Метод максимальної вірогідності

1 Теоретичні відомості

Функція вірогідності кратної вибірки

Для простоти розглянемо (S, Σ) – такий вибірковий простір, де $S \subset \mathbb{R}^n$ та $\Sigma = \mathcal{B}(S)$ – борелева σ-алгебра підмножин S. Відповідно вибірка $X(\omega) : \Omega \to S$ є випадковим вектором на S.

Розглянемо кратну вибірку $X = (X_1, \ldots, X_n) \in S$ з розподілом спостережень $F(t; \theta) = \mathbf{P}_\theta (X_1 < t)$, $\theta \in \Theta \subset \mathbb{R}^d$. Отже сумісний розподіл X має вигляд:

$$F(x_1, \ldots, x_n; \theta) = \mathbf{P}_\theta (X_1 < x_1, \ldots, X_n < x_n) = \prod_{j=1}^{n} \mathbf{P}_\theta (X_1 < x_j) = \prod_{j=1}^{n} F(x_j), x = (x_1, \ldots, x_n)^T \in S$$

Припустимо, що існує така σ-скінченна міра μ на просторі значень вибірки X, відносно якої міра, породжена функцією розподілу $F(x_1, \ldots, x_n; \theta)$, має щільність $f(x_1, \ldots, x_n; \theta)$ (похідна Радона-Никодима якщо формально).

Наприклад,

1. Якщо розподіл X є дискретним, то μ – деяка точкова міра. Тоді
$$f(x_1, \ldots, x_n; \theta) = \mathbf{P}_\theta (X_1 = x_1, \ldots, X_n = x_n)$$
2. Якщо розподіл X є абсолютно неперервним, то μ – міра Лебега. Тоді
$$f(x_1, \ldots, x_n; \theta) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F(x_1, \ldots, x_n; \theta)$$

Функцією вірогідності вибірки X називають сумісну щільність її розподілу, тобто

$$L(x_1, \ldots, x_n; \theta) = f(x_1, \ldots, x_n; \theta).$$

Для спрощення $L(x; \theta) = L(x_1, \ldots, x_n; \theta)$, $x = (x_1, \ldots, x_n)^T \in S$. Грубо та простими словами, $L(x; \theta)$ по x є щільністю розподілу X.

Емпіричною функцією вірогідності вибірки X називають значення функції вірогідності для заданої вибірки, тобто

$$L(X; \theta) = L(x; \theta)\Big|_{x=X} = L(X_1, \ldots, X_n; \theta)$$

Якщо X – кратна вибірка, то спостереження X_j незалежні в сукупності та мають однаковий розподіл $F(x; \theta) = \mathbf{P}_\theta (X_1 < x)$. Відповідно функція вірогідності $L(X, \theta)$ є добутком щільності розподілу спостереження $f(x; \theta)$:

$$L(x_1, \ldots, x_n, \theta) = \prod_{j=1}^{n} f(x_j; \theta)$$

Приклади побудови функції вірогідності дивись у записах до попередніх практичних занять.

Метод максимальної вірогідності

Розглянемо наступний підхід до побудови оцінки невідомого параметра $\theta \in \Theta$, що базується на використанні функції вірогідності $L(x, \theta)$.

Оцінкою методом максимальної вірогідності (далі ОМВ) називають точку максимума $L(x, \theta)$ на Θ із підстановкою спостережуваних значень вибірки, тобто:

$$\hat{\theta}_n^{MLE} = \theta^*(X), \text{ де } L(x, \theta^*(x)) = \max_{\theta \in \Theta} L(x, \theta),\ x \in S. \tag{1}$$

За виконання досить широких умов (як-от консистентність, асимптотична нормальність, оптимальність) ОМВ є хорошою оцінкою параметра, що оцінюється. Інколи буває простіше доводити властивості безпосередньо працюючи з формою оцінки, а не перевіряючи умови.

Інколи знаходження максимума може ускладнюватися внаслідок форми $L(x, \theta)$. Для її спрощення можна спробувати взяти логарифм та знаходити точки максимума власне від $\ln L(x, \theta)$ (завдяки тому, що $\ln(x)$ строго зростаюча функція).

