbf{E}\xi + \varepsilon - 1}{\varepsilon}.$$

■

# Метод Чернова получения вероятностных оценок

**Утверждение 2.2 (Неравенство Чебышёва).**

Пусть $\xi$ – случайная величина, $\mathbf{D}\xi < \infty$ и $\varepsilon > 0$. Тогда

$$\mathsf{P}\{|\xi - \mathbf{E}\xi| \geqslant \varepsilon\} \leqslant \frac{\mathbf{D}\xi}{\varepsilon^2}.$$

# Метод Чернова получения вероятностных оценок

**Утверждение 2.2 (Неравенство Чебышёва).**

Пусть $\xi$ – случайная величина, $\mathbf{D}\xi < \infty$ и $\varepsilon > 0$. Тогда

$$\mathsf{P}\{|\xi - \mathbf{E}\xi| \geqslant \varepsilon\} \leqslant \frac{\mathbf{D}\xi}{\varepsilon^2}.$$

◀ Используя неравенство Маркова, запишем

$$\mathsf{P}\{|\xi - \mathbf{E}\xi| \geqslant \varepsilon\} = \mathsf{P}\{(\xi - \mathbf{E}\xi)^2 \geqslant \varepsilon^2\} \leqslant \frac{\mathbf{E}(\xi - \mathbf{E}\xi)^2}{\varepsilon^2} = \frac{\mathbf{D}\xi}{\varepsilon^2}.$$

■

# Метод Чернова получения вероятностных оценок

Опишем основную идею метод Чернова получения вероятностных оценок (Chernoff bounding trick), который в дальнейшем будет использоваться для доказательства неравенств концентрации меры.

# Метод Чернова получения вероятностных оценок

Опишем основную идею метод Чернова получения вероятностных оценок (Chernoff bounding trick), который в дальнейшем будет использоваться для доказательства неравенств концентрации меры.

Используя монотонность экспоненты и неравенство Маркова, запишем

$$\mathsf{P}\{\xi - \mathbf{E}\xi \geqslant \varepsilon\} = \mathsf{P}\{e^{s(\xi - \mathbf{E}\xi)} \geqslant e^{s\varepsilon}\} \leqslant e^{-s\varepsilon}\,\mathbf{E}\left[e^{s(\xi - \mathbf{E}\xi)}\right].$$

# Метод Чернова получения вероятностных оценок

Опишем основную идею метод Чернова получения вероятностных оценок (Chernoff bounding trick), который в дальнейшем будет использоваться для доказательства неравенств концентрации меры.

Используя монотонность экспоненты и неравенство Маркова, запишем

$$\mathsf{P}\{\xi - \mathbf{E}\xi \geqslant \varepsilon\} = \mathsf{P}\{e^{s(\xi - \mathbf{E}\xi)} \geqslant e^{s\varepsilon}\} \leqslant e^{-s\varepsilon}\,\mathbf{E}\left[e^{s(\xi - \mathbf{E}\xi)}\right].$$

Но тогда

$$\mathsf{P}\{\xi - \mathbf{E}\xi \geqslant \varepsilon\} \leqslant \inf_{s>0} e^{-s\varepsilon}\,\mathbf{E}\left[e^{s(\xi - \mathbf{E}\xi)}\right].$$

# Метод Чернова получения вероятностных оценок

В дальнейшем, мы не будем стремиться точно вычислить инфимум, стоящий в правой части этого неравенства. Достаточно будет подобрать подходящее значение $s$, при котором выражение $\mathbf{E}\left[e^{s(\xi-\mathbf{E}\xi)}\right]$ будет служить приемлемой оценкой для рассматриваемой вероятности.

# Содержание

# Неравенство Хёффдинга

**Лемма 2.1 (Хёффдинг).**

Пусть $\xi$ – случайная величина и $\xi \in [a, b]$ (п.н.) для некоторых $a, b \in \mathbb{R}$, $a < b$. Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] \leqslant e^{\frac{\varepsilon^2 (b-a)^2}{8}}. \tag{1}$$

# Неравенство Хёффдинга

**Лемма 2.1 (Хёффдинг).**

Пусть $\xi$ – случайная величина и $\xi \in [a, b]$ (п.н.) для некоторых $a, b \in \mathbb{R}$, $a < b$. Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] \leqslant e^{\frac{\varepsilon^2 (b-a)^2}{8}}. \tag{1}$$

◀ Зафиксируем произвольное $\varepsilon > 0$ и рассмотрим два случая.

# Неравенство Хёффдинга

**Лемма 2.1 (Хёффдинг).**

Пусть $\xi$ – случайная величина и $\xi \in [a, b]$ (п.н.) для некоторых $a, b \in \mathbb{R}$, $a < b$. Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] \leqslant e^{\frac{\varepsilon^2 (b-a)^2}{8}}. \tag{1}$$

◀ Зафиксируем произвольное $\varepsilon > 0$ и рассмотрим два случая.

1) Предположим, что $\mathbf{E}\xi = 0$. Функция $x \mapsto e^{\varepsilon x}$ является выпуклой, а значит для любого $x \in [a, b]$ выполняется неравенство

$$e^{\varepsilon x} \leqslant \frac{b - x}{b - a}\, e^{\varepsilon a} + \frac{x - a}{b - a}\, e^{\varepsilon b}.$$

# Неравенство Хёффдинга

Следовательно,

$$\mathbf{E}\left[e^{\varepsilon\xi}\right] \leqslant \frac{b-\mathbf{E}\xi}{b-a}\,e^{\varepsilon a} + \frac{\mathbf{E}\xi - a}{b-a}\,e^{\varepsilon b} = \frac{b}{b-a}\,e^{\varepsilon a} - \frac{a}{b-a}\,e^{\varepsilon b} \tag{2}$$
$$= e^{\varphi(\varepsilon(b-a))},$$

# Неравенство Хёффдинга

Следовательно,

$$\mathbf{E}\left[e^{\varepsilon\xi}\right] \leqslant \frac{b-\mathbf{E}\xi}{b-a}\,e^{\varepsilon a} + \frac{\mathbf{E}\xi - a}{b-a}\,e^{\varepsilon b} = \frac{b}{b-a}\,e^{\varepsilon a} - \frac{a}{b-a}\,e^{\varepsilon b} \\ = e^{\varphi(\varepsilon(b-a))}, \tag{2}$$

где

$$\varphi(\lambda) := -p\lambda + \ln\left(pe^{\lambda} + (1-p)\right), \quad p := \frac{-a}{b-a}$$

(заметим, что так как $\mathbf{E}\xi = 0$, то $a \leqslant 0$, а значит $p \geqslant 0$).

# Неравенство Хёффдинга

Следовательно,

$$\mathbf{E}\left[e^{\varepsilon\xi}\right] \leqslant \frac{b-\mathbf{E}\xi}{b-a}\, e^{\varepsilon a} + \frac{\mathbf{E}\xi - a}{b-a}\, e^{\varepsilon b} = \frac{b}{b-a}\, e^{\varepsilon a} - \frac{a}{b-a}\, e^{\varepsilon b} \\ = e^{\varphi(\varepsilon(b-a))}, \tag{2}$$

где

$$\varphi(\lambda) := -p\lambda + \ln\left(pe^{\lambda} + (1-p)\right), \quad p := \frac{-a}{b-a}$$

(заметим, что так как $\mathbf{E}\xi = 0$, то $a \leqslant 0$, а значит $p \geqslant 0$).

Запишем для функции $\varphi$ разложение Тейлора

$$\varphi(\lambda) = \varphi(0) + \varphi'(0)\lambda + \frac{1}{2}\varphi''(u)\lambda^2, \quad u = u(\lambda) \in (0, \lambda). \tag{3}$$

# Неравенство Хёффдинга

Вычислим первую и вторую производные функции $\varphi$. Получим

$$\begin{aligned} \varphi'(\lambda) &= -p + \tfrac{pe^{\lambda}}{pe^{\lambda}+(1-p)}, \\ \varphi''(\lambda) &= \tfrac{p(1-p)e^{\lambda}}{(pe^{\lambda}+(1-p))^2}. \end{aligned}$$

# Неравенство Хёффдинга

Вычислим первую и вторую производные функции $\varphi$. Получим

$$\begin{aligned} \varphi'(\lambda) &= -p + \frac{pe^{\lambda}}{pe^{\lambda}+(1-p)}, \\ \varphi''(\lambda) &= \frac{p(1-p)e^{\lambda}}{(pe^{\lambda}+(1-p))^2}. \end{aligned}$$

Так как $\varphi(0) = \varphi'(0) = 0$, то разложение (3) примет вид

$$\varphi(\lambda) = \frac{1}{2}\varphi''(u)\lambda^2, \quad u = u(\lambda) \in (0, \lambda). \tag{4}$$

# Неравенство Хёффдинга

Вычислим первую и вторую производные функции $\varphi$. Получим

$$\begin{aligned}\varphi'(\lambda) &= -p + \tfrac{pe^\lambda}{pe^\lambda+(1-p)},\\ \varphi''(\lambda) &= \tfrac{p(1-p)e^\lambda}{(pe^\lambda+(1-p))^2}.\end{aligned}$$

Так как $\varphi(0) = \varphi'(0) = 0$, то разложение (3) примет вид

$$\varphi(\lambda) = \frac{1}{2}\varphi''(u)\lambda^2, \quad u = u(\lambda) \in (0, \lambda). \tag{4}$$

Учитывая числовое неравенство $c\,d \leqslant \frac{1}{4}(c+d)^2$ верное для любых $c, d \in \mathbb{R}$, получим оценку

$$\varphi''(u) = \frac{pe^u}{pe^u + (1-p)} \cdot \frac{(1-p)}{pe^u + (1-p)} \leqslant \frac{1}{4}\left[\frac{pe^u + (1-p)}{pe^u + (1-p)}\right]^2 = \frac{1}{4}. \tag{5}$$

# Неравенство Хёффдинга

Объединяя (2), (4) и (5), получим требуемое неравенство (1).

