Теоретичні та вибіркові моменти. Метод моментів.

1 Теоретичні відомості

It's dangerous to go alone! Take this.

Нагадаємо дві теореми з курсу теорії ймовірностей, пов'язані з властивостями збіжних випадкових послідовностей. Нехай "$\to$" позначає збіжність за ймовірністю (або майже напевно).

Теорема (Про арифметичні дії над збіжними послідовностями). Нехай $\{\xi_n\}_{n\geq 1}$ та $\{\eta_n\}_{n\geq 1}$ – такі випадкові послідовності, що $\xi_n \to \xi$, $\eta_n \to \eta$ при $n \to \infty$. Тоді

1. $\xi_n + \eta_n \to \xi + \eta$,
2. $\xi_n \cdot \eta_n \to \xi \cdot \eta$. Зокрема, якщо $\eta \neq 0$, то $\xi_n/\eta_n \to \xi/\eta$.

Теорема (Неперервності). Нехай $\{\xi_n\}_{n\geq 1}$ – послідовність випадкових векторів, що $\xi_n \to \xi$ при $n \to \infty$. Розглянемо $f(x)$ – непреревна функція на множині значень $\{\xi_n\}_{n\geq 1}, \xi$. Тоді $f(\xi_n) \to f(\xi)$ при $n \to \infty$.

Приклад. Нехай $\{X_k\}_{k\geq 1}$ – послідовність н.о.р. випадкових величин, причому $\mathbf{E}\left[|X_1|\right] < \infty$. Покладемо $a = \mathbf{E}\left[X_1\right]$. Для заданої послідовності виконується ПЗВЧ, $S_n = \sum_{k=1}^n X_k/n \to^{P_1} a$. Розглянемо, наприклад, $f(x) = \sin(x)$, $x \in \mathbb{R}$. Тоді $f(S_n) \to^{P_1} f(a) = \sin(a)$.

Оцінювання параметрів розподілу методом моментів.

Розглянемо кратну вибірку $X = (X_1, \ldots, X_n)$, де розподіл спостережень $F(t;\theta) = \mathbf{P}_\theta\left(X_1 < t\right)$ є відомим з точністю до невідомих параметрів $\theta = (\theta_1, \ldots, \theta_d) \in \Theta \subset \mathbb{R}^d$. Невідомі параметри розподілу треба оцінити.

Побудова оцінок методом моментів полягає в наступному. Нехай $h(x) = (h_1(x), \ldots, h_d(x))^T$ – така векторна функція, що $\mathbf{E}_\theta\left[h(X_1)\right] < \infty$, причому $\mathbf{E}_\theta\left[h(X_1)\right] = H(\theta)$ містить усі невідомі параметри, що потрібно оцінити. За законом великих чисел,

$$\hat{h}_n := \frac{1}{n}\sum_{j=1}^n h(X_j) \to \mathbf{E}_\theta\left[h(X_1)\right] = H(\theta),\ n \to \infty.$$

Тобто вибіркові моменти збігаються (за ймовірністю; майже напевно) до теоретичних моментів при збільшенні обсягу вибірки. Можна сподіватися, що для великих n, $\hat{h}_n \approx H(\theta)$. Прирівнюючи вибіркові моменти з теоретичними, отримаємо систему моментних рівнянь

$$\hat{h}_n = H(\theta), \tag{1}$$

яку потрібно розв'язати відносно $\theta \in \Theta$. Розв'язки системи рівнянь (1) називають оцінками методом моментів (які ще називають моментними оцінками).

2 Задачі

2.1 Задача 1

Нехай $X = (X_1, \ldots, X_n)$ – кратна вибірка з пуассонівського розподілу спостережень, тобто $X_j \sim \mathrm{Pois}(\theta)$, де $\theta > 0$ вважається невідомим. Побудуйте оцінку методом моментів для невідомого параметра. Дослідити оцінку на незміщеність, консистентність.

Розв'язання.