Якщо $L(x, \theta)$ є неперервно диференційовною по θ функцією, можна відшукати екстремуми з рівняння

$$\frac{\partial}{\partial \theta} L(x, \theta) = 0 \text{ (або } \operatorname{grad}(L(x, \theta)) = 0 \text{ у разі векторного } \theta),$$

а потім перевірити, чи існує серед знайдених екстремумів точка максимума. Та сама логіка в принципі, якщо досліджувати не $L(x, \theta)$, а $\ln L(x, \theta)$.

Якщо є ОМВ для невідомого параметра, то функція від неї буде ОМВ для функції від невідомого параметра, якщо задана функція є бієкцією:

Теорема (Про інваріантність ОМВ). Нехай $\hat{\theta}$ – ОМВ невідомого параметра θ та розглянемо $q : \Theta \to Q$ – взаємно-однозначне відображення. Тоді $q(\hat{\theta})$ – ОМВ для $\tau(\theta)$.

2 Задачі

2.1 Задача 1

Нехай $X = (X_1, \ldots, X_n)$ – кратна вибірка з розподілом спостережень $X_1 \sim \mathrm{Bern}(\theta)$, $\theta \in (0,1)$.

$$\mathbf{P}_\theta\left(X_1 = k\right) = \theta^k (1-\theta)^{1-k},\ k = 0, 1.$$

Побудувати ОМВ для θ. Дослідити оцінку на консистентність.

Розв'язання. Запишемо функцію вірогідності на $S = \{0,1\}^n$:

$$L(x,\theta) = \prod_{j=1}^{n} \mathbf{P}_\theta\left(X_j = x_j\right) = \theta^{\sum_{j=1}^n x_j} (1-\theta)^{n - \sum_{j=1}^n x_j},\ x = (x_1, \ldots, x_n) \in S$$

Візьмемо логарифм

$$\ln L(x,\theta) = \sum_{j=1}^{n} x_j \cdot \ln\theta + (n - \sum_{j=1}^{n} x_j) \cdot \ln(1-\theta)$$

Знаходимо точки екстремуму з рівняння

$$\frac{\partial}{\partial\theta} \ln L(x,\theta) = \frac{\sum_{j=1}^n x_j}{\theta} - \frac{n - \sum_{j=1}^n x_j}{1-\theta} = \frac{\sum_{j=1}^n x_j (1-\theta) - (n - \sum_{j=1}^n x_j)\theta}{\theta(1-\theta)} = \frac{-n\theta + \sum_{j=1}^n x_j}{\theta(1-\theta)} = 0$$

Розв'язуючи рівняння відносно θ отримаємо розв'язок $\theta_*(x) = \sum_{j=1}^n x_j / n$. Можна перевірити, що $\theta_*(x)$ є точкою максимума $\ln L(x,\theta)$ (функція вірогідності є увігнутою по θ, зростає при $\theta < \theta_*$ та спадає при $\theta > \theta_*$). Значить $\hat{\theta}_n = \theta_*(X) = \sum_{j=1}^n X_j / n$ є ОМВ параметра θ.

Оскільки $\mathbf{E}_\theta\left[X_1\right] = \theta < \infty$, то за критерієм Колмогорова про ПЗВЧ $\hat{\theta}_n \to^{P_1} \theta$, значить оцінка є строго консистентною оцінкою параметра θ.

2.2 Задача 2

Нехай $X = (X_1, \ldots, X_n)$ – кратна вибірка з розподілом спостережень $X_1 \sim \mathrm{Exp}(\theta)$, $\theta > 0$.

$$f(y;\theta) = \theta \exp(-\theta y),\ y > 0.$$

Побудувати ОМВ для θ. Дослідити оцінку на консистентність.

Розв'язання. Запишемо функцію вірогідності на $S = (0, \infty)^n$:

$$L(x,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = \theta^n \exp(-\theta \sum_{j=1}^{n} x_j),\ x = (x_1, \ldots, x_n) \in S$$

Візьмемо логарифм

$$\ln L(x,\theta) = n \ln \theta - \theta \sum_{j=1}^{n} x_j$$

Знаходимо точки екстремуму з рівняння

$$\frac{\partial}{\partial \theta} \ln L(x,\theta) = \frac{n}{\theta} - \sum_{j=1}^{n} x_j = 0$$

Розв'язуючи рівняння відносно θ отримаємо розв'язок $\theta_*(x) = (\sum_{j=1}^{n} x_j/n)^{-1}$. Можна перевірити, що $\theta_*(x)$ є точкою максимума $\ln L(x,\theta)$ (функція вірогідності є увігнутою по θ, зростає при $\theta < \theta_*$ та спадає при $\theta > \theta_*$), значить $\hat{\theta}_n = \theta_*(X) = (\sum_{j=1}^{n} X_j/n)^{-1}$ є ОМВ параметра θ.