# Неравенство Хёффдинга

Объединяя (2), (4) и (5), получим требуемое неравенство (1).

2) Если $\mathbf{E}\xi \neq 0$, то введем новую случайную величину $\zeta := \xi - \mathbf{E}\xi$. Так как $\mathbf{E}\zeta = 0$, то для $\zeta$ выполняется утверждение леммы.

# Неравенство Хёффдинга

Объединяя (2), (4) и (5), получим требуемое неравенство (1).

2) Если $\mathbf{E}\xi \neq 0$, то введем новую случайную величину $\zeta := \xi - \mathbf{E}\xi$. Так как $\mathbf{E}\zeta = 0$, то для $\zeta$ выполняется утверждение леммы.

Следовательно,

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] = \mathbf{E}\left[e^{\varepsilon\zeta}\right] \leqslant e^{\frac{\varepsilon^2((b-\mathbf{E}\xi)-(a-\mathbf{E}\xi))^2}{8}} = e^{\frac{\varepsilon^2(b-a)^2}{8}}.$$

∎

# Неравенство Хёффдинга

**Теорема 2.1 (неравенство Хёффдинга).**

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ – независимые случайные величины и $\xi_i \in [a_i, b_i]$ (п.н.) для некоторых $a_i, b_i \in \mathbb{R}$, $a_i < b_i$ ($i = 1, 2, \ldots, n$). Обозначим $S_n := \sum_{i=1}^{n} \xi_i$.

# Неравенство Хёффдинга

## Теорема 2.1 (неравенство Хёффдинга).

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ – независимые случайные величины и $\xi_i \in [a_i, b_i]$ (п.н.) для некоторых $a_i, b_i \in \mathbb{R}$, $a_i < b_i$ ($i = 1, 2, \ldots, n$). Обозначим $S_n := \sum_{i=1}^n \xi_i$.
Тогда для любого $\varepsilon > 0$ выполняются неравенства

$$\mathsf{P}\big\{S_n - \mathbf{E}S_n \geqslant \varepsilon\big\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right), \tag{6}$$

$$\mathsf{P}\big\{S_n - \mathbf{E}S_n \leqslant -\varepsilon\big\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right), \tag{7}$$

# Неравенство Хёффдинга

## Теорема 2.1 (неравенство Хёффдинга).

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ – независимые случайные величины и $\xi_i \in [a_i, b_i]$ (п.н.) для некоторых $a_i, b_i \in \mathbb{R}$, $a_i < b_i$ ($i = 1, 2, \ldots, n$). Обозначим $S_n := \sum_{i=1}^n \xi_i$.
Тогда для любого $\varepsilon > 0$ выполняются неравенства

$$\mathrm{P}\{S_n - \mathbf{E}S_n \geqslant \varepsilon\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right), \tag{6}$$

$$\mathrm{P}\{S_n - \mathbf{E}S_n \leqslant -\varepsilon\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right), \tag{7}$$

а значит и

$$\mathrm{P}\{|S_n - \mathbf{E}S_n| \geqslant \varepsilon\} \leqslant 2\exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \tag{8}$$

# Неравенство Хёффдинга

◀ Прежде всего заметим, что неравенство (7) является вариантом неравенства (6), которое записано для случайных величин $-\xi_1, -\xi_2, \ldots, -\xi_n$.

# Неравенство Хёффдинга

◀ Прежде всего заметим, что неравенство (7) является вариантом неравенства (6), которое записано для случайных величин $-\xi_1, -\xi_2, \ldots, -\xi_n$.

Неравенство (8) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

# Неравенство Хёффдинга

◀ Прежде всего заметим, что неравенство (7) является вариантом неравенства (6), которое записано для случайных величин $-\xi_1, -\xi_2, \ldots, -\xi_n$.

Неравенство (8) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

Таким образом, достаточно доказать неравенство (6).

# Неравенство Хёффдинга

◀ Прежде всего заметим, что неравенство (7) является вариантом неравенства (6), которое записано для случайных величин $-\xi_1, -\xi_2, \ldots, -\xi_n$.

Неравенство (8) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

Таким образом, достаточно доказать неравенство (6).

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{4\varepsilon}{\sum_{i=1}^{n}(b_i - a_i)^2}. \tag{9}$$

# Неравенство Хёффдинга

◀ Прежде всего заметим, что неравенство (7) является вариантом неравенства (6), которое записано для случайных величин $-\xi_1, -\xi_2, \dots, -\xi_n$.

Неравенство (8) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

Таким образом, достаточно доказать неравенство (6).

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{4\varepsilon}{\sum_{i=1}^{n}(b_i - a_i)^2}. \tag{9}$$

Используя неравенство Маркова, получим оценку

$$\mathsf{P}\{S_n - \mathbf{E}S_n \geqslant \varepsilon\} = \mathsf{P}\{e^{s(S_n - \mathbf{E}S_n)} \geqslant e^{s\varepsilon}\} \leqslant e^{-s\varepsilon}\,\mathbf{E}\left[e^{s(S_n - \mathbf{E}S_n)}\right]. \tag{10}$$

# Неравенство Хёффдинга

Учитывая независимость рассматриваемых случайных величин, запишем

$$\mathbf{E}\left[e^{s(S_n - \mathbf{E}S_n)}\right] = \mathbf{E}\left[\prod_{i=1}^{n} e^{s(\xi_i - \mathbf{E}\xi_i)}\right] = \prod_{i=1}^{n} \mathbf{E}\left[e^{s(\xi_i - \mathbf{E}\xi_i)}\right]. \tag{11}$$

# Неравенство Хёффдинга

Учитывая независимость рассматриваемых случайных величин, запишем

$$\mathbf{E}\left[e^{s(S_n-\mathbf{E}S_n)}\right] = \mathbf{E}\left[\prod_{i=1}^{n} e^{s(\xi_i-\mathbf{E}\xi_i)}\right] = \prod_{i=1}^{n} \mathbf{E}\left[e^{s(\xi_i-\mathbf{E}\xi_i)}\right]. \qquad (11)$$

Из леммы Хёффдинга следует, что

$$\mathbf{E}\left[e^{s(\xi_i-\mathbf{E}\xi_i)}\right] \leqslant e^{\frac{s^2(b_i-a_i)^2}{8}} \quad (i = 1, 2, \ldots, n). \qquad (12)$$

# Неравенство Хёффдинга

Учитывая независимость рассматриваемых случайных величин, запишем

$$\mathbf{E}\left[e^{s(S_n-\mathbf{E}S_n)}\right] = \mathbf{E}\left[\prod_{i=1}^{n} e^{s(\xi_i-\mathbf{E}\xi_i)}\right] = \prod_{i=1}^{n}\mathbf{E}\left[e^{s(\xi_i-\mathbf{E}\xi_i)}\right]. \tag{11}$$

Из леммы Хёффдинга следует, что

$$\mathbf{E}\left[e^{s(\xi_i-\mathbf{E}\xi_i)}\right] \leqslant e^{\frac{s^2(b_i-a_i)^2}{8}} \quad (i = 1, 2, \ldots, n). \tag{12}$$

Объединяя (9), (10), (11) и (12), получим требуемое неравенство (6)

$$\mathrm{P}\{S_n - \mathbf{E}S_n \geqslant \varepsilon\} \leqslant e^{-s\varepsilon}\prod_{i=1}^{n} e^{\frac{s^2(b_i-a_i)^2}{8}} = \exp\left(-s\varepsilon + \frac{s^2}{8}\sum_{i=1}^{n}(b_i-a_i)^2\right)$$
$$= \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^{n}(b_i-a_i)^2}\right).$$

$\blacksquare$

# Неравенство Хёффдинга

## Следствие 2.2.

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ – независимые одинаково распределенные случайные величины и $\xi_1 \in [a, b]$ (п.н.) для некоторых $a, b \in \mathbb{R}$, $a < b$. Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathrm{P}\left\{\left|\frac{S_n}{n} - \mathbf{E}\xi_1\right| \geqslant \varepsilon\right\} \leqslant 2\exp\left(-\frac{2n\varepsilon^2}{(b-a)^2}\right). \tag{13}$$

# Содержание

# Субгауссовские случайные величины

## Определение 2.1.

Случайная величина $\xi$ называется субгауссовской с параметром масштабирования $s > 0$, если для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] \leqslant e^{\frac{\varepsilon^2 s^2}{2}}.$$

# Субгауссовские случайные величины

### Определение 2.1.

Случайная величина $\xi$ называется субгауссовской с параметром масштабирования $s > 0$, если для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] \leqslant e^{\frac{\varepsilon^2 s^2}{2}}.$$

Непосредственно из определения вытекает следующее свойство.

