# Теория 1. Сопряженные распределения и экспоненциальный класс распределений

**Дарья Дятлова**

## Задание 1

$x1, x2, ..., x_N$ – независимая выборка из непрерывного равномерного распределения.

Задача:

- найти оценку максимального правдоподобия $\theta_{ML}$;
- подобрать сопряженное распределение $p(\theta)$;
- найти апостериорное распределение $p(\theta|x1, x2, ..., x_N)$;
- вычислить математическое ожидание, медиану и моду апостериорного распределения.

Решение:

- $L(\theta) = p(X|\theta) = \prod_{i=1}^{N} p(x_i|\theta)$.
  $\theta_{ML} = argmax_\theta L(\theta) = \prod_{i=1}^{N} p(x_i|\Theta) = argmax_\theta \sum_i^N logp(x_i|\theta) = argmax_\theta - Nlog\theta$.
  $\frac{\partial - Nlog\theta}{\partial\theta} = -\frac{N}{\theta}$.
  Функция убывающая, значит, нам нужно выбрать минимальное значение, которое может принимать $\theta$, т.к. $\theta \geq x_i$: $\theta_{ML} = x_N$.
- $p(\theta|a, b) = \frac{a \cdot b^a}{\theta^{a+1}} \cdot \theta \geq b$;
- $p(\theta|X, a, b) = \frac{p(X|\theta)p(\theta|a,b)}{\int p(x|\theta) \cdot p(\theta|a,b) d(\theta)} = \frac{1}{C} \cdot \frac{a \cdot b^a}{\theta^{a+N+1}} \cdot \theta \geq max_{i=1:N}(b, x_i)$.
  $C = \int_0^\infty \frac{a \cdot b^a}{\theta^{a+N+1}} \cdot \theta d(\theta)$, где $(\theta \geq max_{i=1:N}(b, x_i)) = -\frac{a \cdot b^a}{\theta^{a+N+1}} \Big|_{max_{i=1...N}(b,x_i)}^{\inf} = \frac{a \cdot b^a}{(max_{i=1:N}(b,x_i))^{a+N} \cdot (a+N)}$.
  $p(\theta|X, a, b) = \frac{(max_{i=1:N}(b,x_i))^{a+N} \cdot (a+N)}{\theta^{a+N+1}} = Pareto(\theta|a', b')$, где $a' = a + N, b' = max_{i=1:N}(b, x_i)$.
- $E\mu = \int_b^c = \frac{a \cdot b^a}{x^{a+1}} \cdot xdx = \int_b^{\inf} \frac{a \cdot b^a}{x^b} dx = \frac{a \cdot b^a}{(a-1) \cdot b^{a-1}} = \frac{a \cdot b}{a-1}$.

  $Med\mu = c$
  $\int_b^c \frac{a \cdot b^a}{x^{a+1}} dx = \frac{1}{2}$
  $1 - \frac{b^a}{c^a} = \frac{1}{2}$
  $Med\mu = 2^{a^{-1}} \cdot b$.

- $Mod\mu = b$.

## Задание 2

Дано: наблюдение автобуса с номером 100.
Задача:

- оценить общее количество автобустных маршрутов в городе;
- обосновать выбор распределения и его параметров;
- обосновать выбор статистики апостериорного распределения;
- оценить, как изменяться оценки на количество маршрутов при наблюдении автобусов с номерами 50 и 150.

Решение:

- будем оценивать случайную величину $\mu$, используя непрерывное равномерное распределение. В качестве номера автобуса будем брать округленное сверху вещественное значение, которое может принимать наша случайная величина. Задача сводится к оценке параметра $\theta(x \sim U[0, \theta])$;

- $p(\theta|a,b) = \frac{a \cdot b^a}{\theta^{a+1}} \cdot \theta, \theta \geq b$, распределение Парето с параметрами $a, b$. Параметр $b$ можно выбрать на основе нашего жизненного опыта о количестве маршрутов в городе;

- в качестве достаточной статистики возьмем математическое ожидание или медиану, не стоит брать моду, так как это просто значение параметра $b$ – мы не учитываем знание о параметре $a$. Так как математическое ожидание звучит привычнее, я бы остановила свой выбор на нем;

- после наблюдения автобуса с номером 50 параметр $b$ не изменится, но наблюдение автобусов с номером $> 100$, а значит и большее количество автобусных маршрутов в городе (предполагаем, что максимальный номер автобуса $==$ количество автобусных маршрутов в городе). При наблюдении автобуса с номером 150 распределение сдвинется, параметр $b$ может стать равным 150, если был меньше, а значения $\theta \geq 150$ станут более вероятными.

# Задание 3

Дано: распределение Парето с плотностью:$\frac{b \cdot a^b}{x^{b+1}}[x \geq a]$.

Задача:

- записать распределение при фиксированном $a$ в форме экспоненциального класса распределений;

- найти $Elogx$ путем дифференцирования нормировочной константы.

Решение:

- $p(x|\theta) = \frac{f(x)}{g(\theta)} \cdot e^{\theta^{T u(x)}}$
  $p(x|a,b) = \frac{a \cdot b^a}{x} \cdot e^{-alnx}[x \geq b]$, $\theta \Longrightarrow f(x) = x \geq bx; g(\theta = a) = \frac{1}{a \cdot b^a}; u(x) = -ln(x)$.

- $E(logx) = \frac{\partial g(a)}{\partial a} = -\frac{1}{a^2 \cdot b^{2a}}(b^a + 2b^2 + lnb) = -\frac{1+a \cdot lnb}{a^2 \cdot b}$.