Оскільки $\mathbf{E}_\theta\,[X_1] = \theta^{-1} < \infty$, то за критерієм Колмогорова про ПЗВЧ $\sum_{j=1}^{n} X_j/n \to^{P_1} \theta^{-1} > 0$. За теоремою про арифметичні дії над збіжними майже напевне послідовностями миттєво маємо, що $\hat{\theta}_n = (\sum_{j=1}^{n} X_j/n)^{-1} \to^{P_1} (\theta^{-1})^{-1} = \theta$, що і доводить консистентність оцінки.

2.3 Задача 3

Нехай $X = (X_1, \dots, X_n)$ – кратна вибірка з розподілом спостережень $X_1 \sim N(\mu, \sigma^2)$, де $\mu \in \mathbb{R}$ – відоме, але $\sigma^2 > 0$ є невідомим.

$$f(y; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(y-\mu)^2/(2\sigma^2)),\ y \in \mathbb{R}.$$

Побудувати ОМВ для σ^2. Дослідити оцінку на консистентність.

Розв'язання. Запишемо функцію вірогідності на $S = \mathbb{R}^n$:

$$L(x,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = (2\pi\sigma^2)^{-n/2} \exp\left(-(2\sigma^2)^{-1}\sum_{j=1}^{n}(x_j-\mu)^2\right),\ x = (x_1,\dots,x_n) \in S$$

Візьмемо логарифм

$$\ln L(x,\theta) = -(n/2)\ln(2\pi\sigma^2) - (2\sigma^2)^{-1}\sum_{j=1}^{n}(x_j-\mu)^2$$

Знаходимо точки екстремуму з рівняння

$$\frac{\partial}{\partial(\sigma^2)} \ln L(x,\sigma^2) = -(n/2)(\sigma^2)^{-1} + (1/2)(\sigma^2)^{-2}\sum_{j=1}^{n}(x_j-\mu)^2 = 0$$

Розв'язуючи рівняння відносно θ отримаємо розв'язок $\theta_*(x) = \sum_{j=1}^{n}(x_j-\mu)^2/n$. Можна перевірити, що $\theta_*(x)$ є точкою максимума $\ln L(x,\theta)$. Дійсно, перезапишемо логарифм функції вірогідності:

$$\ln L(x,\theta) = -(n/2)\ln(2\pi) - (n/2)\left(\ln(\sigma^2) + \frac{1}{\sigma^2}\left(\frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)^2\right)\right)$$

Максимізувати логарифм вірогідності – значить мінімізувати функцію $h(x) = \ln(x) + a/x$, $a := n^{-1}\sum_{j=1}^{n}(x_j-\mu)^2$ на $x \in (0,\infty)$. Перезапишемо $h(x)$:

$$h(x) = \ln(x) + a\exp(-\ln(x)),\ x > 0$$

Тобто можна спростити собі життя до мінімізації $g(y) = h(e^y) = y + a\exp(-y)$, де $y \in \mathbb{R}$. Можна переконатися, що функція $g(x)$ є опуклою гладкою функцією. Якщо знайдемо її екстремум, значить і знайдемо точку мінімума. Неважко переконатися, що це буде $y_* = \ln(a)$. З іншого боку, $y_* = \ln(x_*)$, тому $x_* = a$.

Значить $\hat{\theta}_n = \theta_*(X) = \sum_{j=1}^{n}(X_j-\mu)^2/n$ є ОМВ параметра σ^2.

Оскільки $\mathbf{E}_{\sigma^2}\left[(X_1-\mu)^2\right] = \mathbf{D}_{\sigma^2}\left[X_1\right] = \sigma^2 < \infty$, то за критерієм Колмогорова про ПЗВЧ $\hat{\theta}_n \to^{P_1} \sigma^2$, значить оцінка є строго консистентною оцінкою параметра σ^2.

