

ТЕОРІЯ ЙМОВІРНОСТЕЙ, СТАТИСТИКА ТА ЙМОВІРНІСНІ ПРОЦЕСИ

Лекція 13.

Розора Ірина Василівна

Київ, 2020

Зміст I

- 1 Вибіркові моменти. Метод моментів
- 2 Метод моментів
- 3 Оцінки максимальної вірогідності

Зміст

- 1 Вибіркові моменти. Метод моментів
- 2 Метод моментів
- 3 Оцінки максимальної вірогідності

Теоретичні та вибіркові моменти

Велику групу статистик утворюють вибіркові моменти, які за ЗВЧ є природними оцінками для теоретичних моментів – математичних сподівань степеневих функцій від випадкової величини.

Момент порядку k

Нехай ξ – випадкова величина. Її (нецентральним) теоретичним моментом порядку $k \in \mathbb{N}$ називається число

$$\mu_k \equiv M\xi^k,$$

за умови інтегрованості величини ξ^k .

Центральний момент порядку k

Центральним теоретичним моментом порядку k називається число

$$\mu_k^0 \equiv \mathbf{M}(\xi - \mu)^k,$$

де $\mu \equiv \mu_1$ – математичне сподівання, $\mu_2^0 = \sigma^2$ – дисперсія ξ .

Значення центральних моментів використовуються в теорії розподілів для означення таких спеціальних характеристик:

Коефіцієнт варіації

$$k_v = \frac{\sigma}{\mu_1}$$

(дорівнює 1 для показникового розподілу)

Коефіцієнт скошеності

$$k_s = \frac{3(\mu_1 - x_{\frac{1}{2}})}{\sigma}$$

(нульовий для симетричних розподілів)

Коефіцієнт асиметрії

$$k_a = \frac{\mu_3^0}{\sigma^3}$$

(нульовий для симетричних розподілів)

Коефіцієнт ексцесу

$$k_e = \frac{\mu_4^0}{\sigma^4} - 3$$

(нульовий для нормальних спостережень).

Всі наведені характеристики є безрозмірними та відображають певні особливості форми відповідного розподілу.

Вибірковий момент порядку k

Нехай $X = (\xi_1, \dots, \xi_n)$ – вибірка. Її (нецентральним) вибірковим моментом порядку k називається статистика

$$\hat{\mu}_{kn} \equiv \frac{1}{n} \sum_{i=1}^n \xi_i^k.$$

Центральний вибірковий момент порядку k

Центральним вибірковим моментом порядку k називається статистика

$$\hat{\mu}_{kn}^0 \equiv \frac{1}{n} \sum_{i=1}^n (\xi_i - \hat{\mu}_n)^k,$$

де $\hat{\mu}_n \equiv \hat{\mu}_{1n}$ – вибіркове середнє, а, $\hat{\mu}_{2n}^0 = \hat{\sigma}_n^2$ – вибіркова дисперсія.

Вибірковий момент порядку k

Нехай $X = (\xi_1, \dots, \xi_n)$ – вибірка. Її (нецентральним) вибірковим моментом порядку k називається статистика

$$\hat{\mu}_{kn} \equiv \frac{1}{n} \sum_{i=1}^n \xi_i^k.$$

Центральний вибірковий момент порядку k

Центральним вибірковим моментом порядку k називається статистика

$$\hat{\mu}_{kn}^0 \equiv \frac{1}{n} \sum_{i=1}^n (\xi_i - \hat{\mu}_n)^k,$$

де $\hat{\mu}_n \equiv \hat{\mu}_{1n}$ – вибіркове середнє, а, $\hat{\mu}_{2n}^0 = \hat{\sigma}_n^2$ – вибіркова дисперсія.