# Субгауссовские случайные величины

## Определение 2.1.

Случайная величина $\xi$ называется субгауссовской с параметром масштабирования $s > 0$, если для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon(\xi - \mathbf{E}\xi)}\right] \leqslant e^{\frac{\varepsilon^2 s^2}{2}}.$$

Непосредственно из определения вытекает следующее свойство.

## Утверждение 2.3.

Пусть $\xi_1, \ldots, \xi_n$ – независимые субгауссовские случайные величины с параметрами масштабирования $s_1, \ldots, s_n$ соответственно.
Тогда их сумма $\xi_1 + \ldots + \xi_n$ является субгауссовской случайной величиной с параметром масштабирования $s$ таким, что $s^2 = s_1^2 + \ldots + s_n^2$.

# Субгауссовские случайные величины

## Лемма 2.2 (максимальное неравенство).

Пусть $\xi_1, \ldots, \xi_n$ – субгауссовские случайные величины с параметром масштабирования $s > 0$ и нулевым математическим ожиданием (от случайных величин не требуется независимости).
Тогда

$$\mathbf{E}\left[\max_{1 \leqslant i \leqslant n} \xi_i\right] \leqslant s\sqrt{2 \ln n}. \tag{14}$$

# Субгауссовские случайные величины

## Лемма 2.2 (максимальное неравенство).

Пусть $\xi_1, \ldots, \xi_n$ – субгауссовские случайные величины с параметром масштабирования $s > 0$ и нулевым математическим ожиданием (от случайных величин не требуется независимости).
Тогда

$$\mathbf{E}\left[\max_{1 \leqslant i \leqslant n} \xi_i\right] \leqslant s\sqrt{2 \ln n}. \tag{14}$$

◀ При $n = 1$ неравенство (14) выполняется. В этом случае его правая и левая части обращаются в нуль. Поэтому, далее, будем считать $n > 1$.

# Субгауссовские случайные величины

**Лемма 2.2 (максимальное неравенство).**

Пусть $\xi_1, \ldots, \xi_n$ – субгауссовские случайные величины с параметром масштабирования $s > 0$ и нулевым математическим ожиданием (от случайных величин не требуется независимости).
Тогда

$$\mathbf{E}\left[\max_{1 \leqslant i \leqslant n} \xi_i\right] \leqslant s\sqrt{2 \ln n}. \tag{14}$$

◀ При $n = 1$ неравенство (14) выполняется. В этом случае его правая и левая части обращаются в нуль. Поэтому, далее, будем считать $n > 1$.

Воспользуемся неравенством Йенсена. Для любого $\varepsilon > 0$ функция $x \longmapsto e^{\varepsilon x}$ $(x \in \mathbb{R})$ является выпуклой. Поэтому

# Субгауссовские случайные величины

$$e^{\varepsilon \mathbf{E}\left[\max\limits_{1\leqslant i\leqslant n} \xi_i\right]} \leqslant \mathbf{E}\left[e^{\varepsilon \max\limits_{1\leqslant i\leqslant n} \xi_i}\right] = \mathbf{E}\left[\max_{1\leqslant i\leqslant n} e^{\varepsilon \xi_i}\right]$$

$$\leqslant \mathbf{E}\left[\sum_{i=1}^{n} e^{\varepsilon \xi_i}\right] = \sum_{i=1}^{n} \mathbf{E}\left[e^{\varepsilon \xi_i}\right] = n\, e^{\frac{\varepsilon^2 s^2}{2}}\,.$$

# Субгауссовские случайные величины

$$e^{\varepsilon \mathbf{E}\left[\max\limits_{1\leqslant i\leqslant n} \xi_i\right]} \leqslant \mathbf{E}\left[e^{\varepsilon \max\limits_{1\leqslant i\leqslant n} \xi_i}\right] = \mathbf{E}\left[\max_{1\leqslant i\leqslant n} e^{\varepsilon \xi_i}\right]$$

$$\leqslant \mathbf{E}\left[\sum_{i=1}^{n} e^{\varepsilon \xi_i}\right] = \sum_{i=1}^{n} \mathbf{E}\left[e^{\varepsilon \xi_i}\right] = n\, e^{\frac{\varepsilon^2 s^2}{2}}.$$

Логарифмируя левую и правую части этого неравенства, получим

$$\mathbf{E}\left[\max_{1\leqslant i\leqslant n} \xi_i\right] \leqslant \frac{\ln n}{\varepsilon} + \frac{\varepsilon s^2}{2}.$$

# Субгауссовские случайные величины

$$
\begin{aligned}
e^{\varepsilon \mathbf{E}\left[\max\limits_{1\leqslant i\leqslant n} \xi_i\right]} &\leqslant \mathbf{E}\left[e^{\varepsilon \max\limits_{1\leqslant i\leqslant n} \xi_i}\right] = \mathbf{E}\left[\max_{1\leqslant i\leqslant n} e^{\varepsilon\xi_i}\right] \\
&\leqslant \mathbf{E}\left[\sum_{i=1}^{n} e^{\varepsilon\xi_i}\right] = \sum_{i=1}^{n} \mathbf{E}\left[e^{\varepsilon\xi_i}\right] = n\, e^{\frac{\varepsilon^2 s^2}{2}}.
\end{aligned}
$$

Логарифмируя левую и правую части этого неравенства, получим

$$
\mathbf{E}\left[\max_{1\leqslant i\leqslant n} \xi_i\right] \leqslant \frac{\ln n}{\varepsilon} + \frac{\varepsilon s^2}{2}.
$$

Выбрав

$$
\varepsilon := \frac{\sqrt{2\ln n}}{s},
$$

получим требуемое неравенство (14).

# Содержание

# Неравенство Азумы-Хёффдинга

## Лемма 2.3.

Пусть $\xi, \alpha$ – случайные величины, $\mathcal{E} \subseteq \mathcal{G}$ – $\sigma$-алгебра, $c > 0$ такие, что

- $\alpha$ – $\mathcal{E}$-измерима;
- $\alpha \leqslant \xi \leqslant \alpha + c$ (п.н.);
- $\mathbf{E}[\xi \mid \mathcal{E}] = 0$ (п.н.).

# Неравенство Азумы-Хёффдинга

## Лемма 2.3.

Пусть $\xi, \alpha$ – случайные величины, $\mathcal{E} \subseteq \mathcal{G}$ – $\sigma$-алгебра, $c > 0$ такие, что

- $\alpha$ – $\mathcal{E}$-измерима;
- $\alpha \leqslant \xi \leqslant \alpha + c$ (п.н.);
- $\mathbf{E}[\xi \mid \mathcal{E}] = 0$ (п.н.).

Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon\xi} \mid \mathcal{E}\right] \leqslant e^{\frac{\varepsilon^2 c^2}{8}} \quad \text{(п.н.)}. \tag{15}$$

# Неравенство Азумы-Хёффдинга

## Лемма 2.3.

Пусть $\xi, \alpha$ – случайные величины, $\mathcal{E} \subseteq \mathcal{G}$ – $\sigma$-алгебра, $c > 0$ такие, что

- $\alpha$ – $\mathcal{E}$-измерима;
- $\alpha \leqslant \xi \leqslant \alpha + c$ (п.н.);
- $\mathbf{E}[\xi \mid \mathcal{E}] = 0$ (п.н.).

Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon\xi} \mid \mathcal{E}\right] \leqslant e^{\frac{\varepsilon^2 c^2}{8}} \quad \text{(п.н.)}. \tag{15}$$

◀ Определим случайную величину $\beta := \alpha + c$ и зафиксируем произвольное $\varepsilon > 0$.

# Неравенство Азумы-Хёффдинга

## Лемма 2.3.

Пусть $\xi, \alpha$ – случайные величины, $\mathcal{E} \subseteq \mathcal{G}$ – $\sigma$-алгебра, $c > 0$ такие, что

- $\alpha$ – $\mathcal{E}$-измерима;
- $\alpha \leqslant \xi \leqslant \alpha + c$ (п.н.);
- $\mathbf{E}[\xi \mid \mathcal{E}] = 0$ (п.н.).