2.4 Задача 4

Нехай $X = (X_1, \ldots, X_n)$ – кратна вибірка з розподілом спостережень $X_1 \sim N(\mu, \sigma^2)$, де $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, \infty)$ – невідомий параметр.

$$f(y; \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(y-\mu)^2/(2\sigma^2)),\ y \in \mathbb{R}.$$

Побудувати ОМВ для θ. Дослідити оцінку на консистентність.

Розв'язання. Запишемо функцію вірогідності на $S = \mathbb{R}^n$:

$$L(x, \theta) = \prod_{j=1}^{n} f(x_j; \theta) = (2\pi\sigma^2)^{-n/2} \exp\left(-(2\sigma^2)^{-1} \sum_{j=1}^{n} (x_j - \mu)^2\right),\ x = (x_1, \ldots, x_n) \in S$$

Візьмемо логарифм $\ln L(x, \theta) = -(n/2)\ln(2\pi\sigma^2) - (2\sigma^2)^{-1} \sum_{j=1}^{n} (x_j - \mu)^2$. Знайдемо стаціонарні точки:

$$\operatorname{grad}(\ln L(x,\theta)) = \begin{pmatrix} \frac{\partial}{\partial\mu} \ln L(x,\theta) \\ \frac{\partial}{\partial\sigma^2} \ln L(x,\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} (2\sigma^2)^{-1} \sum_{j=1}^{n} (x_j - \mu) = 0, \\ -(n/2)(\sigma^2)^{-1} + (1/2)(\sigma^2)^{-2} \sum_{j=1}^{n} (x_j - \mu)^2 = 0 \end{cases}$$

З отриманої вище системи знаходимо розв'язок вигляду

$$\theta_*(x) = \big(\mu_n(x), \sigma_n^2(x)\big) = \left(\frac{1}{n}\sum_{j=1}^{n} x_j, \frac{1}{n}\sum_{j=1}^{n}\left(x_j - \frac{1}{n}\sum_{k=1}^{n} x_k\right)^2\right)$$

Можна перевірити, що $\theta_*(x)$ є точкою максимума $\ln L(x, \theta)$. Дійсно, спочатку враховуємо, що сума квадратів відхилень найменша при середньому арифметичному значень

$$-(n/2)\ln(2\pi\sigma^2) - (2\sigma^2)^{-1} \sum_{j=1}^{n} (x_j - \mu)^2 \le -(n/2)\ln(2\pi\sigma^2) - (2\sigma^2)^{-1} \sum_{j=1}^{n} \left(x_j - \mu_n(x)\right)^2,$$

далі адаптуємо міркування попередньої задачі для оцінювання σ^2:

$$\begin{aligned} &-(n/2)\ln(2\pi\sigma^2) - (2\sigma^2)^{-1} \sum_{j=1}^{n} \left(x_j - \sum_{k=1}^{n} \mu_n(x)/n\right)^2 \le \\ &\le -(n/2)\ln(2\pi\sigma_n^2(x)) - (2\sigma_n^2(x))^{-1} \sum_{j=1}^{n} \left(x_j - \sum_{k=1}^{n} \mu_n(x)/n\right)^2 \end{aligned}$$

Значить

$$\hat{\theta}_n = \theta_*(X) = \left(\frac{1}{n}\sum_{j=1}^{n} X_j, \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \frac{1}{n}\sum_{k=1}^{n} X_k\right)^2\right)$$

є ОМВ параметра θ.

Зауважимо, що $\mathbf{E}_\theta[X_1] = \mu < \infty$, значить за критерієм Колмогорова про ПЗВЧ маємо $\sum_{j=1}^{n} X_j/n \to^{P_1} \mu$. З аналогічних міркувань $\sum_{j=1}^{n} X_j^2/n \to^{P_1} \mu^2 + \sigma^2$. Далі побачимо, що

$$\frac{1}{n}\sum_{j=1}^{n}\left(X_j - \frac{1}{n}\sum_{k=1}^{n} X_k\right)^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)^2 \to^{P_1} \mu^2 + \sigma^2 - (\mu)^2 = \sigma^2$$

внаслідок арифметичних дій над збіжними майже напевно послідовностями. Таким чином ми довели консистентність $\hat{\theta}_n$ для θ (побачте чому).