Зауваження

Важливою властивістю центральних моментів є інваріантність відносно зсувів – вони не змінюються при одночасному зсуві всіх спостережень на сталу:

$$\begin{aligned}\hat{\mu}_{kn}^0 &= \frac{1}{n} \sum_{i=1}^n \left(\xi_i - \frac{1}{n} \sum_{j=1}^n \xi_j \right)^k \\ &= \frac{1}{n} \sum_{i=1}^n \left(\xi_i - c - \frac{1}{n} \sum_{j=1}^n (\xi_j - c) \right)^k.\end{aligned}$$

Зауваження

Як і в теоремі про властивості дисперсії, для вибіркової дисперсії має місце тотожність

$$\hat{\sigma}_n^2 = \hat{\mu}_{2,n} - (\hat{\mu}_n)^2.$$

Дійсно,

$$\begin{aligned}\hat{\sigma}_n^2 &= \frac{1}{n} \sum_{i=1}^n (\xi_i - \hat{\mu}_n)^2 \\ &= \frac{1}{n} \sum_{i=1}^n \xi_i^2 - 2(\hat{\mu}_n)^2 + (\hat{\mu}_n)^2.\end{aligned}$$

Теорема (про моменти вибірових моментів)

Нехай (ξ_1, \dots, ξ_n) – кратна вибірка. Тоді

$$M\hat{\mu}_{kn} = \mu_k,$$

$$D\hat{\mu}_{kn} = \frac{1}{n}(\mu_{2k} - \mu_k^2).$$

Зокрема,

$$M\hat{\mu}_n = \mu, \quad D\hat{\mu}_n = \frac{\sigma^2}{n}.$$

Крім того,

$$M\hat{\sigma}_n^2 = \frac{n-1}{n}\sigma^2, \quad D\hat{\sigma}_n^2 = \frac{n-1}{n^3}((n-1)\mu_4^0 - (n-3)\sigma^4).$$

Доведення

Незміщеність нецентральных моментів є очевидним наслідком лінійності математичного сподівання. Вираз для їх дисперсій впливає з незалежності в сукупності і однакової розподіленості степеневих функцій $(\xi_i^k, i = \overline{1, n})$ та з теореми про дисперсію суми незалежних величин:

$$\begin{aligned} D \left(\frac{1}{n} \sum_{i=1}^n \xi_i^k \right) &= \frac{1}{n^2} \sum_{i=1}^n D(\xi_i^k) \\ &= \frac{1}{n^2} n D(\xi_1^k) = \frac{1}{n} (M \xi_1^{2k} - (M \xi_1^k)^2). \end{aligned}$$

З теореми про моменти вибірових моментів випливає, що вибіркова дисперсія $\hat{\sigma}_n^2$ є зміщеною оцінкою для теоретичної дисперсії σ^2 . Тому часто використовують її скоригований незміщений варіант.

Нормована вибіркова дисперсія

Нормованою вибірковою дисперсією є статистика

$$\hat{s}_n^2 \equiv \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \hat{\mu}_n)^2 = \frac{n}{n-1} \hat{\sigma}_n^2,$$

що є незміщеною оцінкою дисперсії:

$$\mathbf{M}\hat{s}_n^2 = \frac{n}{n-1} \mathbf{M}\hat{\sigma}_n^2 = \sigma^2$$

Теорема (про асимптотичні властивості вибірових моментів)

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка.

(а) Якщо

$$M|\xi_1|^k < \infty,$$

то нецентральний вибіровий момент $\hat{\mu}_{kn}$ є незміщеною та конзистентною оцінкою теоретичного моменту μ_k .

(б) Якщо

$$M\xi_1^{2k} < \infty,$$

то оцінки $\hat{\mu}_{kn}$ є асимптотично нормальними:

$$\sqrt{n}(\hat{\mu}_{kn} - \mu_k) \xrightarrow{W} \eta \cong N(0, \mu_{2k} - \mu_k^2).$$

Зміст

- 1 Вибіркові моменти. Метод моментів
- 2 Метод моментів
- 3 Оцінки максимальної вірогідності

Метод моментів

Метод моментів є спеціальним методом оцінювання невідомих параметрів, який спирається на асимптотичні властивості вибірових моментів.