Тогда для любого $\varepsilon > 0$ выполняется неравенство

$$\mathbf{E}\left[e^{\varepsilon\xi} \mid \mathcal{E}\right] \leqslant e^{\frac{\varepsilon^2 c^2}{8}} \quad \text{(п.н.).} \tag{15}$$

◀ Определим случайную величину $\beta := \alpha + c$ и зафиксируем произвольное $\varepsilon > 0$.

Из выпуклости функции $x \mapsto e^{\varepsilon x}$ следует неравенство

$$e^{\varepsilon\xi} \leqslant \frac{\beta - \xi}{c} e^{\varepsilon\alpha} + \frac{\xi - \alpha}{c} e^{\varepsilon\beta} \quad \text{(п.н.).}$$

# Неравенство Азумы-Хёффдинга

Применяя свойства условного математического ожидания, получим

$$\mathbf{E}\left[e^{\varepsilon\xi}\,\middle|\,\mathcal{E}\right] \leqslant \left| \begin{matrix}\text{Утв. 1.17}\\ \text{(2, 3)}\end{matrix} \right| \leqslant \frac{\beta}{c}\,e^{\varepsilon\alpha} - \frac{e^{\varepsilon\alpha}}{c}\mathbf{E}[\xi\,|\,\mathcal{E}] + \frac{e^{\varepsilon\beta}}{c}\mathbf{E}[\xi\,|\,\mathcal{E}] - \frac{\alpha}{c}\,e^{\varepsilon\beta} \\ = \frac{\beta}{c}\,e^{\varepsilon\alpha} - \frac{\alpha}{c}\,e^{\varepsilon\beta} \quad \text{(п.н.)}. \tag{16}$$

# Неравенство Азумы-Хёффдинга

Применяя свойства условного математического ожидания, получим

$$\mathbf{E}\left[e^{\varepsilon\xi} \,|\, \mathcal{E}\right] \leqslant \left| \begin{smallmatrix}\text{Утв. 1.17}\\ (2,\ 3)\end{smallmatrix} \right| \leqslant \frac{\beta}{c} e^{\varepsilon\alpha} - \frac{e^{\varepsilon\alpha}}{c} \mathbf{E}[\xi \,|\, \mathcal{E}] + \frac{e^{\varepsilon\beta}}{c} \mathbf{E}[\xi \,|\, \mathcal{E}] - \frac{\alpha}{c} e^{\varepsilon\beta}$$
$$= \frac{\beta}{c} e^{\varepsilon\alpha} - \frac{\alpha}{c} e^{\varepsilon\beta} \quad (\text{п.н.}). \tag{16}$$

Повторяя соответствующие шаги из доказательства леммы Хёффдинга, можно показать, что

$$\frac{\beta}{c} e^{\varepsilon\alpha} - \frac{\alpha}{c} e^{\varepsilon\beta} \leqslant e^{\frac{\varepsilon^2 c^2}{8}} \quad (\text{п.н.}). \tag{17}$$

# Неравенство Азумы-Хёффдинга

Применяя свойства условного математического ожидания, получим

$$\mathbf{E}\left[e^{\varepsilon\xi} \,\middle|\, \mathcal{E}\right] \leqslant \left|\begin{smallmatrix}\text{Утв. 1.17}\\ (2,\ 3)\end{smallmatrix}\right| \leqslant \frac{\beta}{c}\, e^{\varepsilon\alpha} - \frac{e^{\varepsilon\alpha}}{c}\mathbf{E}[\xi \,|\, \mathcal{E}] + \frac{e^{\varepsilon\beta}}{c}\mathbf{E}[\xi \,|\, \mathcal{E}] - \frac{\alpha}{c}\, e^{\varepsilon\beta}$$
$$= \frac{\beta}{c}\, e^{\varepsilon\alpha} - \frac{\alpha}{c}\, e^{\varepsilon\beta} \quad (\text{п.н.}). \qquad (16)$$

Повторяя соответствующие шаги из доказательства леммы Хёффдинга, можно показать, что

$$\frac{\beta}{c}\, e^{\varepsilon\alpha} - \frac{\alpha}{c}\, e^{\varepsilon\beta} \leqslant e^{\frac{\varepsilon^2 c^2}{8}} \quad (\text{п.н.}). \qquad (17)$$

Объединяя (16) и (17), получим искомое неравенство (15).

■

# Неравенство Азумы-Хёффдинга

## Определение 2.2.

Последовательность пар $\{(\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$ ($n \in \mathbb{N}$), состоящих из случайной величины и $\sigma$-алгебры, называется мартингалом, если одновременно выполняются следующие условия

- $\mathcal{G}_0 \subseteq \mathcal{G}_1 \subseteq \ldots \mathcal{G}_n \subseteq \mathcal{G}$;
- $\xi_i \in \mathcal{L}(\Omega, \mathcal{G}_i, \mathsf{P})$ ($i = 0, \ldots, n$);
- $\xi_{i-1} = \mathbf{E}[\xi_i \,|\, \mathcal{G}_{i-1}]$ (п.н.) ($i = 1, \ldots, n$).

# Неравенство Азумы-Хёффдинга

**Теорема 2.2 (неравенство Азумы-Хёффдинга).**

Пусть $\{(\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$ ($n \in \mathbb{N}$) – мартингал. Предположим, что существуют положительные числа $c_1, \ldots, c_n$ такие, что

$$|\xi_i - \xi_{i-1}| \leqslant c_i \quad (\text{п.н.}) \quad (i = 1, \ldots, n).$$

# Неравенство Азумы-Хёффдинга

## Теорема 2.2 (неравенство Азумы-Хёффдинга).

Пусть $\{(\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$ ($n \in \mathbb{N}$) – мартингал. Предположим, что существуют положительные числа $c_1, \ldots, c_n$ такие, что

$$|\xi_i - \xi_{i-1}| \leqslant c_i \quad (\text{п.н.}) \quad (i = 1, \ldots, n).$$

Тогда для любого $\varepsilon > 0$ выполняются неравенства

$$\mathsf{P}\{\xi_n - \xi_0 \geqslant \varepsilon\} \leqslant \exp\left(-\frac{\varepsilon^2}{2\sum_{i=1}^n c_i^2}\right), \tag{18}$$

$$\mathsf{P}\{\xi_n - \xi_0 \leqslant -\varepsilon\} \leqslant \exp\left(-\frac{\varepsilon^2}{2\sum_{i=1}^n c_i^2}\right), \tag{19}$$

# Неравенство Азумы-Хёффдинга

## Теорема 2.2 (неравенство Азумы-Хёффдинга).

Пусть $\{(\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$ ($n \in \mathbb{N}$) – мартингал. Предположим, что существуют положительные числа $c_1, \ldots, c_n$ такие, что

$$|\xi_i - \xi_{i-1}| \leqslant c_i \quad \text{(п.н.)} \quad (i = 1, \ldots, n).$$

Тогда для любого $\varepsilon > 0$ выполняются неравенства

$$\mathsf{P}\{\xi_n - \xi_0 \geqslant \varepsilon\} \leqslant \exp\left(-\frac{\varepsilon^2}{2\sum_{i=1}^n c_i^2}\right), \tag{18}$$

$$\mathsf{P}\{\xi_n - \xi_0 \leqslant -\varepsilon\} \leqslant \exp\left(-\frac{\varepsilon^2}{2\sum_{i=1}^n c_i^2}\right), \tag{19}$$

а значит и

$$\mathsf{P}\{|\xi_n - \xi_0| \geqslant \varepsilon\} \leqslant 2\exp\left(-\frac{\varepsilon^2}{2\sum_{i=1}^n c_i^2}\right). \tag{20}$$

# Неравенство Азумы-Хёффдинга

◀ Неравенство (19) представляет собой вариант неравенства (18), записанный для мартингала $\{(-\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$. Неравенство (20) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий. Таким образом, достаточно доказать неравенство (18).

## Неравенство Азумы-Хёффдинга

◀ Неравенство (19) представляет собой вариант неравенства (18), записанный для мартингала $\{(-\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$. Неравенство (20) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий. Таким образом, достаточно доказать неравенство (18).

Определим случайные величины

$$\zeta_i := \xi_i - \xi_{i-1} \quad (i = 1, \ldots, n)$$

и заметим, что $|\zeta_i| \leqslant c_i$ и $\mathbf{E}[\zeta_i \mid \mathcal{G}_{i-1}] = 0$ (п.н.).

# Неравенство Азумы-Хёффдинга

◀ Неравенство (19) представляет собой вариант неравенства (18), записанный для мартингала $\{(-\xi_i, \mathcal{G}_i);\ i = 0, 1, \ldots, n\}$. Неравенство (20) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий. Таким образом, достаточно доказать неравенство (18).