2.5 Задача 5

Нехай $X = (X_1, \dots, X_n)$ – кратна вибірка з розподілом спостережень $X_1 \sim \text{Geom}(\theta)$, $\theta \in (0, 1)$:

$$\mathbf{P}_\theta \left(X_1 = k \right) = \theta (1-\theta)^k, \ k \geq 0.$$

Побудувати ОМВ для θ. Дослідити оцінку на консистентність.

Розв'язання. Запишемо функцію вірогідності на $S = \mathbb{Z}_+^n$:

$$L(x, \theta) = \prod_{j=1}^n \mathbf{P}_\theta \left(X_j = x_j \right) = \theta^n (1-\theta)^{\sum_{j=1}^n x_j}, \ x = (x_1, \dots, x_n) \in S$$

Візьмемо логарифм

$$\ln L(x, \theta) = n \ln \theta + \sum_{j=1}^n x_j \ln(1-\theta)$$

Знаходимо точки екстремуму з рівняння

$$\frac{\partial}{\partial \theta} \ln L(x, \theta) = \frac{n}{\theta} - \frac{\sum_{j=1}^n x_j}{\theta(1-\theta)} = \frac{n(1-\theta) - \sum_{j=1}^n x_j \theta}{\theta(1-\theta)} = \frac{n - (n + \sum_{j=1}^n x_j)\theta}{\theta(1-\theta)} = 0$$

Розв'язуючи рівняння відносно θ отримаємо $\theta_*(x) = n/(n + \sum_{j=1}^n x_j) = 1/(1 + \sum_{j=1}^n x_j/n)$. Можна перевірити, що $\theta_*(x)$ є точкою максимума $\ln L(x, \theta)$ (функція вірогідності є увігнутою по θ, зростає при $\theta < \theta_*$ та спадає при $\theta > \theta_*$), значить

$$\hat{\theta}_n = \theta_*(X) = \frac{1}{1 + \sum_{j=1}^n X_j/n}$$

є ОМВ параметра θ.

Оскільки $\mathbf{E}_\theta \left[X_1 \right] = (1-\theta)/\theta < \infty$, то за критерієм Колмогорова про ПЗВЧ $\sum_{j=1}^n X_j/n \to^{P_1} (1-\theta)/\theta$, а за теоремою про арифметичні дії над збіжними майже напевно послідовностями маємо

$$\hat{\theta}_n = \frac{1}{1 + \sum_{j=1}^n X_j/n} \to^{P_1} \frac{1}{1 + (1-\theta)/\theta} = \frac{1}{(\theta + 1 - \theta)/\theta} = \theta,$$

що і доводить строгу консистентність ОМВ для θ.

2.6 Задача 6

Нехай $X = (X_1, \ldots, X_n)$ – кратна вибірка з розподілом спостережень $U[0, \theta]$, $\theta > 0$.

$$f(y; \theta) = \frac{1}{\theta}, \; y \in [0, \theta].$$

Побудувати ОМВ для θ. Дослідити оцінку на консистентність.

Розв'язання. Запишемо функцію вірогідності на $S = [0, \infty)^n$:

$$L(x, \theta) = \prod_{j=1}^{n} f(x_j; \theta) = \frac{1}{\theta^n} \cdot \mathbf{1}\{\max_{1 \le j \le n} x_j \le \theta\} = \begin{cases} \dfrac{1}{\theta^n}, & \max_{1 \le j \le n} x_j \le \theta \\ 0, & \max_{1 \le j \le n} x_j > \theta \end{cases}, \; x = (x_1, \ldots, x_n) \in S$$

Функція вірогідності є негладкою, тому треба трохи інакше діяти. Наприклад, зауважимо, що $L(x, \theta)$ спадна по θ при $\theta \ge \max\limits_{1 \le j \le n} x_j$ та $L(x, \theta) < L(x, \max_{1 \le j \le n} x_j)$ для всіх $\theta < \max\limits_{1 \le j \le n} x_j$. Отже, точкою максимума $L(x, \theta)$ є $\theta_*(x) = \max\limits_{1 \le j \le n} x_j$. Значить,

$$\hat{\theta}_n = \theta_*(X) = \max_{1 \le j \le n} X_j$$

є ОМВ параметра $\theta > 0$.

Консистентність оцінки було досліджено у попередніх роботах (доведення за означенням).