Припустимо, що параметричний простір є d -вимірним:

$$\Theta \in \mathbb{R}^d.$$

Оскільки розподіл вибірки $X = (\xi_1, \dots, \xi_n)$ відомий повністю при заданому значенні $\theta \in \Theta$, то повністю відомими є функції

$$\mu_k(\theta) \equiv \mathbf{M}_{\theta} \xi_1^k.$$

Розглянемо векторну функцію

$$\mu^{(d)}(\theta) \equiv (\mu_k(\theta), k = \overline{1, d}): \Theta \rightarrow \mathbb{R}^d.$$

Припустимо, що існує неперервне відображення

$$T_d(\mu): \mathbf{R}^d \rightarrow \Theta,$$

яке є оберненим до $\mu^{(d)}(\theta)$, тобто

$$T_d(\mu^{(d)}(\theta)) = \theta, \quad \forall \theta \in \Theta.$$

Ця умова виконується, зокрема, за теоремою про обернене відображення з курсу математичного аналізу, якщо функція $\mu^{(d)}(\theta)$ неперервно диференційовна, якобіан

$$\det \left| \frac{d}{d\theta} \mu^{(d)}(\theta_0) \right| \neq 0$$

для деякого $\theta_0 \in \Theta$ і простір Θ звужено до деякого околу точки θ_0 .

Оцінка методу моментів

Оцінкою методу моментів параметра - називається така статистика від вектора вибірових моментів

$\hat{\mu}_n^{(d)} = (\hat{\mu}_{nk}, k = \overline{1, d})$, що містить значення перших d вибірових моментів:

$$\hat{\theta}_n \equiv T_d(\hat{\mu}_n^{(d)}),$$

де $T_d : \mathbf{R} \rightarrow \Theta$ – обернена функція до вектора моментів $\mu^{(d)}(\theta)$.

Зауваження

З означення оберненої функції випливає, що оцінка методу моментів $\hat{\theta}_n$ є єдиним розв'язком системи рівнянь методу моментів

$$\mu^{(d)}(\hat{\theta}_n) = \hat{\mu}_n^{(d)}.$$

Зауваження

У деяких випадках кількість дійсно залежних від θ координат вектора $\mu^{(d)}(\theta)$ може бути меншою за d , тому відображення $T_d(\mu)$ не існує – наприклад, при рівномірному на $[-\theta, \theta]$ розподілу спостережень та $d = 1$. У цьому разі необхідно збільшити розмірність d вектора $\mu^{(d)}(\theta)$.

Теорема (про конзистентність оцінок методу моментів)

Якщо $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка,

$$M_{\theta} |\xi_1|^d < \infty,$$

і функція T_d неперервна, то оцінка методу моментів є конзистентною оцінкою параметра θ .

Зауваження

Якщо

$$M_{\theta} \xi_1^{2d} < \infty$$

і функція T_d неперервно диференційовна, то можна довести, що оцінка методу моментів є асимптотично нормальною з матрицею

$$t_d = \frac{d}{d\mu} T_d(\mu) \Big|_{\mu=\mu^{(d)}(\theta)},$$

тобто

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{w} \eta \cong N_d(0, t_d V^{(d)} t_d'), \quad n \rightarrow \infty.$$

Приклад 1. Оцінка параметрів гама-розподілу

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка з гама-розподілом $\xi_1 \cong \Gamma(\lambda, \alpha)$, невідомий параметр $\theta = (\lambda, \alpha)$, $\dim \theta = 2$. Тоді теоретичні моменти мають вигляд

$$\mu_1(\theta) = \frac{\alpha}{\lambda}, \quad \mu_2(\theta) = \frac{\alpha}{\lambda^2} + \frac{\alpha^2}{\lambda^2}.$$

Оцінку методу моментів $(\hat{\alpha}, \hat{\lambda})$ знаходимо з системи рівнянь

$$\frac{\hat{\alpha}}{\hat{\lambda}} = \hat{\mu}_n, \quad \frac{\hat{\alpha}}{\hat{\lambda}^2} + \frac{\hat{\alpha}^2}{\hat{\lambda}^2} = \hat{\mu}_{2n} = \hat{\sigma}_n^2 + (\hat{\mu}_n)^2,$$

звідки дістанемо

$$\hat{\lambda} = \frac{\hat{\mu}_n}{\hat{\sigma}_n^2}, \quad \hat{\alpha} = \frac{(\hat{\mu}_n)^2}{\hat{\sigma}_n^2}.$$

Приклад 1. Оцінка параметрів гама-розподілу

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка з гама-розподілом $\xi_1 \cong \Gamma(\lambda, \alpha)$, невідомий параметр $\theta = (\lambda, \alpha)$, $\dim \theta = 2$. Тоді теоретичні моменти мають вигляд

$$\mu_1(\theta) = \frac{\alpha}{\lambda}, \quad \mu_2(\theta) = \frac{\alpha}{\lambda^2} + \frac{\alpha^2}{\lambda^2}.$$