У розподілі спостережень лише один невідомий параметр, який треба оцінити. Тому достатньо побудувати моментне рівняння, звідки знайдемо моментну оцінку. Для побудови моментного рівняння підберемо таку функцію $h(x)$, щоб $\mathbf{E}_\theta\left[h(X_1)\right] = H(\theta)$ містило θ. Візьмемо, наприклад, $h(x) = x$. Тоді

$$\mathbf{E}_\theta\left[h(X_1)\right] = \mathbf{E}_\theta\left[X_1\right] = \sum_{k=0}^{+\infty} k\mathbf{P}_\theta\left(X_1 = k\right) = \sum_{k=0}^{+\infty} k\frac{\theta^k}{k!}e^{-\theta} = \sum_{k=1}^{+\infty} k\frac{\theta^k}{k!}e^{-\theta} =$$

$$= \theta e^{-\theta}\sum_{k=1}^{+\infty}\frac{\theta^{k-1}}{(k-1)!} = \left| e^\theta = \sum_{k=1}^{+\infty}\frac{\theta^{k-1}}{(k-1)!}\right| = \theta =: H(\theta).$$

Отже при заданій функції $h(x) = x$ маємо, що $\mathbf{E}_\theta\left[h(X_1)\right]$ містить θ. Будуємо моментне рівняння, прирівнявши теоретичний момент з вибірковим та розв'язуємо відносно θ:

$$H(\theta) = \frac{1}{n}\sum_{j=1}^{n} h(X_j) \Leftrightarrow \theta = \frac{1}{n}\sum_{j=1}^{n} X_j$$

Розв'язком рівняння є вибіркове середнє, $\hat{\theta}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$. Оцінка, очевидно, незміщена (випливає з властивостей математичного сподівання та підрахунків вище). Очевидною є (строга) консистентність, бо за критерієм Колмогорова про ПЗВЧ (який має місце, бо $\mathbf{E}_\theta\left[X_1\right] < \infty$) $\hat{\theta}_n \to^{P_1} \theta$ при $n \to +\infty$.

Для допитливих. Для отримання моментної оцінки достатньо підібрати моментну функцію $h(x)$ так, щоби $\mathbf{E}_\theta\left[h(X_1)\right]$ містило θ та рівняння $H(\theta) = \frac{1}{n}\sum_{j=1}^{n} h(X_j)$ було розв'язним відносно θ. Значить, форма моментної оцінки залежить від того яку моментну функцію $h(x)$ ми підберемо (відповідно й властивості оцінки).

Можна взяти $h(x) = \mathbf{1}\{x = 0\}$. Тоді $\mathbf{E}_\theta\left[h(X_1)\right] = \mathbf{E}_\theta\left[\mathbf{1}\{X_1 = 0\}\right] = \mathbf{P}_\theta\left(X_1 = 0\right) = e^{-\theta}$, $\frac{1}{n}\sum_{j=1}^{n} h(X_j) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}\{X_j = 0\}$. Звідси $\hat{\theta}^*_n = -\ln(\frac{1}{n}\sum_{j=1}^{n}\mathbf{1}\{X_j = 0\})$ – моментна оцінка для θ.

А можна взяти $h(x) = x^2$: $\mathbf{E}_\theta\left[h(X_1)\right] = \mathbf{E}_\theta\left[(X_1)^2\right] = \theta^2 + \theta$, $\frac{1}{n}\sum_{j=1}^{n} h(X_j) = \frac{1}{n}\sum_{j=1}^{n}(X_j)^2$. Тоді $\theta^2 + \theta - \frac{1}{n}\sum_{j=1}^{n}(X_j)^2 = 0$, звідки беремо додатний розв'язок: $\hat{\theta}^{**}_n = \frac{1}{2}(\sqrt{\frac{4}{n}\sum_{j=1}^{n}(X_j)^2 + 1} - 1)$.

Тому вибір, справді кажучи, широкий. Бажано підбирати таку моментну функцію, щоб моментна оцінка була в певному сенсі *кращою* (наприклад, має хороші статистичні властивості; на практиці її можна 'спокійно' обчислити тощо). Серед $\hat{\theta}_n, \hat{\theta}^*_n, \hat{\theta}^{**}_n$ кращою буде $\hat{\theta}_n$ (подумати).