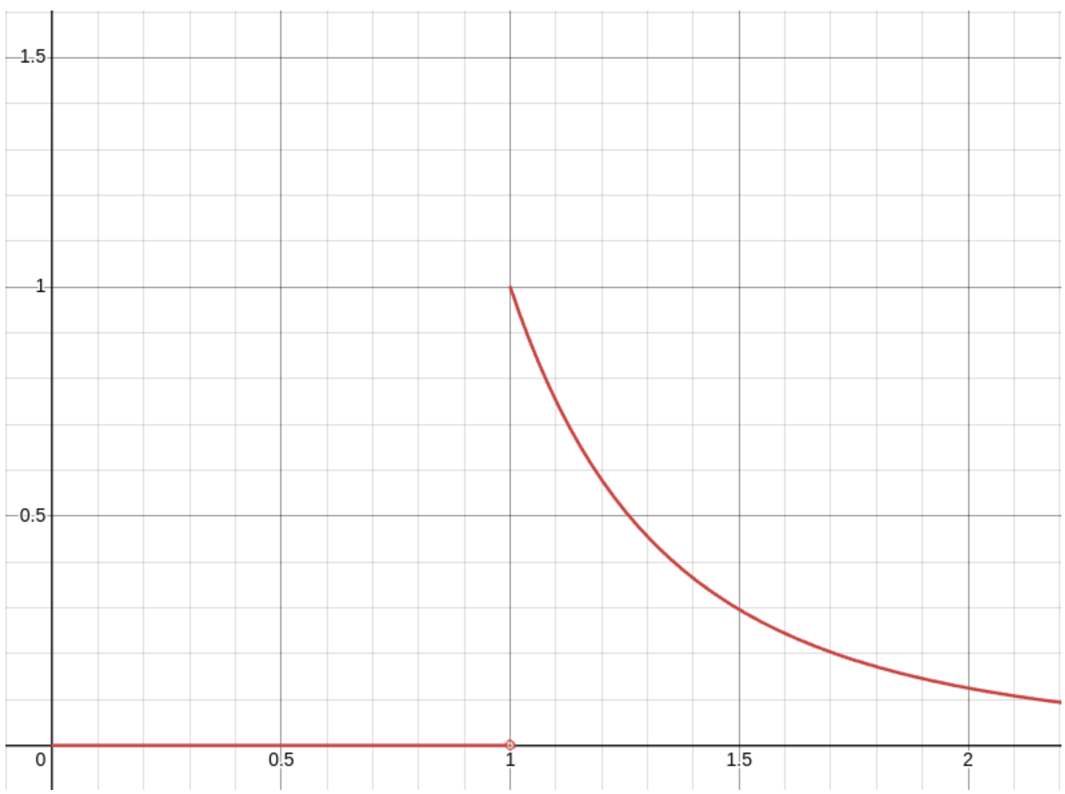

Рис. 1: Графік $L(x, \theta)$ по θ при $\max\limits_{1 \le j \le n} x_j = 1$.

2.7 Задача 7

Нехай $X = (X_1, X_2, X_3)$ – кратна вибірка з розподілом спостережень Лапласа:

$$f(y;\theta) = \frac{1}{2}\exp(-|y-\theta|),\ y \in \mathbb{R}.$$

Тут $\theta \in \mathbb{R}$.

Побудувати ОМВ для θ.

Запишемо функцію вірогідності на $S = \mathbb{R}^n$:

$$L(x,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = 8^{-1}\exp(-\sum_{j=1}^{3}|x_j-\theta|),\ x = (x_1,\ldots,x_n) \in S$$

Ще один приклад негладкої функції, розглянемо наступне дослідження на максимум.

Для того, що максимізувати $L(x,\theta)$, потрібно максимізувати $-\sum_{j=1}^{3}|x_j-\theta|$, тобто мінімізувати $g(\theta) = \sum_{j=1}^{3}|x_j-\theta|$ по θ (випливає з монотонності функцій).