Оцінку методу моментів $(\hat{\alpha}, \hat{\lambda})$ знаходимо з системи рівнянь

$$\frac{\hat{\alpha}}{\hat{\lambda}} = \hat{\mu}_n, \quad \frac{\hat{\alpha}}{\hat{\lambda}^2} + \frac{\hat{\alpha}^2}{\hat{\lambda}^2} = \hat{\mu}_{2n} = \hat{\sigma}_n^2 + (\hat{\mu}_n)^2,$$

звідки дістанемо

$$\hat{\lambda} = \frac{\hat{\mu}_n}{\hat{\sigma}_n^2}, \quad \hat{\alpha} = \frac{(\hat{\mu}_n)^2}{\hat{\sigma}_n^2}.$$

Приклад 2. Оцінка параметрів рівномірного розподілу

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка з рівномірним розподілом $\xi_1 \cong U(a, b)$, невідомий параметр $\theta = (a, b)$, $\dim \theta = 2$. Тоді

$$\mu_1(\theta) = \frac{a+b}{2}, \quad \mu_2(\theta) = \mu_1^2(\theta) + \frac{(b-a)^2}{12}.$$

Оцінку методу моментів (\hat{a}, \hat{b}) знаходимо з системи

$$\frac{\hat{a} + \hat{b}}{2} = \hat{\mu}_n, \quad \frac{(\hat{b} - \hat{a})^2}{12} = \hat{\sigma}_n^2,$$

$$\hat{b} = \hat{\mu}_n + \sqrt{3}\hat{\sigma}_n, \quad \hat{a} = \hat{\mu}_n - \sqrt{3}\hat{\sigma}_n.$$

Приклад 2. Оцінка параметрів рівномірного розподілу

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка з рівномірним розподілом $\xi_1 \cong U(a, b)$, невідомий параметр $\theta = (a, b)$, $\dim \theta = 2$. Тоді

$$\mu_1(\theta) = \frac{a+b}{2}, \quad \mu_2(\theta) = \mu_1^2(\theta) + \frac{(b-a)^2}{12}.$$

Оцінку методу моментів (\hat{a}, \hat{b}) знаходимо з системи

$$\frac{\hat{a} + \hat{b}}{2} = \hat{\mu}_n, \quad \frac{(\hat{b} - \hat{a})^2}{12} = \hat{\sigma}_n^2,$$

$$\hat{b} = \hat{\mu}_n + \sqrt{3}\hat{\sigma}_n, \quad \hat{a} = \hat{\mu}_n - \sqrt{3}\hat{\sigma}_n.$$

Приклад 3. Оцінка параметрів логнормального розподілу

В.в. ξ має логнормальний розподіл, $\xi \cong LN(\mu, \sigma^2)$, якщо її логарифм має нормальний розподіл:

$$\ln \xi \cong N(\mu, \sigma^2).$$

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка з логнормальним розподілом спостережень: $\xi_1 \cong LN(\mu, \sigma^2)$.

Для оцінки параметрів методом моментів можна скористатися перетворенням вибірки за допомогою логарифмічної функції. За теоремою про перетворення незалежних величин статистика $Y = (\eta_1, \dots, \eta_n)$ з координатами

$$\eta_k = \ln \xi_k, \quad k = \overline{1, n},$$

є кратною вибіркою, а її елементи мають нормальний розподіл.

Приклад 3 (продовження)

За теоремою про інтерпретацію параметрів нормального розподілу вибіркові середнє та дисперсія вектора Y :

$$\hat{\mu}_n = \frac{1}{n} \sum_{k=1}^n \ln \xi_k,$$

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{k=1}^n (\ln \xi_k - \hat{\mu}_n)^2$$

є конзистентними оцінками параметрів μ , σ^2 .

Зміст

- 1 Вибіркові моменти. Метод моментів
- 2 Метод моментів
- 3 Оцінки максимальної вірогідності**

Визначення оцінки максимальної вірогідності ґрунтується на принципі максимальної вірогідності:

"те, що спостерігається, є найбільш імовірним серед усіх можливих альтернатив".

