Точкове оцінювання.

Лекція № 2

2.1. Метод моментів
2.2. Метод макимальної вірогідності

ВСТУП

Позначимо через $\mathcal{F}_\theta$ – деякий клас розподілів, який цілком визначається значенням скалярного чи векторного параметра θ. θ приймає значення з деякої множини Θ.

Приклад 1

Наведемо декілька прикладів таких розподілів:

- розподіл Бернуллі $p \in (0, 1)$;
- біноміальний розподіл $Bin(n, p)$, $n \in \mathbb{N}$, $p \in (0, 1)$;
- рівномірний розподіл $U(a, b)$, $a < b$;
- рівномірний розподіл $U(0, \theta)$, $\theta > 0$;
- Гауссівський розподіл $N(a, \sigma^2)$, $a \in \mathbb{R}$, $\sigma^2 > 0$.

ВСТУП

Нехай $\xi_1, \ldots, \xi_n$ є вибіркою об'єма n із параметричної сім'ї розподілів $\mathcal{F}_\theta$, $\theta \in \Theta$.

Означення 1

Статистикою називається довільна борелева функція $\theta^* = \theta^*(\xi_1, \ldots, \xi_n)$ від елементів вибірки.

Зауваження 1

Зазначимо, що статистика є функцією від емпіричних даних, але ніяк не від параметра θ. Вона, як правило, призначена саме для оцінювання невідомого параметра θ (тому її також називають оцінкою)) і вже тому від нього залежати не може.

Функція $f : \mathbb{R}^n \to \mathbb{R}$ називається борелевою, якщо прообраз будь-якої борелевої множини з $\mathbb{R}$ при такому перетворенні є борелевою множиною з $\mathbb{R}^n$.

Розглянемо деякі стандартні методи отримання точкових оцінок. Метод моментів пропонує для знаходження оцінки невідомого параметра використовувати вибіркові моменти замість істинних. Цей метод полягає в наступному: деякий момент випадкової величини ξ_1 може бути функцією від параметра θ. Але тоді параметр θ, в свою чергу, може бути функцією від теоретичного k-го момента. Підставляючи в цю функцію замість невідомого торетичного k-го момента його вибірковий аналог, отримаємо замість параметhа θ його оцінку θ^*.

2.1. Метод моментів

Нехай $\xi_1, \ldots, \xi_n$ є вибіркою об'єма n із параметричної сім'ї розподілів $\mathcal{F}_\theta$, $\theta \in \Theta \subset \mathbb{R}$. Оберемо деяку функцію $g : \mathbb{R} \to \mathbb{R}$ так, щоб існував момент

$$Eg(\xi_1) = h(\theta), \tag{1}$$

і функцію h можна обернути на множині Θ.

Розв'язавши (1) відносно θ отримаємо, що $\theta = h^{-1}\left(Eg(\xi_1)\right)$.
Підставивши в отриману формулу замість істинного момента відповідний вибірковий момент будемо мати

$$\theta^* = h^{-1}\left(\overline{g(\xi)}\right) = h^{-1}\left(\frac{1}{n}\sum_{j=1}^{n} g(\xi_j)\right).$$

Отримана таким чином оцінка θ^* називається оцінкою метода моментів (ОММ) параметра θ.
Частіш за все, в якості функції $g(y)$ обирають $g(y) = y^k$ тобто, використовують моменти k-го порядку.

Зауваження 2

Якщо параметр генеральної сукупності є векторним $\theta = (\theta_1, \ldots, \theta_q)$, тоді обирають низку функцій $g_1, \ldots, g_q$ і розв'язують систему рівнянь, яка повинна однозначно розв'язуватись відносно $\theta_1, \ldots, \theta_q$.

Розглянемо приклади знаходження оцінок методом моментів.

Приклад 2

Нехай маємо рівномірний розподіл $\xi_j \sim U(a,b)$.