Припустимо, що $x_1 < x_2 < x_3$. Запишемо суму модулів явно:

$$g(\theta) = |x_1-\theta|+|x_2-\theta|+|x_3-\theta| = \begin{cases} (x_1-\theta)+(x_2-\theta)+(x_3-\theta) = (x_1+x_2+x_3)-3\theta, & \theta < x_1 \\ (\theta-x_1)+(x_2-\theta)+(x_3-\theta) = (-x_1+x_2+x_3)-\theta, & x_1 \le \theta < x_2 \\ (\theta-x_1)+(\theta-x_2)+(x_3-\theta) = (-x_1-x_2+x_3)+\theta, & x_2 \le \theta < x_3 \\ (\theta-x_1)+(\theta-x_2)+(\theta-x_3) = -(x_1+x_2+x_3)+3\theta, & x_3 \le \theta \end{cases}$$

Отримана функція є неперервною по θ, спадна на $(-\infty, x_2)$ та зростаюча на $(x_2, +\infty)$. Отже, точкою мінімуму суми модулів є $\theta_*(x) = x_2$, що, власне, буде точкою максимума для $L(x,\theta)$. Але ж ми вважали, що x_j впорядковані за зростанням. Значить, ОМВ в даному випадку є медіаною вибірки $\hat{\theta} = \theta_*(X) = \mathrm{median}(X)$.

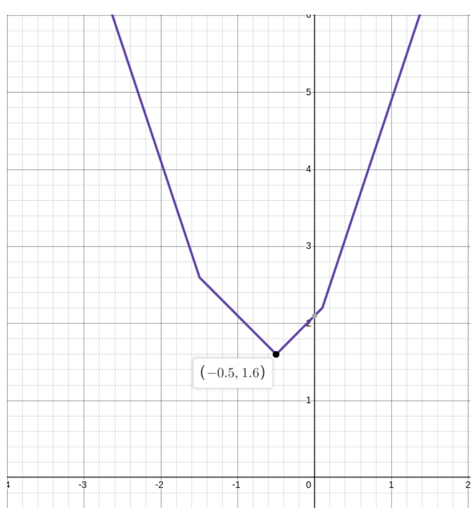

Рис. 2: Графік $g(\theta)$ при $(x_1, x_2, x_3) = (-1.5, -0.5, 0.1)$. Мінімум досягається на x_2.

2.8 Задача 8

Нехай $X = (X_1, X_2, X_3)$ – кратна вибірка з розподілом спостережень Пуассона:

$$\mathbf{P}_\theta (X_1 = k) = \frac{\theta^k}{k!} \exp(-\theta),\ k \geq 0.$$

Тут $\theta > 0$.

Побудувати ОМВ для θ. Дослідити оцінку на консистентність. Побудувати ОМВ для $\mathbf{P}_\theta (X_1 = 0)$.

Розв'язання. Запишемо функцію вірогідності на $S = \mathbb{Z}_+^n$:

$$L(x, \theta) = \prod_{j=1}^{n} \mathbf{P}_\theta (X_j = x_j) = \prod_{j=1}^{n} (x_j!)^{-1} \cdot \theta^{\sum_{j=1}^n x_j} \exp(-n\theta),\ x = (x_1, \ldots, x_n) \in S$$

Візьмемо логарифм

$$\ln L(x, \theta) = \sum_{j=1}^{n} x_j \ln \theta - n\theta$$

Знаходимо точки екстремуму з рівняння

$$\frac{\partial}{\partial \theta} \ln L(x, \theta) = \frac{\sum_{j=1}^n x_j}{\theta} - n = 0$$

Розв'язуючи рівняння відносно θ отримаємо розв'язок $\theta_*(x) = \sum_{j=1}^n x_j/n$. Можна перевірити, що $\theta_*(x)$ є точкою максимума $\ln L(x, \theta)$ (функція вірогідності є увігнутою по θ, зростає при $\theta < \theta_*$ та спадає при $\theta > \theta_*$), значить

$$\hat{\theta}_n = \theta_*(X) = \sum_{j=1}^{n} X_j/n$$

є ОМВ параметра θ.

Оскільки $\mathbf{E}_\theta [X_1] = \theta < \infty$, то за критерієм Колмогорова про ПЗВЧ $\hat{\theta}_n \to^{P_1} \theta$, тобто оцінка є строго консистентною для θ.

Далі, скористаємося теоремою про інваріантність ОМВ. Функція $\tau(\theta) = \mathbf{P}_\theta (X_1 = 0) = \exp(-\theta)$ є взаємно однозначною на $(0, \infty)$ (переконайтеся в цьому). Отже, оцінка $\tau(\hat{\theta}_n) = \exp(-\hat{\theta}_n)$ є ОМВ функції $\tau(\theta) = \mathbf{P}_\theta (X_1 = 0)$.