Определим случайные величины

$$\zeta_i := \xi_i - \xi_{i-1} \quad (i = 1, \ldots, n)$$

и заметим, что $|\zeta_i| \leqslant c_i$ и $\mathbf{E}[\zeta_i \,|\, \mathcal{G}_{i-1}] = 0$ (п.н.).

Действительно, используя определение мартингала и свойства условного математического ожидания, получим

$$\begin{aligned}\mathbf{E}[\zeta_i \,|\, \mathcal{G}_{i-1}] &= \mathbf{E}[\xi_i - \xi_{i-1} \,|\, \mathcal{G}_{i-1}] \\ &= \mathbf{E}[\xi_i \,|\, \mathcal{G}_{i-1}] - \mathbf{E}[\xi_{i-1} \,|\, \mathcal{G}_{i-1}] \\ &= \xi_{i-1} - \xi_{i-1} = 0 \quad \text{(п.н.)} \quad (i = 1, \ldots, n).\end{aligned}$$

# Неравенство Азумы-Хёффдинга

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{\varepsilon}{\sum_{i=1}^{n} c_i^2}. \tag{21}$$

# Неравенство Азумы-Хёффдинга

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{\varepsilon}{\sum_{i=1}^{n} c_i^2}. \tag{21}$$

Из леммы 2.3 (неравенство Хёффдинга для условного математического ожидания) следуют неравенства

$$\mathbf{E}\left[e^{s\zeta_i} \,\middle|\, \mathcal{G}_{i-1}\right] \leqslant \exp\left(\frac{s^2 c_i^2}{2}\right) \quad \text{(п.н.)} \quad (i = 1, \ldots, n). \tag{22}$$

# Неравенство Азумы-Хёффдинга

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{\varepsilon}{\sum_{i=1}^{n} c_i^2}. \tag{21}$$

Из леммы 2.3 (неравенство Хёффдинга для условного математического ожидания) следуют неравенства

$$\mathbf{E}\left[e^{s\zeta_i} \,\middle|\, \mathcal{G}_{i-1}\right] \leqslant \exp\left(\frac{s^2 c_i^2}{2}\right) \quad (\text{п.н.}) \quad (i = 1, \ldots, n). \tag{22}$$

Заметим, что

$$\xi_n - \xi_0 = \sum_{i=1}^{n} \zeta_i.$$

# Неравенство Азумы-Хёффдинга

Используя неравенство Маркова, запишем

$$\mathsf{P}\left\{\xi_n - \xi_0 \geqslant \varepsilon\right\} = \mathsf{P}\left\{e^{s(\xi_n - \xi_0)} \geqslant e^{s\varepsilon}\right\} \leqslant \mathbf{E}\left[\exp\left(-s\varepsilon + s\sum_{i=1}^{n} \zeta_i\right)\right]. \quad (23)$$

# Неравенство Азумы-Хёффдинга

Используя неравенство Маркова, запишем

$$\mathsf{P}\Big\{\xi_n - \xi_0 \geqslant \varepsilon\Big\} = \mathsf{P}\Big\{e^{s(\xi_n - \xi_0)} \geqslant e^{s\varepsilon}\Big\} \leqslant \mathbf{E}\left[\exp\left(-s\varepsilon + s\sum_{i=1}^{n}\zeta_i\right)\right]. \quad (23)$$

С помощью свойств условного математического ожидания оценим правую часть этого неравенства

$$\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right] = \left|\begin{array}{c}\text{Утв. 1.17}\\ (6)\end{array}\right| = \mathbf{E}\left[\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right]\,\Big|\,\mathcal{G}_{n-1}\right]$$

$$= \left|\begin{array}{c}\text{Утв. 1.17}\\ (1)\end{array}\right| = \mathbf{E}\left[\exp\left(s\sum_{i=1}^{n-1}\zeta_i\right)\mathbf{E}\left[\exp(s\zeta_n)\,\big|\,\mathcal{G}_{n-1}\right]\right]$$

$$\leqslant \left|\begin{array}{c}\text{Неравенство}\\ (22)\end{array}\right| \leqslant \mathbf{E}\left[\exp\left(s\sum_{i=1}^{n-1}\zeta_i\right)\right]\exp\left(\frac{s^2c_n^2}{2}\right),$$

# Неравенство Азумы-Хёффдинга

а значит

$$\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right] \leqslant \prod_{i=1}^{n}\exp\left(\frac{s^2c_i^2}{2}\right) = \exp\left(\frac{s^2}{2}\sum_{i=1}^{n}c_i^2\right). \qquad (24)$$

# Неравенство Азумы-Хёффдинга

а значит

$$\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right] \leqslant \prod_{i=1}^{n}\exp\left(\frac{s^2c_i^2}{2}\right) = \exp\left(\frac{s^2}{2}\sum_{i=1}^{n}c_i^2\right). \qquad (24)$$

Объединяя (21), (23) и (24), получим искомое неравенство (18).

■

# Содержание

# Неравенство МакДиармида

## Определение 2.3.

Пусть X – непустое множество. Функция $g : X^n \longrightarrow \mathbb{R}$ обладает свойством ограниченности покоординатных приращений, если существуют положительные числа $c_1, \ldots, c_n$ такие, что

$$\left| g(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - g(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \right| < c_i \tag{25}$$

$(i = 1, \ldots, n)$ для любых $x, x_1, \ldots, x_n \in X$.

# Неравенство МакДиармида

**Теорема 2.3 (неравенство МакДиармида).**

Пусть независимые случайные элементы $\xi_1, \dots, \xi_n$ принимают свои значения в измеримом пространстве $(\mathsf{X}, \mathcal{X})$, а измеримая функция $g : \mathsf{X}^n \longrightarrow \mathbb{R}$ обладает свойством ограниченности покоординатных приращений (25). Определим случайную величину $\zeta := g(\xi_1, \dots, \xi_n)$.

# Неравенство МакДиармида

### Теорема 2.3 (неравенство МакДиармида).

Пусть независимые случайные элементы $\xi_1, \ldots, \xi_n$ принимают свои значения в измеримом пространстве $(\mathsf{X}, \mathcal{X})$, а измеримая функция $g : \mathsf{X}^n \longrightarrow \mathbb{R}$ обладает свойством ограниченности покоординатных приращений (25). Определим случайную величину $\zeta := g(\xi_1, \ldots, \xi_n)$. Тогда для любого $\varepsilon > 0$ выполняются неравенства

$$\mathsf{P}\{\zeta - \mathbf{E}\zeta \geqslant \varepsilon\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right), \tag{26}$$

$$\mathsf{P}\{\zeta - \mathbf{E}\zeta \leqslant -\varepsilon\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right), \tag{27}$$

# Неравенство МакДиармида

**Теорема 2.3 (неравенство МакДиармида).**

Пусть независимые случайные элементы $\xi_1, \ldots, \xi_n$ принимают свои значения в измеримом пространстве $(\mathsf{X}, \mathcal{X})$, а измеримая функция $g : \mathsf{X}^n \longrightarrow \mathbb{R}$ обладает свойством ограниченности покоординатных приращений (25). Определим случайную величину $\zeta := g(\xi_1, \ldots, \xi_n)$. Тогда для любого $\varepsilon > 0$ выполняются неравенства

$$\mathsf{P}\{\zeta - \mathbf{E}\zeta \geqslant \varepsilon\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right), \tag{26}$$

$$\mathsf{P}\{\zeta - \mathbf{E}\zeta \leqslant -\varepsilon\} \leqslant \exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right), \tag{27}$$

а значит и

$$\mathsf{P}\{|\zeta - \mathbf{E}\zeta| \geqslant \varepsilon\} \leqslant 2\exp\left(-\frac{2\varepsilon^2}{\sum_{i=1}^n c_i^2}\right). \tag{28}$$

# Неравенство МакДиармида

◀ Прежде всего заметим, что функция $-g$ также является измеримой и обладает свойством ограниченности покоординатных приращений (25). Записывая для неё неравенство (26), получим неравенство (27).

# Неравенство МакДиармида

◀ Прежде всего заметим, что функция $-g$ также является измеримой и обладает свойством ограниченности покоординатных приращений (25). Записывая для неё неравенство (26), получим неравенство (27).

Неравенство (28) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

# Неравенство МакДиармида

◀ Прежде всего заметим, что функция $-g$ также является измеримой и обладает свойством ограниченности покоординатных приращений (25). Записывая для неё неравенство (26), получим неравенство (27).

Неравенство (28) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

Таким образом, достаточно доказать неравенство (26).

# Неравенство МакДиармида

◀ Прежде всего заметим, что функция $-g$ также является измеримой и обладает свойством ограниченности покоординатных приращений (25). Записывая для неё неравенство (26), получим неравенство (27).

Неравенство (28) является непосредственным следствием предыдущих двух неравенств и известного свойства вероятности объединения событий.

Таким образом, достаточно доказать неравенство (26).