Надалі будемо припускати, що вибірка X задовольняє умову підпорядкованості її розподілу деякій мірі у вибіркового просторі. За такої умови повністю визначена вибіркова функція вірогідності $L(X, \theta)$. Значення цієї функції і дають критерій "найбільшої вірогідності".

Оцінка максимальної вірогідності

Оцінкою максимальної вірогідності (ОМВ) параметра θ за вибіркою X називається статистика, що максимізує вибірккову функцію вірогідності $L(X, \theta)$:

$$\hat{\theta} = \hat{\theta}(X) \equiv \arg \max_{\theta \in \Theta} L(X, \theta),$$

тобто це така статистика $\hat{\theta} = \hat{\theta}(X)$, що задовольняє умову:

$$L(X, \theta) \leq L(X, \hat{\theta}), \forall \theta \in \Theta.$$

Для кратної вибірки ОМВ позначається як $\hat{\theta}_n$, де n – об'єм вибірки.

Оцінка максимальної вірогідності

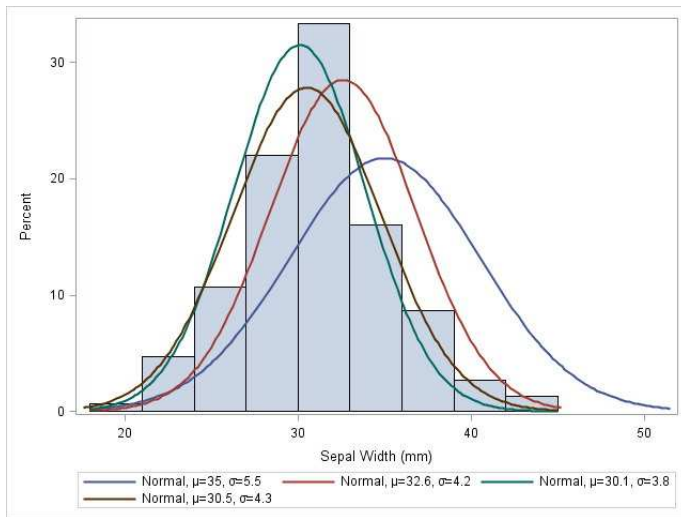
Оцінкою максимальної вірогідності (ОМВ) параметра θ за вибіркою X називається статистика, що максимізує вибірккову функцію вірогідності $L(X, \theta)$:

$$\hat{\theta} = \hat{\theta}(X) \equiv \arg \max_{\theta \in \Theta} L(X, \theta),$$

тобто це така статистика $\hat{\theta} = \hat{\theta}(X)$, що задовольняє умову:

$$L(X, \theta) \leq L(X, \hat{\theta}), \forall \theta \in \Theta.$$

Для кратної вибірки ОМВ позначається як $\hat{\theta}_n$, де n – об'єм вибірки.



Іноді простішим є обчислення ОМВ з еквівалентного означення

$$\hat{\theta} = \hat{\theta}(X) \equiv \arg \max_{\theta \in \Theta} \ln L(X, \theta),$$

що збігається з основним через монотонність логарифмічної функції.

Необхідною умовою існування ОМВ є припущення існування точки максимуму. ОМВ визначена однозначно за умови єдиності цієї точки.

Іноді простішим є обчислення ОМВ з еквівалентного означення

$$\hat{\theta} = \hat{\theta}(X) \equiv \arg \max_{\theta \in \Theta} \ln L(X, \theta),$$

що збігається з основним через монотонність логарифмічної функції.

Необхідною умовою існування ОМВ є припущення існування точки максимуму. ОМВ визначена однозначно за умови єдиності цієї точки.

Теорема (про рівняння максимальної вірогідності)

Якщо параметр є векторним: $\Theta \subset R^d$, максимум в означенні ОМВ досягається всередині параметричної множини, а функція вірогідності диференційовна, то ОМВ задовольняє рівняння максимальної вірогідності:

$$\frac{\partial}{\partial \theta} \ln L(X, \theta)|_{\theta=\hat{\theta}} = 0,$$

тобто ОМВ є коренем функції впливу $U(X, \theta)$:

$$U(X, \hat{\theta}) = 0.$$

Якщо $\dim \theta = d > 1$, то рівняння МВ є векторним, і перетворюється на систему рівнянь МВ для кожної координати вектора-градієнта.

