Задание номер 1

Н. К. Животовский
nikita.zhivotovskiy@phystech.edu

6 марта 2017 г.

Задание принимается до 2.00 утра 6 марта по адресу slt.fupm.2017@gmail.com. В начале текста задания *обязательно* указывается:

- С кем вы делали это задание.
- Какие источники (кроме материалов лекций) вы использовали.

Задание оформляется в формате pdf (текст набирается в latex/Word) и в таком виде, чтобы ваши коллеги могли разобрать текст решения. Задания, оформленные не в соответствии с указанными правилами, не принимаются. Желательно оставлять зазоры между задачами для пометок.

Упражнение 1.

- Привести пример случайной величины $X \geq 0$, для которой в неравенстве Маркова достигается точное равенство.
- Привести пример случайной величины Y, для которой в неравенстве Чебышева достигается точное равенство.

Упражнение 2 [От хвостов к математическим ожиданиям]

- Пусть для с.в. $Y \geq 0$, некоторых $A \geq 2, B > 0$ и для любого $\varepsilon \geq 0$ выполнено

  $$P(Y \geq \varepsilon) \leq A \exp\left(-\frac{\varepsilon^2}{B^2}\right).$$

  Докажите, что существует абсолютная константа $C > 0$, такая что

  $$\mathbb{E}Y \leq CB\sqrt{\log(A)}.$$

- Какая оценка на $\mathbb{E}Y$ получится в случае $P(Y \geq \varepsilon) \leq A \exp\left(-\frac{\varepsilon}{B}\right)$?

Указание. Используйте, что для с.в. $Y \geq 0$ имеет место $\mathbb{E}Y = \int\limits_0^\infty P(Y > \varepsilon) d\varepsilon$.

■

Упражнение 3

- Докажите, что сумма двух субгассовских случайных величин с параметрами σ_1, σ_2 также является субгауссовской случайной величиной. Оценить сверху параметр σ для суммы через σ_1, σ_2. Достигается ли ваша оценка?
- Верно ли, что произведение субгауссовских также субгауссовское?
- Докажите, что условие $\mathbb{E}X = 0$ является необходимым условием того, что $\mathbb{E}\exp(\lambda X) \le \exp(\frac{\lambda^2\sigma^2}{2})$ для всех $\lambda \in \mathbb{R}$.
- Докажите, что если X — субгассовская случайная величина с параметром σ, то
$$\|X\|_{L_p} = (\mathbb{E}|X|^p)^{\frac{1}{p}} \le K\sqrt{p},$$
где K зависит только от σ для всех $p > 1$.

Упражнение 4 Пусть $\psi(x)$ — плотность распределения с. в. $X \sim \mathcal{N}(0,1)$.

- Докажите, что $\psi'(x) + x\psi(x) = 0$.
- Докажите, что для всех $x > 0$
$$\psi(x)\left(\frac{1}{x} - \frac{1}{x^3}\right) \le P(X \ge x) \le \psi(x)\left(\frac{1}{x} - \frac{1}{x^3} + \frac{3}{x^5}\right).$$
- Сравнить верхнюю оценку в последнем неравенстве с оценкой, получаемой с помощью метода Чернова.
- Сравнить для всех значений параметра p для схемы испытаний Бернулли хвост, получаемый с помощью неравенства Хеффдинга с асимптотическим хвостом, получаемым с помощью предельной теоремы Муавра-Лапласа. Верно ли, что хвосты Хеффдинга всегда слабее?

Упражнение 5 В задаче классификации с двумя классами $\mathcal{Y} = \{1, -1\}$ и бинарной функцией потерь

- Докажите что $f^*(x) = \operatorname{sign}(\eta(x))$, где $\eta(x) = E[Y|X = x]$.
- Пусть известно, что $|\eta(x)| \ge h$, где $h \in [0,1]$. Докажите, что $L(f^*) \le \frac{1}{2}(1-h)$.

Задача 1. Приведите пример класса функций $\mathcal{F}$, для которого алгоритм голосования большинства (Halving) неоптимален. То есть пример такого класса, для которого число ошибок на любой конечной выборке в худшем случае будет строго меньше, если не всегда следовать большинству.

Задача 2 Рассмотрим задачу классификации с двумя классами $\mathcal{Y} = \{1, -1\}$ и бинарной функцией потерь такую, что для некоторого $f \in \mathcal{F}$ выполнено $Y = f(X)$. Предположим, что мы заменили определение обучаемости следующим образом. Если ранее требовался малый риск с вероятностью $1 - \delta$, то теперь требуется лишь с вероятностью $\frac{1}{2}$. Докажите, что множество обучаемых классов не поменяется от такого изменения определения. Другими словами, если $\mathcal{F}$ обучаем в постановке $\frac{1}{2}$, то он обучаем и в постановке $1 - \delta$.