1. Найдите оценки по методу моментов для следующих параметрических распределений (каждый раз нам задана выборка $X_1, \dots, X_n$):

$a)$ $Pois(\theta), \theta > 0$;

$b)$ $\Gamma(\theta, \theta), \theta > 0$. Здесь подразумевается распределение $\Gamma(\alpha, \beta)$ с плотностью

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\beta x} \cdot \mathbb{I}_{\{x>0\}};$$

$c)$ Можно ли построить оценку по методу моментов с помощью одной из пробных функций $y, y^2, y^3, \dots$ для параметра сдвига θ распределения Коши

$$p(x) = \frac{1}{\pi(1 + (x - \theta)^2)}, \qquad x \in \mathbb{R}$$

Решение.

$$a) \ \mathbb{E}_\theta X_1 = \int_0^{+\infty} x \cdot \frac{e^{-\theta} \cdot \theta^x}{x!} dx = \theta, \qquad \mathbb{E}_\theta X_1^2 = \int_0^{+\infty} x^2 \cdot \frac{e^{-\theta} \cdot \theta^x}{x!} dx = \theta^2 + \theta - \text{второй момент не понадобился}$$

$$\boxed{\hat{\theta} = \bar{X}}$$

$$b) \ \mathbb{E}_\theta X_1 = \frac{\alpha}{\beta} = \frac{\theta}{\theta} = 1, \qquad \mathbb{E}_\theta X_1^2 = \frac{\alpha + \alpha^2}{\beta^2} = \frac{\theta + \theta^2}{\theta^2}$$

$$\bar{X} = \frac{\theta + \theta^2}{\theta^2} \quad \Rightarrow \quad \frac{\theta + 1}{\theta} = \bar{X} \quad \Rightarrow \quad \boxed{\hat{\theta} = \frac{1}{\bar{X} - 1}}$$

$c)$ Так как интеграл Лебега $\int_{-\infty}^{+\infty} x^\alpha f_X(x) dx$ не определен для $\alpha \geq 1$, то математическое ожидание (моменты любых порядков) этого распределения вырождено (не определено), а значит и метод моментов нам ничего не даст.

2. Пусть $X_1, \dots, X_n$ — выборка из распределения Лапласа со сдвигом $\theta \in \mathbb{R}$, т. е. плотность имеет вид

$$p_\theta(x) = \frac{1}{2} e^{-|x - \theta|}, \qquad x \in \mathbb{R}.$$

Найдите оценку максимального правдоподобия для параметра θ.

Решение.

Найдем функцию правдоподобия:

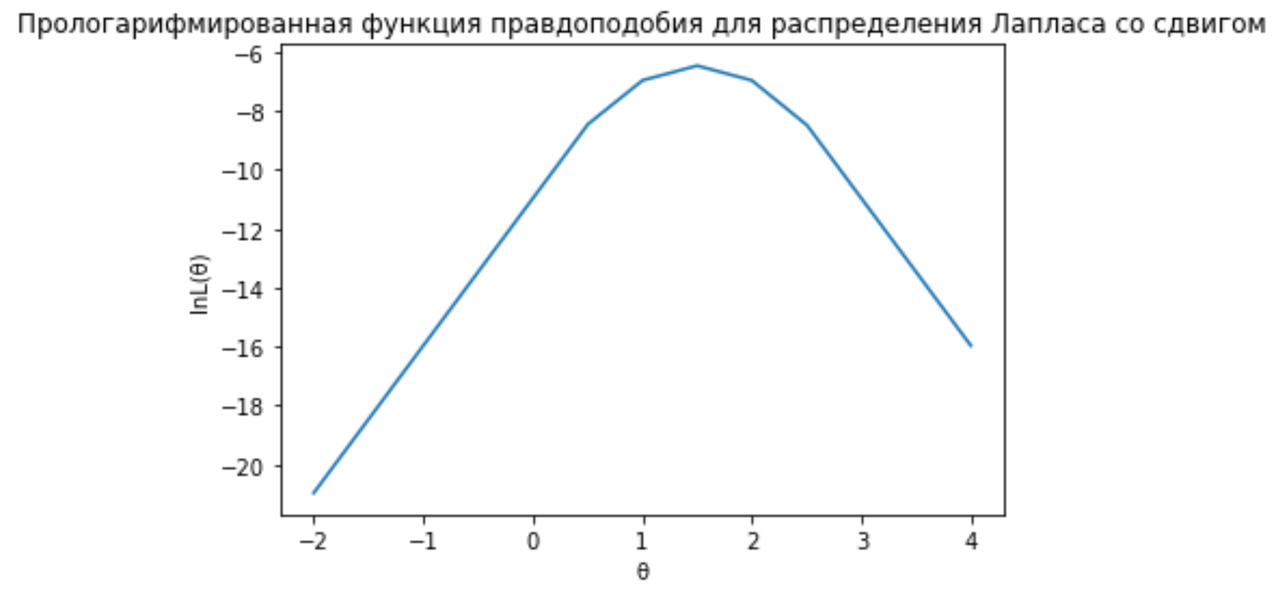
$$L_\theta(X) = \prod_{i=1}^n p_\theta(X_i) = \left(\frac{1}{2}\right)^n \cdot e^{-\sum_{i=1}^n |X_i - \theta|} \quad \Rightarrow \quad \ln L = -n \cdot \ln 2 - \sum_{i=1}^n |X_i - \theta| = -n \cdot \ln 2 - |\theta - \overline{X_1}| - \dots - |\theta - \overline{X_n}|, \qquad \overline{X_1} \leq \overline{X_2} \leq \dots \leq \overline{X_n}$$