Теорема (про рівняння максимальної вірогідності)

Якщо параметр є векторним: $\Theta \subset R^d$, максимум в означенні ОМВ досягається всередині параметричної множини, а функція вірогідності диференційовна, то ОМВ задовольняє рівняння максимальної вірогідності:

$$\frac{\partial}{\partial \theta} \ln L(X, \theta)|_{\theta=\hat{\theta}} = 0,$$

тобто ОМВ є коренем функції впливу $U(X, \theta)$:

$$U(X, \hat{\theta}) = 0.$$

Якщо $\dim \theta = d > 1$, то рівняння МВ є векторним, і перетворюється на систему рівнянь МВ для кожної координати вектора-градієнта.

Співвідношення з ефективними оцінками. Інваріантність

Теорема (про співвідношення ОМВ з ефективними оцінками)

- 1 Якщо параметр θ – скалярний і існує його ефективна оцінка $\hat{\theta}^*$, то існує ОМВ $\hat{\theta}$, причому

$$\hat{\theta} = \hat{\theta}^*.$$

Теорема (про інваріантність оцінки максимальної вірогідності)

Нехай існує оцінка максимальної вірогідності $\hat{\theta}$ параметра θ , а функція $q : \Theta \rightarrow Q$ є взаємно-однозначною. Тоді оцінка $q(\hat{\theta})$ є оцінкою максимальної вірогідності для значення функції $q(\theta)$.

Приклади обчислення ОМВ

Схема Бернуллі

$$\arg \max \ln L(X, \theta) = \arg \max \ln \theta^{\nu_n(X)} (1 - \theta)^{n - \nu_n(X)}$$

$$= \arg \max (\nu_n(X) \ln \theta + (n - \nu_n(X)) \ln(1 - \theta))$$

Якщо прирівняти похідну до 0 $\frac{\partial \ln L(X, \theta)}{\partial \theta} = 0$, то ОМВ

$$\frac{\nu_n(X)}{n} = \hat{\theta}_n.$$

Приклади обчислення ОМВ

Схема Бернуллі

$$\arg \max \ln L(X, \theta) = \arg \max \ln \theta^{\nu_n(X)} (1 - \theta)^{n - \nu_n(X)}$$

$$= \arg \max (\nu_n(X) \ln \theta + (n - \nu_n(X)) \ln(1 - \theta))$$

Якщо прирівняти похідну до 0 $\frac{\partial \ln L(X, \theta)}{\partial \theta} = 0$, то ОМВ

$$\frac{\nu_n(X)}{n} = \hat{\theta}_n.$$

Приклади обчислення ОМВ

Пуассонівська вибірка

$$\ln L(X, \theta) = \ln \prod_{k=1}^n \frac{\theta^{\xi_k}}{\xi_k!} e^{-\theta}$$

$$= \ln \theta \sum_{k=1}^n \xi_k - n\theta + \ln h(X),$$

Якщо прирівняти похідну до 0 $\frac{\partial \ln L(X, \theta)}{\partial \theta} = 0$, то ОМВ

$$\hat{\theta}_n = \arg \max \ln L(X, \theta) = \frac{1}{n} \sum_{k=1}^n \xi_k = \hat{\mu}_n.$$

Приклади обчислення ОМВ

Пуассонівська вибірка

$$\begin{aligned}\ln L(X, \theta) &= \ln \prod_{k=1}^n \frac{\theta^{\xi_k}}{\xi_k!} e^{-\theta} \\ &= \ln \theta \sum_{k=1}^n \xi_k - n\theta + \ln h(X),\end{aligned}$$

Якщо прирівняти похідну до 0 $\frac{\partial \ln L(X, \theta)}{\partial \theta} = 0$, то ОМВ

$$\hat{\theta}_n = \arg \max \ln L(X, \theta) = \frac{1}{n} \sum_{k=1}^n \xi_k = \hat{\mu}_n.$$

Приклади обчислення ОМВ

Пуассонівська вибірка

$$\begin{aligned}\ln L(X, \theta) &= \ln \prod_{k=1}^n \frac{\theta^{\xi_k}}{\xi_k!} e^{-\theta} \\ &= \ln \theta \sum_{k=1}^n \xi_k - n\theta + \ln h(X),\end{aligned}$$

