Домашняя работа

Котов Артем, МОиАД ВШЭ СПб

26 октября 2021 г.

Содержание

Task 1

Условие: $X = (x_1, \ldots, x_n) \sim U[0, \theta]$ — независимые. Найти θ_{ML}, $p^\dagger(\theta)$, $\mathbb{E}\theta$, медиану и моду апостериорного $p(\theta|X)$.

Max. likelihood

Решение.

$$L(\theta) = p(X, \theta) = [\text{независимость}] = \prod_{i=1}^{n} p(x_i|\theta)$$

$$\theta_{ML} = \underset{\theta}{\operatorname{argmax}}\, L(\theta) = \cdots = \underset{\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \log p(x_i|\theta)$$

Из условия мы знаем, что $p(x_i|\theta) = \frac{1}{\theta}$, тогда

$$\theta_{ML} = \underset{\theta}{\operatorname{argmax}}(-n \log \theta)$$

Видно, что в аргументе стоит убывающая функция, следовательно, θ для максимизирования правдоподобия должна быть наименьшей из возможных. Так как мы пронаблюдали какие-то значения X, то наименьшей из возможных будет $\max X = x_{(n)}$. ■

Сопряженное к равномерному

Решение.

Рассмотрим $p(\theta|\alpha, \beta) = \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}[\beta \leqslant \theta]$ — распределение Парето. Покажем, что апостериорное так же будет иметь такой же вид, но с другими $\tilde{\alpha}, \tilde{\beta}$.

Рассмотим $p(\theta|X) = \frac{p(X|\theta)p(\theta|\alpha,\beta)}{\int\limits_0^{+\infty} p(X|\theta)p(\theta|\alpha,\beta)d\theta}$.

Сначала разберемся с нормировочным интегралом:

$$\int\limits_0^{+\infty} p(X|\theta)p(\theta|\alpha, \beta)d\theta = (*)$$

Здесь возникнет произведение двух индикаторных функций: $[\beta \leqslant \theta][x_{(n)} \leqslant \theta]$, что можно переписать, как $[m \leqslant \theta]$, где $m = \max(\beta, x_{(n)})$, тогда

$$(*) = \alpha\beta^\alpha \int_m^\infty \theta^{-n-\alpha-1} d\theta = (**)$$

Если $-n - \alpha - 1 \geqslant -1$, то интеграл расходится, т.е. необходимо, чтобы $n + \alpha < 0$

$$(**) = \frac{\alpha\beta^\alpha}{-n-\alpha} \theta^{-n-\alpha}\Big|_m^{+\infty} = \frac{\alpha\beta^\alpha}{(n+\alpha)m^{n+\alpha}}$$

Таким образом,
$$p(\theta|X) = (n+\alpha)m^{n+\alpha}\theta^{-n-\alpha-1}[m \leqslant \theta] = (***),$$
где $m = \max(\beta, x_{(n)})$.

Если $\tilde{\alpha} = \alpha + n$, а $\tilde{\beta} = m$, то
$$(***) = \frac{\tilde{\alpha}\tilde{\beta}^{\tilde{\alpha}}}{\theta^{\tilde{\alpha}+1}}[\tilde{\beta} \leqslant \theta],$$
что есть распределение Парето, т.е. оно действительно является сопряженным к равномерному. ■

Статистики

Мат. ожидание

Решение.

$$\mu = \mathbb{E}\theta = \int_0^{+\infty} \theta p(\theta|X)d\theta = (n+\alpha)m^{n+\alpha}\int_m^{+\infty}\theta^{-n-\alpha}d\theta$$
$$= [\tilde{\alpha} = n+\alpha] = \tilde{\alpha}m^{\tilde{\alpha}}\int_m^{+\infty}\theta^{-\tilde{\alpha}}d\theta = \tilde{\alpha}m^{\tilde{\alpha}}\frac{1}{\tilde{\alpha}-1}m^{-\tilde{\alpha}+1} = \frac{\tilde{\alpha}m}{\tilde{\alpha}-1}$$

■

Медиана

Решение.

c — медиана, причем $c \geqslant m$
$$P(c \leqslant \theta < +\infty) = (n+\alpha)m^{n+\alpha}\int_c^{+\infty}\theta^{-n-\alpha-1}d\theta = \left(\frac{m}{c}\right)^{n+\alpha} = \frac{1}{2}$$
$$\Downarrow$$
$$(n+\alpha)(\ln m - \ln c) = -\ln 2 \Longrightarrow c = \exp(\ln m + \frac{\ln 2}{n+\alpha}) = m2^{\frac{1}{n+\alpha}}$$

■

Мода

Решение.

$$\operatorname*{argmax}_\theta p(\theta|X) = \operatorname*{argmax}_\theta \theta^{-n-\alpha-1}[m \leqslant \theta]$$

Тут также возникает убывающая функция от θ, следовательно, берем наименьшее доступное, т.е. $\theta = m$, где $m = \max(\beta, x_{(n)})$ ■