Тогда

При четном размере выборки $n = 2k$: максимум достигается при любом $\theta \in [\bar{X}_k, \bar{X}_{k+1}]$ ...

При нечетном размере выборки $n = 2k - 1$: максимум достигается при $\theta = \text{median}(X) = \bar{X}_k$ ...

Медиана является оценкой максимального правдоподобия в распределении Лапласа со сдвигом θ, так как это распределение является симметричным и имеет вырожденный момент. Медианой мы оцениваем центральное значение распределения, когда когда мат ожидание не может быть вычислено.



3. $X_1, \dots, X_n$ — выборка из дискретного распределения

$$P_\theta(X_1 = 0) = \theta, \qquad P_\theta(X_1 = 1) = \theta, \qquad P_\theta(X_1 = 2) = 1 - 2\theta, \qquad \theta \in \left(0, \frac{1}{2}\right).$$

Найдите оценку параметра θ по методу максимального правдоподобия. Проверьте полученную оценку на несмещенность и состоятельность.

Решение.

Найдем функцию правдоподобия:

$$L_\theta(X) = \prod_{i=1}^n P_\theta(X_i) = P_\theta(X_1) \cdot P_\theta(X_2) \cdot \ldots \cdot P_\theta(X_n)$$

так как значения $X_i = \{0,1,2\}$, то функция правдоподобия может иметь сдедующий вид:

$$L_\theta(X) = \prod_{i=1}^n (\theta^{n_0} \cdot \theta^{n_1} \cdot (1 - 2\theta)^{n_2})$$, где n_0, n_1, n_2 — количество результатов X_i, равных 0, 1, 2 соответственно, $\quad n = n_1 + n_2 + n_3$.

Значит функция логарифмического правдоподобия:

$$\ln\left(L_\theta(X)\right) = \ln\left(\prod_{i=1}^n (\theta^{n_0} \cdot \theta^{n_1} \cdot (1 - 2\theta)^{n_2})\right) = n_0 \ln \theta + n_1 \ln \theta + n_2 \ln(1 - 2\theta)$$

$$\frac{\partial}{\partial \theta}\left(\ln\left(L_\theta(X)\right)\right) = \frac{n_1 + n_2}{\theta} - \frac{2n_3}{1 - 2\theta} \quad \Rightarrow \quad \frac{n_1 + n_2}{\theta} - \frac{2n_3}{1 - 2\theta} = 0 \quad \Rightarrow \quad \hat{\theta} = \frac{n_0 + n_1}{2(n_1 + n_2 + n_3)} = \frac{n_0 + n_1}{2n} = \frac{\sum_{i=1}^n (I[X_i = 0] + I[X_i = 1])}{2n}$$

Проверим на несмещенность:

$$\mathbb{E}\hat{\theta} = \mathbb{E}\left(\frac{\sum_{i=1}^n (I[X_i = 0] + I[X_i = 1])}{2n}\right) = \left(\frac{\sum_{i=1}^n \mathbb{E}(I[X_i = 0] + I[X_i = 1])}{2n}\right) = \frac{\sum_{i=1}^n (\theta + \theta)}{2n} = \frac{2n\theta}{2n} = \theta - \text{оценка является несмещенной.}$$

Проверим на состоятельность:

$$\mathbb{D}\hat{\theta} = \mathbb{D}\left(\frac{\sum_{i=1}^n (I[X_i = 0] + I[X_i = 1])}{2n}\right) = \frac{1}{4n^2} \cdot \mathbb{D}\left(\sum_{i=1}^n (I[X_i = 0] + I[X_i = 1])\right) = \frac{1}{4n^2} \cdot \left(\sum_{i=1}^n \mathbb{D}(I[X_i = 0] + I[X_i = 1])\right) =$$

$$= \frac{1}{4n^2} \cdot \left(\sum_{i=1}^n \theta(1 - \theta)\right) = \frac{\theta(1 - \theta)}{4n} \xrightarrow{n \to +\infty} 0 - \text{оценка является состоятельной.}$$