Тоді, використовуючи те, що $E\xi_1 = \dfrac{a+b}{2}$, $D\xi = \dfrac{(b-a)^2}{12}$, отримаємо

$$\begin{cases} \bar{\xi} = \dfrac{a^* + b^*}{2} \\ s^2 = \dfrac{(b^* - a^*)^2}{12} \end{cases} \Leftrightarrow \begin{cases} \bar{\xi} = \dfrac{a^* + b^*}{2} \\ \sqrt{3s^2} = \dfrac{b^* - a^*}{2} \end{cases} \Leftrightarrow \begin{cases} a^* = \bar{\xi} - \sqrt{3s^2} \\ b^* = \bar{\xi} + \sqrt{3s^2} \end{cases}.$$

Приклад 3

Нехай маємо рівномірний розподіл $\xi_j \sim U(0, \theta)$.

Тоді, використовуючи те, що $E\xi_1 = \dfrac{\theta}{2}$, отримаємо

$$\overline{\xi} = \frac{\theta^*}{2} \Leftrightarrow \theta^* = 2\overline{\xi}.$$

Розглянемо, як можна знайти оцінку методом моментів через k-й момент. Можна показати, що $E\xi_1^k = \dfrac{\theta^k}{k+1}$, звідки випливає, що

$$\overline{\xi^k} = \frac{(\theta^*)^k}{k+1} \Leftrightarrow \theta^* = \sqrt[k]{(k+1) \cdot \overline{\xi^k}}.$$

Приклад 4

Нехай маємо розподіл Пуассона $\xi_j \sim Poiss(\lambda)$.
Тоді, оскільки $E\xi_1 = D\xi_1 = \lambda$, отримаємо дві оцінки для параметра λ

$$\lambda^* = \overline{\xi}, \quad \lambda^{**} = s^2.$$

Приклад 5

Нехай маємо біноміальний розподіл $\xi_j \sim Bin(N, p)$.
Враховуючи те, що $E\xi = Np$, $D\xi = Np(1-p)$, будемо мати

$$\begin{cases} \bar{\xi} = N^* p^* \\ s^2 = N^* p^* (1 - p^*) \end{cases} \Leftrightarrow \begin{cases} p^* = 1 - \frac{s^2}{\bar{\xi}} \\ N^* = \frac{\bar{\xi}}{p^*} = \frac{\left(\bar{\xi}\right)^2}{\bar{\xi} - s^2} \end{cases} .$$

Метод максимальної вірогідності полягає в тому, що у якості "найбільш вірогідного" значення параметра беруть значення θ, яке максимізує ймовірність отримати при n випробуваннях дану вибірку

$$(x_1, \ldots, x_n).$$

Воно залежить від вибірки і є шуканою оцінкою.
З'ясуємо спочатку, що таке"ймовірність отримати дану вибірку", тобто що ми маємо максимізувати. Для дискретних розподілів $\mathcal{F}_\theta$ ймовірність попасти в точку y дорівнює $P\{\xi_1 = y\}$. Для абсолютно неперервних розподілів розглядається "майже" ймовірність попадання в точку y (з точністю до dy):

$$P\{\xi_1 \in (y, y + dy\} \approx f_\theta(y)dy,$$

де $f_\theta(y)$ - щільність розподілу $\mathcal{F}_\theta$.

Означення 2

Функція

$$L(\vec{x};\theta) = \begin{cases} \prod_{j=1}^{n} P\{\xi_j = x_j\}, & \text{якщо } \mathcal{F}_\theta \text{ має дискретний розподіл ;} \\ \prod_{j=1}^{n} f_\theta(x_j), & \text{якщо } \mathcal{F}_\theta \text{ має неперервний розподіл,} \end{cases}$$

називається функцією вірогідності.

При фіксованому θ дана функція є в.в.

Наша задача полягає у максимізації функції вірогідності $f(\vec{x};\theta)$ відносно параметру θ. Введемо функцію

$$\ln L(\vec{x};\theta) = \begin{cases} \sum_{j=1}^{n} \ln P\{\xi_j = x_j\}, & \text{якщо } \mathcal{F}_\theta \text{ має дискретний розподіл ;} \\ \sum_{j=1}^{n} \ln f_\theta(x_j), & \text{якщо } \mathcal{F}_\theta \text{ має неперервний розподіл,} \end{cases}$$

яка називається логарифмічною функцією вірогідності.

Означення 3

Оцінкою максимальної вірогідності(ОМВ) $\hat{\theta}$ невідомого параметра θ називається таке значення θ, при якому досягається максимум функції $L(\vec{x};\theta)$.