Якщо прирівняти похідну до 0 $\frac{\partial \ln L(X, \theta)}{\partial \theta} = 0$, то ОМВ

$$\hat{\theta}_n = \arg \max \ln L(X, \theta) = \frac{1}{n} \sum_{k=1}^n \xi_k = \hat{\mu}_n.$$

Приклади обчислення ОМВ

Показникова вибірка (зміщеність ОМВ)

$$\ln L(X, \theta) = \ln \prod_{k=1}^n \theta \exp(-\theta \xi_k) = n \ln \theta - \theta \sum_{k=1}^n \xi_k,$$

$$\hat{\theta}_n = \arg \max \ln L(X, \theta) = \frac{n}{\sum_{k=1}^n \xi_k} = \frac{1}{\hat{\mu}_n}.$$

Оскільки $\sum_{k=1}^n \xi_k$ має розподіл Ерланга з параметрами n, θ , то

$$M_{\theta} \hat{\theta}_n = \int_0^{\infty} \frac{n}{x} \theta \frac{(\theta x)^{n-1}}{(n-1)!} \exp(-\theta x) dx = \theta \frac{n}{n-1} \neq \theta \quad \forall n \geq 1 \forall \theta > 0.$$

Отже, загалом не можна розраховувати на незсунутість ОМВ.

Приклади обчислення ОМВ

Показникова вибірка (зміщеність ОМВ)

$$\ln L(X, \theta) = \ln \prod_{k=1}^n \theta \exp(-\theta \xi_k) = n \ln \theta - \theta \sum_{k=1}^n \xi_k,$$

$$\hat{\theta}_n = \arg \max \ln L(X, \theta) = \frac{n}{\sum_{k=1}^n \xi_k} = \frac{1}{\hat{\mu}_n}.$$

Оскільки $\sum_{k=1}^n \xi_k$ має розподіл Ерланга з параметрами n, θ , то

$$M_{\theta} \hat{\theta}_n = \int_0^{\infty} \frac{n}{x} \theta \frac{(\theta x)^{n-1}}{(n-1)!} \exp(-\theta x) dx = \theta \frac{n}{n-1} \neq \theta \quad \forall n \geq 1 \forall \theta > 0.$$

Отже, загалом не можна розраховувати на незсунутість ОМВ.

Консистентність ОМВ

Теорема (про консистентність ОМВ)

За певних умов оцінка максимальної вірогідності $\hat{\theta}_n$ параметра θ за кратною вибіркою X існує і є консистентною.

Асимптотична нормальність і ефективність ОМВ

У даному розділі припустимо, що $\Theta \subset \mathbb{R}^1$

Теорема (про асимптотичну нормальність ОМВ)

Нехай $X = (\xi_1, \dots, \xi_n)$ – кратна вибірка, щільність одного спостереження $f(y, \theta)$ задовольняє умови регулярності та тричі неперервно диференційовна за θ , причому відповідні похідні мажоруються за модулем інтегровною величиною η :

$$\left| \frac{\partial^{(i)} \ln f(\xi_1, \theta)}{\partial^{(i)} \theta} \right| \leq \eta, i = \overline{1, 3}, \quad \mathbf{M}_{\theta_0} \eta < \infty.$$

Теорема (продовження)

Якщо ОМВ $\hat{\theta}_n$ невідомого параметра θ_0 є розв'язком рівняння максимальної вірогідності, то ця оцінка є асимптотично нормальною:

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{W} \zeta \cong N(0, 1/i(\theta_0)), \quad n \rightarrow \infty,$$

де асимптотична дисперсія визначається кількістю інформації за Фішером, що міститься в одному спостереженні:

$$i(\theta) \equiv \mathbf{M}_{\theta} \left(\frac{\partial}{\partial \theta} \ln f(\xi_1, \theta) \right)^2$$

Зауваження

За теоремою про адитивність інформації за Фішером повна інформація за Фішером у вибірці дорівнює

$$I_n(\theta) = ni(\theta),$$

тому з даної теореми випливає асимптотична ефективність ОМВ, оскільки

$$\frac{1}{ni(\theta_0)} = \frac{1}{I_n(\theta_0)}$$

збігається з найменшою можливою межею для дисперсій незміщених оцінок параметра θ в теоремі про нерівність та критерій Крамера – Рао.

ПИТАННЯ?