

# Введение в машинное обучение для Java-разработчиков

## Практическое задание 4. Статистика

Дата выдачи: 16.11.2023

Дедлайн: 23:59MSK 30.11.2023

## Оценивание и штрафы

Оценка считается по формуле: число выполненных задач + 2

Задание выполняется самостоятельно. «Похожие» решения считаются плагиатом и все задействованные студенты (в том числе те, у кого списали) не могут получить за него больше 0 баллов, что автоматически ведет к несдаче курса. Если вы нашли решение какого-то из заданий (или его часть) в открытом источнике, необходимо указать ссылку на этот источник в комментариях.

## Формат сдачи

Задания сдаются путем форка основного репозитория, коммита решения в мастер-ветку вашего форка и оповещении преподавателя о выполнении ДЗ. 1 и 2 задачу оформите в виде PDF файла. 3 задачу в виде ноутбука.

## Задачи

1. Известно, что генеральная совокупность распределена нормально с известным среднеквадратическим отклонением 16. Найдите доверительный интервал для оценки математического ожидания с надежностью 0.95, если выборочное среднее равно 80, а размер выборки равен 256.

hint: какой параметр взять для оценки мат. ожидания нормального распределения?

2. Продавец утверждает, что средний вес пачки печенья составляет 200 грамм. Из партии извлечена выборка из 10 пачек. Вес каждой пачки составляет 202, 203, 199, 197, 195, 201, 200, 204, 194, 190 грамм соответственно. Известно, что их веса распределены нормально. Верно ли утверждение продавца, если учитывать, что уровень значимости равен 1%?

3. Напишите функцию на языке Python, реализующую критерий Стьюдента. Для проверки сгенерируйте две выборки из двух нормальных распределений с разными параметрами (100 элементов будет достаточно) и примените критерий Стьюдента. Далее сгенерируйте две выборки из одного нормального распределения, задайте уровень значимости 5% и посчитайте количество ошибок первого рода.  
hint: `scipy.stats.t`