Зауваження 3

В силу монотонності функції $\ln y$ максимуми функцій $L(\vec{x};\theta)$ та $\ln L(\vec{x};\theta)$. Тому ОМВ можна назвати точку максимуму функції $\ln L(\vec{x};\theta)$ за змінною θ.

Зауваження 4

В силу монотонності функції $\ln y$ максимуми функцій $L(\vec{x};\theta)$ та $\ln L(\vec{x};\theta)$. Тому ОМВ можна назвати точку максимуму функції $\ln L(\vec{x};\theta)$ за змінною θ.

Приклад 6

Нехай маємо рівномірний розподіл $\xi_j \sim U(0,\theta)$.
Знайдемо функцію вірогідності. Оскільки щільність розподілу має вигляд

$$f_{U(0,\theta)}(x) = \begin{cases} \frac{1}{\theta}, & x \in (0,\theta), \\ 0, & x \notin (0,\theta), \end{cases}$$

то функція вірогідності має бути вигляду

$$L(\vec{x},\theta) = \begin{cases} \frac{1}{\theta^n}, & \text{якщо для всіх } j=\overline{1,n}\ x_j \in (0,\theta), \\ 0, & \text{якщо хоча б для одного } j\ x_j \notin (0,\theta). \end{cases}$$

Очевидно, що шукана оцінка θ^* має бути більшою за найбільше значення у вибірці. Тоді $L(\vec{x},\theta) = \frac{1}{\theta^n}$ і вона буде набувати найбільшого значення при найменшому можливому значенню θ, тобто, за наших обмежень, $\theta^* = \max\{x_1,\ldots,x_n\}$.

Приклад 7

Нехай маємо розподіл Пуассона $\xi_j \sim Poiss(\lambda)$. Тоді функція вірогідності буде мати вигляд

$$L(\vec{x}, \lambda) = \prod_{j=1}^{n} \frac{\lambda^{x_j}}{x_j!} e^{-\lambda},$$

а логарифмічна функція вірогідності

$$\ln L(\vec{x}, \lambda) = \ln \lambda \sum_{j=1}^{n} x_j - \sum_{j=1}^{n} \ln(x_j!) - \lambda n.$$

Для знаходження максимуму маємо розв'язати рівняння

$$\frac{d \ln L(\vec{x}, \lambda)}{d\lambda} = 0 \;\Leftrightarrow\; \frac{1}{\lambda} \sum_{j=1}^{n} x_j - n = 0 \;\Leftrightarrow\; \lambda = \frac{1}{n} \sum_{j=1}^{n} x_j = \overline{x}.$$

Таким чином ОМВ $\lambda^* = \overline{x}$.

Приклад 8

Нехай маємо біноміальний розподіл $\xi_j \sim Bin(N, p)$, причому N будемо вважати відомим. Тоді функція вірогідності буде мати вигляд

$$L(\vec{x}, p) = \prod_{j=1}^{n} C_N^{x_j} p^{x_j} (1-p)^{N-x_j} = \left(\prod_{j=1}^{n} C_N^{x_j} \right) \cdot p^{\sum\limits_{j=1}^{n} x_j} (1-p)^{Nn - \sum\limits_{j=1}^{n} x_j},$$

а логарифмічна функція вірогідності

$$\ln L(\vec{x}, p) = \sum_{j=1}^{n} \ln C_N^{x_j} + \ln p \sum_{j=1}^{n} x_j + \ln(1-p) \left(Nn - \sum_{j=1}^{n} x_j \right).$$

Для знаходження максимуму маємо розв'язати рівняння

$$\frac{d \ln L(\vec{x}, p)}{dp} = 0 \;\Leftrightarrow\; \frac{1}{p} \sum_{j=1}^{n} x_j - \frac{1}{1-p} \left(Nn - \sum_{j=1}^{n} x_j \right) \;\Leftrightarrow\; p = \frac{1}{Nn} \sum_{j=1}^{n} x_j = \frac{\overline{x}}{N}.$$

Таким чином ОМВ $p^* = \dfrac{\overline{x}}{N}$.

Дякуємо за увагу!