2.2 Задача 2

Нехай $X = (X_1, \ldots, X_n)$ – кратна вибірка з експоненційним розподілом спостережень, тобто $X_j \sim \text{Exp}(\theta)$, де $\theta > 0$ вважається невідомим. Побудуйте оцінку методом моментів для невідомого параметра. Дослідити оцінку на незміщеність, консистентність.

Розв'язання.

В цьому випадку теж один невідомий параметр для оцінювання, тому треба побудувати одне моментне рівняння. В якості моментної функції оберемо $h(x) = x$, бо

$$\mathbf{E}_\theta\left[h(X_1)\right] = \mathbf{E}_\theta\left[X_1\right] = \int\limits_0^\infty t\theta e^{-\theta t}dt = \theta^{-1}\int\limits_0^\infty (\theta t)e^{-(\theta t)}d(\theta t) = \theta^{-1}\int\limits_0^\infty ue^{-u}du = \theta^{-1}\Gamma(2) = \theta^{-1} = H(\theta).$$

Будуємо моментне рівняння, прирівнявши теоретичний момент з вибірковим та розв'язуємо відносно θ:

$$H(\theta) = \frac{1}{n}\sum_{j=1}^n h(X_j) \Leftrightarrow \theta^{-1} = \frac{1}{n}\sum_{j=1}^n X_j$$

Розв'язком рівняння є $\hat{\theta}_n = 1/(\frac{1}{n}\sum_{j=1}^n X_j) = n/\sum_{j=1}^n X_j$.

Перевіримо, чи є $\hat{\theta}_n$ незміщеною. Для цього зауважимо, що $S_n = \sum_{j=1}^n X_j \sim \Gamma(n, \theta)$ (побачте чому). Тоді

$$\mathbf{E}_\theta\left[\hat{\theta}_n\right] = n\mathbf{E}_\theta\left[S_n^{-1}\right] = n\int\limits_0^\infty t^{-1}\frac{\theta^n}{\Gamma(n)}t^{n-1}e^{-\theta t}dt = \frac{n}{\Gamma(n)}\cdot\theta\int\limits_0^\infty (\theta t)^{n-2}e^{-(\theta t)}d(\theta t) =$$

$$= \frac{n}{\Gamma(n)}\cdot\theta\int\limits_0^\infty u^{(n-1)-1}e^{-u}du = \frac{n}{\Gamma(n)}\cdot\theta\cdot\Gamma(n-1) = \frac{n(n-2)!}{(n-1)!}\cdot\theta = \frac{n}{n-1}\cdot\theta \to \theta,\ n\to\infty.$$

Отже $\hat{\theta}_n$ є зміщеною оцінкою θ, але буде асимптотично неміщеною для θ.

Далі, оскільки критерій Колмогорова ПЗВЧ виконується (бо перший теоретичний момент розподілу спостережень інтегровний), то внаслідок ПЗВЧ $\frac{1}{n}\sum_{j=1}^n X_j \to^{P_1} \theta^{-1} > 0$, а тому й $\hat{\theta}_n \to^{P_1} \theta$ при $n \to \infty$ (арифметичні дії над збіжними з імовірністю 1 послідовностями). Отже моментна оцінка $\hat{\theta}_n$ є строго консистентною оцінкою θ.

2.3 Задача 3

Нехай $X = (X_1, \dots, X_n)$ – кратна вибірка з нормальним розподілом спостережень, тобто $X_j \sim N(\mu, \sigma^2)$, де обидва параметри $\theta = (\mu, \sigma^2) \in \mathbb{R} \times (0, +\infty)$ вважаються невідомими. Побудуйте оцінки методом моментів для невідомих параметрів. Дослідити оцінки на незміщеність, консистентність.

Розв'язання.

У заданій моделі даних маємо два невідомих параметри для оцінювання, тобто μ і σ^2. Тому потрібно побудувати систему з (принаймні) двух моментних рівнянь та розв'язати її відносно μ і σ^2.

