

Теория 1. Сопряжённые распределения и экспоненциальный класс распределений

Курс: Байесовские методы в машинном обучении, 2020

1. Пусть x_1, x_2, \dots, x_N – независимая выборка из непрерывного равномерного распределения $U[0, \theta]$. Требуется найти оценку максимального правдоподобия θ_{ML} , подобрать сопряжённое распределение $p(\theta)$, найти апостериорное распределение $p(\theta|x_1, \dots, x_N)$ и вычислить его статистики: мат.ожидание, медиану и моду. Формулы для статистик нужно вывести, а не взять готовые. *Подсказка: задействовать распределение Парето.*
2. Предположим, что вы приезжаете в новый город и видите автобус с номером 100. Требуется с помощью байесовского подхода оценить общее количество автобусных маршрутов в городе. Каким априорным распределением стоит воспользоваться (обоснуйте выбор его параметров)? Какая из статистик апостериорного распределения будет наиболее адекватной (обоснуйте свой выбор)? Как изменятся оценки на количество автобусных маршрутов при последующем наблюдении автобусов с номерами 50 и 150? *Подсказка: воспользоваться результатами предыдущей задачи. При этом обдумать как применить непрерывное распределение к дискретным автобусам.*
3. Записать распределение Парето с плотностью $\text{Pareto}(x|a, b) = \frac{ba^b}{x^{b+1}} [x \geq a]$ при фиксированном a в форме экспоненциального класса распределений. Найти $E \log x$ путём дифференцирования нормировочной константы.