

# Лабораторная работа №1

## Задание 1

Выберите непрерывное распределение, у которого существуют первые четыре момента, и экспериментально убедитесь в асимптотической нормальности выборочного среднего, выборочной дисперсии, выборочной квантили порядка 0.5 для данного распределения. Также эмпирически убедитесь в том, что  $nF(X_{(2)}) \rightarrow U_1 \sim \Gamma(2, 1)$  и  $n(1 - F(X_{(n-2)})) \rightarrow U_2 \sim \Gamma(3, 1)$ .

*Указание:* сгенерируйте достаточно большое количество выборок достаточно большого объема, для каждой сгенерированной выборки вычислите соответствующие статистики (функции от выборок), постройте гистограммы результатов для каждой статистики, для наглядности рядом с гистограммой можно нарисовать соответствующую плотность (пока это метод "на глаз", но в дальнейшем мы разберем статистическую процедуру, позволяющую проверить согласованность распределения выборки с заданным вероятностным законом), также можно помимо гистограммы вывести мат. ожидание, дисперсию (или стандартное отклонение) и медиану.

## Задание 2

Второе задание представлено в пяти вариантах.

1. В файле <https://drive.google.com/file/d/1CSCheMzjberRwgcf90BBu-J6uxMg-Qf7/> представлены данные о параметрах различных экземплярах цветка ириса. Какой вид в датасете представлен больше всего, какой – меньше? Рассчитайте выборочное среднее, выборочную дисперсию, выборочную медиану и выборочную квантиль порядка 3/5 для суммарной площади чашелистика и лепестка (точнее оценки площадей) всей совокупности и отдельно для каждого вида. Построить график эмпирической функции распределения, гистограмму, box-plot и violinplot суммарной площади чашелистика и лепестка для всей совокупности и каждого вида.
2. В файле <https://drive.google.com/file/d/1gzPRqj7gZetjsipo3xpogYGL76enZDNO/> приведены данные (пол, ИМТ, курит/не курит) о более 1000 испытуемых. Какая категория граждан наиболее представлена в датасете: курящие мужчины, некурящие мужчины, курящие женщины, некурящие женщины? Рассчитайте выборочное среднее, выборочную дисперсию, выборочную медиану и выборочную квантиль порядка 7/10 ИМТ всех наблюдателей и отдельно для каждой возможной комбинации пол-курение. Построить график эмпирической функции распределения, гистограмму, box-plot и violinplot ИМТ для всех наблюдателей и отдельно для каждой возможной комбинации пол-курение.
3. В файле <https://drive.google.com/file/d/1vv2jGNp6EO8HHRoscDRQU90faR3j8iTN/> представлены данные об автомобилях. Какие типы автомобилей представлены в датасете? Какой тип наиболее распространен, какой – менее? Рассчитайте выборочное среднее, выборочную дисперсию, выборочную медиану и межквартильный размах для разности между максимальной и минимальной ценой для всей совокупности автомобилей и отдельно для каждого типа автомобиля. Построить график эмпирической функции распределения, гистограмму, box-plot и violinplot для этого же показателя для всей совокупности и отдельно для каждого типа авто.
4. В файле [https://drive.google.com/file/d/1O4rFr9xg9aFmkjx4-hl\\_XOc5O9q65\\_EW/](https://drive.google.com/file/d/1O4rFr9xg9aFmkjx4-hl_XOc5O9q65_EW/) приведены данные о мобильных телефонах. В сколько моделей можно вставить 2 сим-карты, сколько поддерживают 3-G, каково наибольшее число ядер у процессора? Рассчитайте выборочное среднее, выборочную дисперсию, выборочную медиану и выборочную квантиль порядка 3/5, построить график эмпирической функции распределения, гистограмму, box-plot и violinplot для емкости аккумулятора для всей совокупности и в отдельности для поддерживающих/не поддерживающих Wi-Fi.
5. В файле <https://drive.google.com/file/d/1EOoNW8I1ZH0KQ6lwuv2TZH0pNfBHVIZ/> приведены данные о шахматных партиях, сыгранных на одном шахматном сайте. Каких матчей сыграно больше: рейтинговых или обычных? Сколько матчей закончились быстрее чем за 20 ходов? Рассчитайте выборочное среднее, дисперсию, выборочную медиану и IQR для количества ходов в партии, построить график эмпирической функции распределения, гистограмму, box-plot и violinplot для количества ходов партии для всего датасета и в отдельности для рейтинговых/нерейтинговых игр.