Построим последовательность вложенных друг в друга $\sigma$-алгебр

$$\mathcal{G}_0 := \{\varnothing, \Omega\}, \quad \mathcal{G}_i := \sigma\{\xi_1, \ldots, \xi_i\} \quad (i = 1, \ldots, n),$$

для которых выполняется условие $\mathcal{G}_0 \subseteq \mathcal{G}_1 \subseteq \ldots \subseteq \mathcal{G}_n$.

# Неравенство МакДиармида

Определим последовательность случайных величин

$$\zeta_i := \mathbf{E}[\zeta \mid \mathcal{G}_i] - \mathbf{E}[\zeta \mid \mathcal{G}_{i-1}] \quad (i = 1, \ldots, n) \tag{29}$$

# Неравенство МакДиармида

Определим последовательность случайных величин

$$\zeta_i := \mathbf{E}[\zeta \,|\, \mathcal{G}_i] - \mathbf{E}[\zeta \,|\, \mathcal{G}_{i-1}] \quad (i = 1, \ldots, n) \tag{29}$$

и заметим, что

$$\mathbf{E}[\zeta \,|\, \mathcal{G}_0] = \mathbf{E}\zeta, \quad \mathbf{E}[\zeta \,|\, \mathcal{G}_n] = \zeta \quad (\text{п.н.}). \tag{30}$$

# Неравенство МакДиармида

Определим последовательность случайных величин

$$\zeta_i := \mathbf{E}[\zeta \,|\, \mathcal{G}_i] - \mathbf{E}[\zeta \,|\, \mathcal{G}_{i-1}] \quad (i = 1, \ldots, n) \tag{29}$$

и заметим, что

$$\mathbf{E}[\zeta \,|\, \mathcal{G}_0] = \mathbf{E}\zeta, \quad \mathbf{E}[\zeta \,|\, \mathcal{G}_n] = \zeta \quad (\text{п.н.}). \tag{30}$$

Из равенств (29) и (30) следует, что

$$\zeta - \mathbf{E}\zeta = \sum_{i=1}^{n} \zeta_i \quad (\text{п.н.}). \tag{31}$$

# Неравенство МакДиармида

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{4\varepsilon}{\sum_{i=1}^{n} c_i^2}. \tag{32}$$

# Неравенство МакДиармида

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{4\varepsilon}{\sum_{i=1}^{n} c_i^2}. \tag{32}$$

Основная идея доказательства теоремы состоит в получении с помощью леммы 2.3 (неравенство Хёффдинга для условного математического ожидания) и использовании следующих неравенств

$$\mathbf{E}\left[e^{s\zeta_i} \,\middle|\, \mathcal{G}_{i-1}\right] \leqslant \exp\left(\frac{s^2 c_i^2}{8}\right) \quad \text{(п.н.)} \quad (i = 1, \dots, n). \tag{33}$$

# Неравенство МакДиармида

Зафиксируем произвольное $\varepsilon > 0$ и определим число

$$s := \frac{4\varepsilon}{\sum_{i=1}^{n} c_i^2}. \tag{32}$$

Основная идея доказательства теоремы состоит в получении с помощью леммы 2.3 (неравенство Хёффдинга для условного математического ожидания) и использовании следующих неравенств

$$\mathbf{E}\left[e^{s\zeta_i} \mid \mathcal{G}_{i-1}\right] \leqslant \exp\left(\frac{s^2 c_i^2}{8}\right) \quad (\text{п.н.}) \quad (i = 1, \ldots, n). \tag{33}$$

Предположим, что эти неравенства выполнены. Тогда с помощью свойств условного математического ожидания получим оценку

# Неравенство МакДиармида

$$\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right] = \left|\begin{array}{c}\text{Утв. 1.17}\\ (6)\end{array}\right| = \mathbf{E}\left[\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right]\middle|\,\mathcal{G}_{n-1}\right]$$

$$= \left|\begin{array}{c}\text{Утв. 1.17}\\ (1)\end{array}\right| = \mathbf{E}\left[\exp\left(s\sum_{i=1}^{n-1}\zeta_i\right)\mathbf{E}\left[\exp(s\zeta_n)\,\middle|\,\mathcal{G}_{n-1}\right]\right]$$

$$\leqslant \left|\begin{array}{c}\text{Неравенство}\\ (33)\end{array}\right| \leqslant \mathbf{E}\left[\exp\left(s\sum_{i=1}^{n-1}\zeta_i\right)\right]\exp\left(\frac{s^2c_n^2}{8}\right),$$

# Неравенство МакДиармида

$$\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right] = \left|\begin{array}{c}\text{Утв. 1.17}\\ (6)\end{array}\right| = \mathbf{E}\left[\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right]\,\middle|\,\mathcal{G}_{n-1}\right]$$

$$= \left|\begin{array}{c}\text{Утв. 1.17}\\ (1)\end{array}\right| = \mathbf{E}\left[\exp\left(s\sum_{i=1}^{n-1}\zeta_i\right)\mathbf{E}\left[\exp(s\zeta_n)\,\middle|\,\mathcal{G}_{n-1}\right]\right]$$

$$\leqslant \left|\begin{array}{c}\text{Неравенство}\\ (33)\end{array}\right| \leqslant \mathbf{E}\left[\exp\left(s\sum_{i=1}^{n-1}\zeta_i\right)\right]\exp\left(\frac{s^2c_n^2}{8}\right),$$

а значит

$$\mathbf{E}\left[\exp\left(s\sum_{i=1}^{n}\zeta_i\right)\right] \leqslant \prod_{i=1}^{n}\exp\left(\frac{s^2c_i^2}{8}\right) = \exp\left(\frac{s^2}{8}\sum_{i=1}^{n}c_i^2\right). \qquad (34)$$

# Неравенство МакДиармида

Используя неравенство Маркова и (31), запишем

$$\mathsf{P}\Big\{\zeta - \mathbf{E}\zeta \geqslant \varepsilon\Big\} = \mathsf{P}\Big\{e^{s(\zeta-\mathbf{E}\zeta)} \geqslant e^{s\varepsilon}\Big\} \leqslant \mathbf{E}\left[\exp\left(-s\varepsilon + s\sum_{i=1}^{n}\zeta_i\right)\right]. \quad (35)$$

# Неравенство МакДиармида

Используя неравенство Маркова и (31), запишем

$$\mathrm{P}\left\{\zeta - \mathbf{E}\zeta \geqslant \varepsilon\right\} = \mathrm{P}\left\{e^{s(\zeta - \mathbf{E}\zeta)} \geqslant e^{s\varepsilon}\right\} \leqslant \mathbf{E}\left[\exp\left(-s\varepsilon + s\sum_{i=1}^{n}\zeta_i\right)\right]. \quad (35)$$

Объединяя (32), (34) и (35), получим требуемое неравенство (26).

# Неравенство МакДиармида

Используя неравенство Маркова и (31), запишем

$$\mathrm{P}\left\{\zeta - \mathbf{E}\zeta \geqslant \varepsilon\right\} = \mathrm{P}\left\{e^{s(\zeta - \mathbf{E}\zeta)} \geqslant e^{s\varepsilon}\right\} \leqslant \mathbf{E}\left[\exp\left(-s\varepsilon + s\sum_{i=1}^{n}\zeta_i\right)\right]. \quad (35)$$

Объединяя (32), (34) и (35), получим требуемое неравенство (26).

Перейдем к доказательству неравенств (33).

# Неравенство МакДиармида

Используя неравенство Маркова и (31), запишем

$$\mathsf{P}\Big\{\zeta - \mathbf{E}\zeta \geqslant \varepsilon\Big\} = \mathsf{P}\Big\{e^{s(\zeta - \mathbf{E}\zeta)} \geqslant e^{s\varepsilon}\Big\} \leqslant \mathbf{E}\left[\exp\left(-s\varepsilon + s\sum_{i=1}^{n}\zeta_i\right)\right]. \quad (35)$$

Объединяя (32), (34) и (35), получим требуемое неравенство (26).

Перейдем к доказательству неравенств (33).

При выполнении следующих двух условий эти неравенства будут являться следствием леммы 2.3 (неравенство Хёффдинга для условного математического ожидания).

# Неравенство МакДиармида

Во-первых, условные математические ожидания $\mathbf{E}[\zeta_i \,|\, \mathcal{G}_{i-1}]$ должны быть равны нулю.

# Неравенство МакДиармида

Во-первых, условные математические ожидания $\mathbf{E}[\zeta_i \,|\, \mathcal{G}_{i-1}]$ должны быть равны нулю.

Во-вторых, должны существовать $\mathcal{G}_{i-1}$-измеримые случайные величины $\alpha_i$, удовлетворяющие неравенствам

$$\alpha_i \leqslant \zeta_i \leqslant \alpha_i + c_i \quad (\text{п.н.}) \quad (i = 1, \dots, n). \tag{36}$$

# Неравенство МакДиармида

Во-первых, условные математические ожидания $\mathbf{E}[\zeta_i \mid \mathcal{G}_{i-1}]$ должны быть равны нулю.

