

Инструкция к выполнению лабораторной работы №2

Лабораторная работа состоит из одного большого задания. Допускается выполнение лабораторной на любом языке программирования, но мы настоятельно рекомендуем использовать Python (или R). В ином случае, количество работы увеличится в разы! Дедлайн по первой лабораторной работе — 16 апреля (оформленные решения нужно прислать мне в телеграм или на почту до 23:59).

Задание. Пусть X_1, X_2, \dots, X_n — выборка из непрерывного распределения.

1. **Понять, из какого распределения нам дана эта выборка.** Для этого можно построить графики гистограммы и ЭФР и сравнить их с какими-нибудь графиками известных функций распределения. Для этого, например, пригодится вот такая [полезная страничка](#).

На этом этапе предполагается, что мы вообще не умеем строить никаких оценок. И тип распределения и его параметры нужно просто **угадать**. Известно, что все параметры распределения имеют не более одного знака после запятой.

После нахождения оценок применить критерий Колмогорова (см. [полезные ссылки](#)) для проверки гипотезы о том, что выборка взята из **угаданного распределения**.

2. **Построить ОММ (с заданной функцией) и ОМП неизвестных параметров угаданного распределения.** Можно сделать это руками. Можно воспользоваться встроенными функциями Python. Можно найти их в литературе и использовать.
3. **Графически исследовать оценки на состоятельность и асимптотическую нормальность.** Сделать это в духе первой лабораторной работы.

(а) **Состоятельность:** интересуется график, демонстрирующий, сходится ли наша оценка к параметру с ростом объёма выборки. Для этого предлагается последовательно «отрезать» от нашей выборки начальные участки, вычислять по этому участку оценку и отмечать на плоскости точки с координатами (объём начального участка, построенная по нему оценка), соединяя каждую с предыдущей прямой линией. Провести прямые, соответствующие угаданным параметрам.

(б) **Асимптотическая нормальность.** Предлагается разбить данную выборку объёма 10^6 на 10^3 выборок объёма 10^3 . Для каждой из них вычислить значение статистики $\sqrt{n}(\theta^* - \theta)$ и построить гистограмму и ЭФР для получившихся значений статистики. Сравнить их с плотностью и функцией распределения стандартного нормального закона, соответственно. Оценить дисперсию получившейся выборки из значений статистики. Совпадает ли она с коэффициентом асимптотической нормальности?

4. **Построить асимптотические доверительные интервалы уровней доверия 0.95 и 0.99.** Рекомендуется использовать теорему о построении асимптотических доверительных интервалов с помощью асимптотически нормальных оценок (нам ведь по крайней мере одна такая оценка известна?).
-

Полезные ссылки:

- [Вся необходимая теория.](#)

- [SciPy documentation](#).
- [Continuous distributions in SciPy](#).
- [Kolmogorov test in SciPy](#).
- [Bokeh documentation](#).
- [ЭФР в Python, пример](#).