

## ВОПРОСЫ НА ПОНИМАНИЕ

Перед тем как приступить к решению домашних задач, попробуйте ответить на следующие вопросы. Это простые вопросы на понимание. Ответы на них включать в домашнюю работу не нужно.

1. Что такое функция правдоподобия? Что она означает в дискретном и непрерывном случаях?
2. Что такое функция распределения случайной величины?
3. Как по функции распределения  $F(x)$  случайной величины  $X$  найти ее плотность?
4. Как по функции распределения  $F(x)$  случайной величины  $X$  найти вероятность  $\mathbb{P}(X > 3)$ ?
5. Как по неравенству Чебышёва оценить вероятность  $\mathbb{P}(|X - \mathbb{E}X| \leq t)$ , если  $t > 0$  и  $\text{Var}X$  существует? Будет ли это оценка сверху или снизу?
6. Куда сходится среднее арифметическое независимых и одинаково распределённых случайных величин с конечной дисперсией?
7. Чему равно математическое ожидание и дисперсия величины  $\frac{S_n - \mathbb{E}S_n}{\sqrt{\text{Var}S_n}}$ ?
8. К какому распределению в условиях ЦПТ приближается распределение величины  $\frac{S_n - \mathbb{E}S_n}{\sqrt{\text{Var}S_n}}$ ?

## МЕТОД МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

**Упражнение 1** (12 баллов). Пусть дана реализация выборки  $x_1, \dots, x_n$  из равномерного распределения на отрезке  $[\theta; \theta + 1]$ . Найдите оценку для неизвестного параметра  $\theta$  методом максимального правдоподобия. (Тут нужно найти оценку теоретически, ничего реализовывать в Python не нужно.)

**Упражнение 2** (20 баллов). Пусть дана реализация выборки  $x_1, \dots, x_n$  их нормального распределения  $\mathcal{N}(\theta_1, \theta_2^2)$ . Найдите оценки для неизвестных параметров  $\theta_1$  и  $\theta_2^2$  методом максимального правдоподобия. Реализуйте эту задачу в Python:

- (1) сгенерируйте  $\theta_1$  из равномерного распределения на  $[-5, 5]$ , а  $\theta_2^2$  — из равномерного распределения на  $[0.5, 10]$ ;
- (2) сгенерируйте выборку из нормального распределения  $\mathcal{N}(\theta_1, \theta_2^2)$  размера  $n = 10, 100, 1\,000, 10\,000$ ;
- (3) найдите значения полученных оценок (и посчитав значения оценок максимального правдоподобия, и численно с помощью метода `fit()` из `SciPy`);
- (4) выведите отклонения полученных оценок от параметров  $\theta_1$  и  $\theta_2^2$ . Что происходит с ростом  $n$ ?

## ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

**Упражнение 3** (12 баллов). Случайная величина  $X$  равномерно распределена на отрезке  $[0, 1]$ . Докажите, что случайная величина

$$Y = -\frac{1}{\lambda} \ln(1 - X)$$

имеет экспоненциальное распределение с параметром  $\lambda > 0$ . Этот факт может использоваться для генерации выборки экспоненциального распределения с помощью равномерного: чтобы сгенерировать экспоненциальную случайную величину можно сгенерировать равномерно распределённую на  $[0, 1]$  случайную величину и применить к ней вышеприведённое преобразование.

**Упражнение 4** (12 баллов). Пусть  $X$  имеет стандартное нормальное распределение  $\mathcal{N}(0, 1)$ . Найдите плотность распределения случайной величины  $X^2$ . В статистике оно известно под именем «распределения хи-квадрат» с одной степенью свободы.

## ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ И КОНЦЕНТРАЦИЯ

**Упражнение 5** (12 баллов). Допустим, Вы являетесь владельцем небольшого магазина свадебных платьев. Ваши покупатели ведут себя случайным образом:

- с вероятностью 0.8 им ничего не нравится, и они покидают ваш магазин;
- с вероятностью 0.18 они примеряют хотя бы одно платье, но ничего не покупают;
- с оставшейся вероятностью они покупают платье, которое примеряли.

Что можно сказать о вероятности продать от 38 до 62 свадебных платьев за месяц, если Ваш магазин посетит 2 500 (стохастически независимых) женщин? Предполагается, что более одного платья никто не покупает. Решите задачу всеми изученными методами (неравенство Чебышева, ЦПТ). Сравните полученные результаты.

**Упражнение 6** (12 баллов). Последнее время автобус, на котором я добираюсь до университета, ходит не так регулярно, как раньше. И теперь я в среднем 1 раз из 3-х вынужден садиться в маршрутку и платить 25 руб. С какой вероятностью тогда мне хватит на месяц 250 руб., если проездной на автобус мне покупают родители, а ездить приходится 25 раз? Решите задачу с помощью ЦПТ.

**Упражнение 7** (20 баллов). В городе за год рождается 20 000 детей и считается, что вероятность рождения мальчика  $p = 0.51$ . В этом случае существует такое число  $d$ , что среди рожденных за год детей разница числа мальчиков и числа девочек будет не больше  $d$  с вероятностью 0.99. Найдите примерное значение  $d$  с помощью ЦПТ.