Во-вторых, должны существовать $\mathcal{G}_{i-1}$-измеримые случайные величины $\alpha_i$, удовлетворяющие неравенствам

$$\alpha_i \leqslant \zeta_i \leqslant \alpha_i + c_i \quad (\text{п.н.}) \quad (i = 1, \ldots, n). \tag{36}$$

Проверим выполнение первого условия.

# Неравенство МакДиармида

Во-первых, условные математические ожидания $\mathbf{E}[\zeta_i \,|\, \mathcal{G}_{i-1}]$ должны быть равны нулю.

Во-вторых, должны существовать $\mathcal{G}_{i-1}$-измеримые случайные величины $\alpha_i$, удовлетворяющие неравенствам

$$\alpha_i \leqslant \zeta_i \leqslant \alpha_i + c_i \quad (\text{п.н.}) \quad (i = 1, \ldots, n). \tag{36}$$

Проверим выполнение первого условия. Используя линейность и телескопическое свойство условного математического ожидания, а также вложенность рассматриваемых $\sigma$-алгебр, получим

$$\begin{aligned}
\mathbf{E}[\zeta_i \,|\, \mathcal{G}_{i-1}] &= \mathbf{E}[\, \mathbf{E}[\zeta \,|\, \mathcal{G}_i] - \mathbf{E}[\zeta \,|\, \mathcal{G}_{i-1}] \,|\, \mathcal{G}_{i-1}] \\
&= \mathbf{E}[\, \mathbf{E}[\zeta \,|\, \mathcal{G}_i] \,|\, \mathcal{G}_{i-1}] - \mathbf{E}[\zeta \,|\, \mathcal{G}_{i-1}] \\
&= \mathbf{E}[\zeta \,|\, \mathcal{G}_{i-1}] - \mathbf{E}[\zeta \,|\, \mathcal{G}_{i-1}] = 0 \quad (\text{п.н.}) \quad (i = 1, \ldots, n).
\end{aligned}$$

# Неравенство МакДиармида

Проверим выполнение второго условия.

# Неравенство МакДиармида

Проверим выполнение второго условия.

Учитывая независимость случайных элементов $\xi_1, \ldots, \xi_n$, из утверждения 1.18 следует, что для случайных величин $\zeta_1, \ldots, \zeta_n$ имеет место следующее представление

$$\zeta_i = g_i(\xi_1, \ldots, \xi_i) \quad (\text{п.н.}) \quad (i = 1, \ldots, n),$$

в котором измеримые функции $g_i : \mathsf{X}^i \longrightarrow \mathbb{R}$ задаются по правилу

$$g_i(x_1, \ldots, x_i) := \mathbf{E}g(x_1, \ldots, x_i, \xi_{i+1}, \ldots, \xi_n) - \mathbf{E}g(x_1, \ldots, x_{i-1}, \xi_i, \ldots, \xi_n). \tag{37}$$

# Неравенство МакДиармида

Проверим выполнение второго условия.

Учитывая независимость случайных элементов $\xi_1, \ldots, \xi_n$, из утверждения 1.18 следует, что для случайных величин $\zeta_1, \ldots, \zeta_n$ имеет место следующее представление

$$\zeta_i = g_i(\xi_1, \ldots, \xi_i) \quad (\text{п.н.}) \quad (i = 1, \ldots, n),$$

в котором измеримые функции $g_i : \mathsf{X}^i \longrightarrow \mathbb{R}$ задаются по правилу

$$g_i(x_1, \ldots, x_i) := \mathbf{E}g(x_1, \ldots, x_i, \xi_{i+1}, \ldots, \xi_n) - \mathbf{E}g(x_1, \ldots, x_{i-1}, \xi_i, \ldots, \xi_n). \tag{37}$$

Учитывая лемму 1.2, будем предполагать измеримыми следующие функции

$$a_i(x_1, \ldots, x_{i-1}) := \inf_{x \in \mathsf{X}} g_i(x_1, \ldots, x_{i-1}, x) \quad (i = 1, \ldots, n).$$

# Неравенство МакДиармида

По определению положим

$$\alpha_i := a_i(\xi_1, \ldots, \xi_{i-1}) \quad (i = 1, \ldots, n). \tag{38}$$

# Неравенство МакДиармида

По определению положим

$$\alpha_i := a_i(\xi_1, \dots, \xi_{i-1}) \quad (i = 1, \dots, n). \tag{38}$$

Из свойства ограниченности покоординатных приращений (25) функции $g$ и определения (37) следует, что для любых $x, x', x_1, \dots, x_n \in \mathrm{X}$ выполняются неравенства

$$g_i(x_1, \dots, x_{i-1}, x) - g_i(x_1, \dots, x_{i-1}, x') \leqslant c_i \quad (i = 1, \dots, n),$$

# Неравенство МакДиармида

По определению положим

$$\alpha_i := a_i(\xi_1, \ldots, \xi_{i-1}) \quad (i = 1, \ldots, n). \tag{38}$$

Из свойства ограниченности покоординатных приращений (25) функции $g$ и определения (37) следует, что для любых $x, x', x_1, \ldots, x_n \in \mathrm{X}$ выполняются неравенства

$$g_i(x_1, \ldots, x_{i-1}, x) - g_i(x_1, \ldots, x_{i-1}, x') \leqslant c_i \quad (i = 1, \ldots, n),$$

а значит и

$$\sup_{x \in \mathrm{X}} g_i(x_1, \ldots, x_{i-1}, x) - \inf_{x \in \mathrm{X}} g_i(x_1, \ldots, x_{i-1}, x) \leqslant c_i \quad (i = 1, \ldots, n).$$

# Неравенство МакДиармида

Таким образом, для любых $x_1, \ldots, x_n \in \mathrm{X}$ выполняются неравенства

$$a_i(x_1, \ldots, x_{i-1}) \leqslant g_i(x_1, \ldots, x_i) \leqslant a_i(x_1, \ldots, x_{i-1}) + c_i \quad (i = 1, \ldots, n),$$

из которых, учитывая определения (38), вытекает справедливость неравенств (36).

$\blacksquare$

# Содержание

# Заключительный пример

В заключение продемонстрируем эффективность подхода, базирующегося на использовании неравенств концентрации меры, на примере решения классической задачи оценки параметра семейства вероятностных распределений.

Данный пример позволит осветить некоторые ключевые идеи, лежащие в основе статистической теории обучения.

# Заключительный пример

Рассмотрим независимые и одинаково распределённые бернулливские случайные величины $\xi_1, \xi_2, \ldots$, принимающие значение 1 с вероятностью $\theta$ и значение 0 с вероятностью $1 - \theta$.

# Заключительный пример

Рассмотрим независимые и одинаково распределённые бернулливские случайные величины $\xi_1, \xi_2, \ldots$, принимающие значение 1 с вероятностью $\theta$ и значение 0 с вероятностью $1 - \theta$.

Заметим, что $\mathbf{E}\xi_1 = \theta$ и $\mathbf{D}\xi_1 = \theta(1 - \theta)$.

# Заключительный пример

Рассмотрим независимые и одинаково распределённые бернулливские случайные величины $\xi_1, \xi_2, \ldots$, принимающие значение 1 с вероятностью $\theta$ и значение 0 с вероятностью $1 - \theta$.

Заметим, что $\mathbf{E}\xi_1 = \theta$ и $\mathbf{D}\xi_1 = \theta(1 - \theta)$.

В качестве статистической оценки для параметра $\theta$ будем использовать случайную величину

$$\widehat{\theta}_n := \frac{1}{n} \sum_{i=1}^{n} \xi_i \quad (n \in \mathbb{N}).$$

# Заключительный пример

Рассмотрим независимые и одинаково распределённые бернулливские случайные величины $\xi_1, \xi_2, \ldots$, принимающие значение 1 с вероятностью $\theta$ и значение 0 с вероятностью $1 - \theta$.

Заметим, что $\mathbf{E}\xi_1 = \theta$ и $\mathbf{D}\xi_1 = \theta(1 - \theta)$.

В качестве статистической оценки для параметра $\theta$ будем использовать случайную величину

$$\widehat{\theta}_n := \frac{1}{n} \sum_{i=1}^{n} \xi_i \quad (n \in \mathbb{N}).$$

Зафиксируем произвольное $\varepsilon > 0$.

# Заключительный пример

Рассмотрим независимые и одинаково распределённые бернулливские случайные величины $\xi_1, \xi_2, \ldots$, принимающие значение 1 с вероятностью $\theta$ и значение 0 с вероятностью $1 - \theta$.

Заметим, что $\mathbf{E}\xi_1 = \theta$ и $\mathbf{D}\xi_1 = \theta(1 - \theta)$.