Оберемо такі моментні функції: $h_1(x) = x$ та $h_2(x) = x^2$. Тоді

$$\mathbf{E}_\theta\left[h_1(X_1)\right] = \mathbf{E}_\theta\left[X_1\right] = \mu \text{ та } \mathbf{E}_\theta\left[h_2(X_1)\right] = \mathbf{E}_\theta\left[X_1^2\right] = \mu^2 + \sigma^2$$

Отже перші два теоретичні моменти містять потрібні невідомі параметри. На їх основі будуємо моментні рівняння та, відповідно, систему

$$\begin{cases} \mathbf{E}_\theta\left[h_1(X_1)\right] = \frac{1}{n}\sum_{j=1}^n h_1(X_j) \\ \mathbf{E}_\theta\left[h_2(X_1)\right] = \frac{1}{n}\sum_{j=1}^n h_2(X_j) \end{cases} \Leftrightarrow \begin{cases} \mu = \frac{1}{n}\sum_{j=1}^n X_j \\ \mu^2 + \sigma^2 = \frac{1}{n}\sum_{j=1}^n X_j^2 \end{cases}$$

Розв'язком рівняння є пара $(\hat{\mu}_n, \hat{\sigma}_n^2) = (\sum_{j=1}^n X_j/n, \sum_{j=1}^n X_j^2/n - (\sum_{j=1}^n X_j/n)^2)$, тобто вибіркове середнє та дисперсія відповідно.

Консистентність оцінок випливає з виконання ПЗВЧ та властивістю арифметичних дій над збіжними з імовірністю 1 послідовностями:

$$\sum_{j=1}^n X_j/n \to^{P_1} \mu,\ \sum_{j=1}^n X_j^2/n \to^{P_1} \mu^2 + \sigma^2 \Rightarrow \hat{\mu}_n \to^{P_1} \mu,\ \hat{\sigma}_n^2 \to^{P_1} (\mu^2 + \sigma^2) - \mu^2 = \sigma^2,\ n \to \infty.$$

Очевидно, що $\hat{\mu}_n$ – незміщена оцінка для μ, але $\hat{\sigma}_n^2$ – зміщена для σ^2:

$$\mathbf{E}_\theta\left[\hat{\sigma}_n^2\right] = (\mu^2 + \sigma^2) - \mathbf{E}_\theta\left[\left(\sum_{j=1}^n X_j/n\right)^2\right] = \dots$$

$$\mathbf{E}_\theta\left[\left(\sum_{j=1}^n X_j/n\right)^2\right] = \frac{\sigma^2}{n} + \mu^2, \text{ бо } \sum_{j=1}^n X_j/n \sim N(\mu, \sigma^2/n)$$

$$\dots = \sigma^2\left(1 - \frac{1}{n}\right) \to \sigma^2,\ n \to \infty.$$

Але $\hat{\sigma}_n^2$ – асимптотично незміщена для σ^2.

2.4 Задача 4

Нехай $X = (X_1, \dots, X_n)$ – кратна вибірка, де спостереження мають розподіл Лапласа, тобто

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

де $\theta \in \mathbb{R}$ вважається невідомим. Побудуйте оцінку методом моментів для невідомого параметра θ. Дослідити оцінку на незміщеність, консистентність.

Розв'язання.

В цьому випадку теж один невідомий параметр для оцінювання, тому треба побудувати одне моментне рівняння. В якості моментної функції оберемо $h(x) = x$. Тоді

$$\begin{aligned}
\mathbf{E}_\theta\left[h(X_1)\right] = \mathbf{E}_\theta\left[X_1\right] &= \frac{1}{2} \cdot \int\limits_{-\infty}^{+\infty} t \exp(-|t-\theta|)dt = \\
&= \frac{1}{2} \cdot \left(\int\limits_{\theta}^{+\infty} t \exp(-(t-\theta))dt + \int\limits_{-\infty}^{\theta} t \exp(-(\theta-t))dt \right) = \\
&= \frac{1}{2} \cdot \left(\int\limits_{0}^{+\infty} (\theta+u) \exp(-u)du + \int\limits_{0}^{+\infty} (\theta-u) \exp(-u)du \right) = \\
&= \theta \int\limits_{0}^{+\infty} \exp(-u)du + \frac{1}{2} \int\limits_{0}^{+\infty} u \exp(-u)du \cdot (1-1) = \theta + 1 \cdot (1-1) = \theta.
\end{aligned}$$

