1. Сгенерируйте выборку из нормального распределения $\mathcal{N}(\theta, 1)$ для $\theta = 0$, записав ее в виде матрицы $\{X_{ij}\}_{i=1...N, j=1...K}$, где $N = 1000$ и $K = 10000$.

   Выполните следующую процедуру для всех $n = 1...N$. Для всех $j = 1...K$ по выборке $\{X_{ij}\}_{i=1...n}$ оцените параметр θ с помощью оценки максимального правдоподобия $\widehat{\theta}_j^n$ и оценки $\widetilde{\theta}_j^n = \frac{1}{2}(\min_i X_{ij} + \max_i X_{ij})$. По выборкам $\{\widehat{\theta}_j^n\}_{j=1...K}$ и $\{\widetilde{\theta}_j^n\}_{j=1...K}$ для этих двух типов оценок найдите оценки дисперсий по методу максимального правдоподобия $\widehat{\sigma}^n$ и $\widetilde{\sigma}^n$.

   Нарисуйте график оценок дисперсий в зависимости от n. При необходимости не забывайте выставлять значение `plt.ylim`. Какая из оценок получилась лучше и почему?

   Постарайтесь решить эту задачу без циклов, используя только функции библиотеки `numpy`.

2. Предположим, что зависимость имеет вид $Y = \theta_0 X_0 + \theta_1 X_1 + \theta_2 X_2 + \varepsilon$, где $X_0 = 1$. Скачайте данные с диска, получив соответствующую выборку $(X^i, Y^i)_{i=1...N}$, где $X^i = (X_0^i, X_1^i, X_2^i)$. Оценкой параметра θ по методу наименьших квадратов называется величина $\widehat{\theta} = (X^T X)^{-1} X^T Y$. Посчитайте эту оценку для выданных данных и значение $\det(X^T X)$. Сделайте вывод.

   Рассмотрим оценку $\widehat{\theta}(\lambda) = (X^T X + \lambda I)^{-1} X^T Y$. Возьмите сетку значений λ от 0 до 100 с шагом 0.1 (нам: числа подобрать). Для каждого значения λ посчитать оценки $\{\widehat{\theta}_i(\lambda)\}_{i=1..N}$, где оценка $\widehat{\theta}_i(\lambda)$ получена по приведенной выше формуле, где из данных исключена пара (X_i, Y_i).

   Для каждого значения λ посчитайте оценку дисперсии $\widehat{\sigma}(\lambda)$ по выборке $\{\widehat{\theta}_i(\lambda)\}_{i=1..N}$. Постройте график $\widehat{\sigma}(\lambda)$ в зависимости от λ. По графику выберите оптимальное значение λ.

   Постройте график доверительных интервалов для $\{\widehat{\theta}_i(\lambda)\}$ в зависимости от λ. Фактически для этого нужно построить графики верхней и нижней границ интервалов. Для наглядности можно закрасить область между этими границами. Сделайте выводы.

   Напечатать лучшее значение λ, для которого так же напечатать значения оценки дисперсий и доверительных интервалов.

3. Целый семестр вы решали практические задачи. Теперь пришло время вам проставить оценки самим себе за практику в течении семестра. Эта оценка не будет являться вашей итоговой оценкой за практику. Скачайте таблицу с результатами по практике. В удобном формате ее можно скачать на диске (файл marks.csv).

   Поскольку еще не все задачи проверены, оцените, сколько баллов вы получите за непроверенные задачи на основе данных по вашим проверенным работам и по проверенным работам остальных по этой задаче следующим образом. Пусть у вас проверены n задач, за которые вы получили баллы $X_1, ..., X_n$, где $\mathsf{P}(X_i = j) = \theta_j$. Оценим $\theta = (\theta_1, ..., \theta_k)$ с помощью байесовской оценкой с априорным распределением, в качестве которого возмем сопряженное в данной модели распределение Дирихле. Вектор θ, определенный на симплексе, имеет распределение Дирихле $Dir(\alpha_1, ..., \alpha_k)$, если его плотность имеет вид

$$q(t) = \frac{\Gamma\left(\sum_j \alpha_j\right)}{\prod_{j=1}^k \Gamma(\alpha_j)} \prod_{j=1}^k t_j^{\alpha_j - 1},$$

причем $\mathsf{E}\theta_j = \frac{\alpha_j}{\alpha_1+...+\alpha_k}$.

Апостериорным распределением будет $Dir\left(\alpha_1 + \sum_i I\{X_i = 1\}, ..., \alpha_k + \sum_i I\{X_i = k\}\right)$. Для получения оценки балла X за непроверенную задачу оцените величину $\mathsf{E}X = \sum_j j\,\mathsf{P}(X = j)$. В качестве параметра α_j априорного распределения возьмите частоту появления балла j среди проверенных работ по этой задаче. Описанным выше способом оцените баллы за непроверенные задачи у себя, своего лучшего друга, врага, а так же самой(го) красивой(го) девушки(парня). При написании кода удобно пользоваться `collections.Counter`.

По задачам в семестре можно получить от 0 до 10 баллов. Для каждого студента X определим метку $T \in \{2, 3, 4, 5\}$ в соответствии с правилами выставления оценок в МФТИ по 5-балльной системе. Считается, что 2,9999 это неуд. Поставьте оценки всем студентам. Сколько вы получаете? Для каждой оценки напечатайте количество человек, которые получают эту оценку.

Посчитайте, как решает задачи среднестатистический студент, который получает метку t, то есть $\mathsf{E}(X|T = t)$. Посчитайте, как решает задачи среднестатистический студент, который не получает метку $t = 2$, то есть $\mathsf{E}(X|T \neq 2)$. При подсчете считайте, что все студенты равновероятны. Постройте график 20-мерной плотности. Сделайте выводы.