В качестве статистической оценки для параметра $\theta$ будем использовать случайную величину

$$\widehat{\theta}_n := \frac{1}{n} \sum_{i=1}^{n} \xi_i \quad (n \in \mathbb{N}).$$

Зафиксируем произвольное $\varepsilon > 0$.

Заметим, что $\mathbf{E}\,\widehat{\theta}_n = \theta$ и $\mathbf{D}\,\widehat{\theta}_n = \frac{1}{n}\,\theta(1 - \theta)$.

# Заключительный пример

Применяя неравенство Чебышёва, получим

$$\mathsf{P}\left\{\left|\widehat{\theta}_n - \theta\right| \geqslant \varepsilon\right\} \leqslant \frac{\theta(1-\theta)}{n\varepsilon^2}. \tag{39}$$

# Заключительный пример

Применяя неравенство Чебышёва, получим

$$\mathsf{P}\left\{\left|\widehat{\theta}_n - \theta\right| \geqslant \varepsilon\right\} \leqslant \frac{\theta(1-\theta)}{n\varepsilon^2}. \tag{39}$$

Неравенство (39) показывает, что вероятность построить неудовлетворительную оценку приближаемого параметра, убывает обратно пропорционально размеру случайной выборки.

# Заключительный пример

Применяя неравенство Чебышёва, получим

$$\mathsf{P}\left\{\left|\widehat{\theta}_n - \theta\right| \geqslant \varepsilon\right\} \leqslant \frac{\theta(1-\theta)}{n\varepsilon^2}. \tag{39}$$

Неравенство (39) показывает, что вероятность построить неудовлетворительную оценку приближаемого параметра, убывает обратно пропорционально размеру случайной выборки.

Скорость убывания достаточно медленная. Улучшить положение можно за счёт использования центральной предельной теоремы.

# Заключительный пример

**Теорема 2.4 (центральная предельная теорема).**

Пусть $\{\zeta_n\}_{n\in\mathbb{N}}$ – последовательность независимых одинаково распределённых и невырожденных случайных величин с $\mathbf{E}\zeta_1^2 < \infty$. Положим $S_n := \zeta_1 + \ldots + \zeta_n$. Тогда

$$\mathsf{P}\left\{\frac{S_n - \mathbf{E}S_n}{\sqrt{\mathbf{D}S_n}} \leqslant x\right\} \longrightarrow \Phi(x) \quad \text{при} \quad n \longrightarrow \infty \quad (x \in \mathbb{R}),$$

где

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-y^2/2} dy.$$

# Заключительный пример

Обозначим

$$t := \varepsilon \sqrt{\frac{n}{\theta(1-\theta)}},$$

# Заключительный пример

Обозначим

$$t := \varepsilon\sqrt{\frac{n}{\theta(1-\theta)}},$$

тогда

$$P\left\{\sqrt{\frac{n}{\theta(1-\theta)}}\cdot(\widehat{\theta}_n-\theta)\geqslant t\right\}\longrightarrow 1-\Phi(t)=\frac{1}{\sqrt{2\pi}}\int\limits_t^{\infty}e^{-x^2/2}dx \qquad (40)$$

при $n\longrightarrow\infty$.

# Заключительный пример

Обозначим

$$t := \varepsilon \sqrt{\frac{n}{\theta(1-\theta)}},$$

тогда

$$\mathsf{P}\left\{\sqrt{\frac{n}{\theta(1-\theta)}} \cdot (\widehat{\theta}_n - \theta) \geqslant t\right\} \longrightarrow 1 - \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-x^2/2} dx \qquad (40)$$

при $n \longrightarrow \infty$.

Оценим интеграл, фигурирующий в (40). Получим

$$\int_t^{\infty} e^{-x^2/2} dx \leqslant \frac{1}{t} \int_t^{\infty} x e^{-x^2/2}\, dx = \frac{1}{t} \int_{t^2/2}^{\infty} e^{-y}\, dy = \frac{1}{t} \left[-e^{-y}\right]\Big|_{t^2/2}^{\infty} = \frac{1}{t} e^{-t^2/2}.$$

# Заключительный пример

Таким образом,

$$\mathsf{P}\left\{\widehat{\theta}_n - \theta \geqslant \varepsilon\right\} \lesssim \exp\left\{\frac{-n\varepsilon^2}{2\theta(1-\theta)}\right\}. \tag{41}$$

# Заключительный пример

Таким образом,

$$\mathsf{P}\left\{\widehat{\theta}_n - \theta \geqslant \varepsilon\right\} \lesssim \exp\left\{\frac{-n\varepsilon^2}{2\theta(1-\theta)}\right\}. \qquad (41)$$

Неравенство (41) гарантирует экспоненциальное убывание вероятности построния неудовлетворительной оценки приближаемого параметра относительно размера случайной выборки.

# Заключительный пример

Таким образом,

$$\mathrm{P}\left\{\widehat{\theta}_n - \theta \geqslant \varepsilon\right\} \lesssim \exp\left\{\frac{-n\varepsilon^2}{2\theta(1-\theta)}\right\}. \tag{41}$$

Неравенство (41) гарантирует экспоненциальное убывание вероятности построния неудовлетворительной оценки приближаемого параметра относительно размера случайной выборки.

К сожалению это неравенство является асимптотическим и не позволяет точно указать число наблюдений $n$, при котором вероятность рассматриваемого события не будет превосходить заданное значение $\delta$.

# Заключительный пример

Применяя неравенство Хёффдинга, можно получить точную оценку

$$\mathsf{P}\left\{\left|\widehat{\theta}_n - \theta\right| \geqslant \varepsilon\right\} \leqslant 2\,e^{-2n\varepsilon^2}. \tag{42}$$

# Заключительный пример

Применяя неравенство Хёффдинга, можно получить точную оценку

$$\mathsf{P}\left\{\left|\widehat{\theta}_n - \theta\right| \geqslant \varepsilon\right\} \leqslant 2\,e^{-2n\varepsilon^2}. \tag{42}$$

Ограничивая правую часть (42) величиной $\delta$, получим решение

$$n \geqslant n(\varepsilon, \delta) := \left\lceil \frac{1}{2\varepsilon^2} \ln\left(\frac{2}{\delta}\right) \right\rceil.$$

# Заключительный пример

Применяя неравенство Хёффдинга, можно получить точную оценку

$$\mathsf{P}\left\{\left|\widehat{\theta}_n - \theta\right| \geqslant \varepsilon\right\} \leqslant 2\,e^{-2n\varepsilon^2}. \tag{42}$$

Ограничивая правую часть (42) величиной $\delta$, получим решение

$$n \geqslant n(\varepsilon, \delta) := \left\lceil \frac{1}{2\varepsilon^2} \ln\left(\frac{2}{\delta}\right) \right\rceil.$$

Из описанного примера может быть выделена общая схема построения решения, которая обобщается и на другие классы задач.

# Заключительный пример

В этих задачах требуется построить некоторый объект.

# Заключительный пример

В этих задачах требуется построить некоторый объект.

В рассмотренном примере это была числовая характеристика $\theta$. Однако изначально на природу такого объекта не накладываются никакие ограничения. В качестве такого объекта может выступать, например, функция.

# Заключительный пример

В этих задачах требуется построить некоторый объект.

В рассмотренном примере это была числовая характеристика $\theta$. Однако изначально на природу такого объекта не накладываются никакие ограничения. В качестве такого объекта может выступать, например, функция.

Объект строится приближенно с помощью некоторого алгоритма, на вход которому подается набор значений (наблюдений), имеющих статистическую природу. При этом учитывается размер $n$ этого набора.

# Заключительный пример

В этих задачах требуется построить некоторый объект.

В рассмотренном примере это была числовая характеристика $\theta$. Однако изначально на природу такого объекта не накладываются никакие ограничения. В качестве такого объекта может выступать, например, функция.

Объект строится приближенно с помощью некоторого алгоритма, на вход которому подается набор значений (наблюдений), имеющих статистическую природу. При этом учитывается размер $n$ этого набора.

В рассмотренном примере это было значение, которое принимает случайна величина $\widehat{\theta}_n$.

# Заключительный пример

Имеется возможность оценить точность построенного приближения с помощью вещественного параметра $\varepsilon$.

# Заключительный пример

Имеется возможность оценить точность построенного приближения с помощью вещественного параметра $\varepsilon$.

Для рассматриваемого алгоритма существует функция сложности $n(\varepsilon, \delta)$, зависящая от значений параметров точности и достоверности и обладающая следующим свойством.

# Заключительный пример

Имеется возможность оценить точность построенного приближения с помощью вещественного параметра $\varepsilon$.

Для рассматриваемого алгоритма существует функция сложности $n(\varepsilon, \delta)$, зависящая от значений параметров точности и достоверности и обладающая следующим свойством.

Если размер набора входных значений $n \geqslant n(\varepsilon, \delta)$, то с вероятностью не меньшей, чем $1 - \delta$, алгоритм построит приближения с точностью $\varepsilon$.