Будуємо моментне рівняння, прирівнявши теоретичний момент з вибірковим та розв'язуємо відносно θ:

$$H(\theta) = \frac{1}{n} \sum_{j=1}^{n} h(X_j) \Leftrightarrow \theta = \frac{1}{n} \sum_{j=1}^{n} X_j$$

Розв'язком рівняння є $\hat{\theta}_n = \frac{1}{n} \sum_{j=1}^{n} X_j$. Оцінка, очевидно, незміщена та строго консистентна для параметра θ.

2.5 Задача 5

Оцінити значення інтеграла $I = \int\limits_0^2 y(1+[y])^{-1}dy$, де $[y]$ – ціла частина y, за вибіркою:

$$1.82, 0.56, 1.63, 1.46, 0.09, 1.19, 1.88, 1.06, 1.07, 0.67, 0.50, 1.89, 1.33, 1.86, 0.69, 1.27$$

Знайти дисперсію використаної оцінки для I.

Розв'язання.

Зауважимо, що I можна виразити через математичне сподівання деякої функції від спостереження з вибірки:

$$I = \int\limits_0^2 y(1+[y])^{-1}dy = \int\limits_0^2 2y(1+[y])^{-1}\frac{1}{2}dy = \int\limits_0^2 f(y)\frac{1}{2}dy = \mathbf{E}\left[f(X_1)\right],\ f(y) = 2y(1+[y])^{-1}.$$

Очевидно, що $I < \infty$:

$$I = \int\limits_0^2 y(1+[y])^{-1}dy = \int\limits_0^1 ydy + \int\limits_1^2 y/2dy = \frac{1}{2} + \frac{1}{4}(2^2 - 1^2) = \frac{2}{4} + \frac{3}{4} = \frac{5}{4}$$

Отже, випадкові величини $\{f(X_j)\}_{j=1}^n$ є незалежними в сукупності та інтегровними. Має місце ПЗВЧ:

$$\hat{I}_n := \frac{1}{n}\sum_{j=1}^n f(X_j) \to^{P_1} \mathbf{E}\left[f(X_1)\right] = I,\ n \to \infty.$$

В якості оцінки для I беремо вибіркове середнє перетворення величин, тобто $\hat{I}_n$. Підрахуємо це значення за вибіркою (її реалізацією вище):

$$f(X) = (1.82, 1.12, 1.63, 1.46, 0.18, 1.19, 1.88, 1.06, 1.07, 1.34, 1, 1.89, 1.33, 1.86, 1.38, 1.27)$$

Отже $\hat{I}_n = 1.3425$.

Обчислимо $\mathbf{D}\left[\hat{I}_n\right] = \mathbf{E}\left[\hat{I}_n^2\right] - \left(\mathbf{E}\left[\hat{I}_n\right]\right)^2$:

$$\mathbf{E}\left[f(X_1)^2\right] = \int\limits_0^2 (f(y))^2\frac{1}{2}dy = \int\limits_0^2 2y^2(1+[y])^{-2}dy =$$

$$= \int\limits_0^1 2y^2dy + \int\limits_1^2 \frac{1}{2}y^2dy = \frac{2}{3} + \frac{7}{6} = \frac{11}{6}$$

$$\Rightarrow \mathbf{D}\left[\hat{I}_n\right] = \frac{11}{6} - \left(\frac{5}{4}\right)^2 = \frac{11}{6} - \frac{25}{16} = \frac{88}{48} - \frac{75}{48} = \frac{3}{88}$$