Упражнение 11. Точкови оценки. Метод на моментите. Свойства на точковите оценки.

Методът на моментите се използва за намиране на точкови оценки. Той се състои в следното. Нека плътността на наблюдаваната сл. величина X зависи от k параметъра $f(x;\theta) = f_X(x;\theta_1,\theta_2,\dots,\theta_k)$. Нека предположим, че моментите до ред k на сл. величина са крайни и се изразяват, чрез параметрите по следния начин:

$$
\begin{aligned}
&E[X] = \mu_1(\theta_1,\theta_2,\dots,\theta_k)\\
&E[X^2] = \mu_2(\theta_1,\theta_2,\dots,\theta_k)\\
&\dots\dots\\
&E[X^k] = \mu_k(\theta_1,\theta_2,\dots,\theta_k)
\end{aligned}
$$

Нека извадката се състои от N наблюдения на X: $(x_1,x_2,\dots,x_N)$. Определяме емпиричните моменти

$$
\begin{aligned}
&\bar{m}_1 = \frac{1}{N}\sum_{j=1}^{N} x_j\\
&\bar{m}_2 = \frac{1}{N}\sum_{j=1}^{N} x_j^2\\
&\dots\dots\\
&\bar{m}_k = \frac{1}{N}\sum_{j=1}^{N} x_j^k
\end{aligned}
$$

Решаваме системата

$$
\begin{aligned}
&\mu_1(\theta_1,\theta_2,\dots,\theta_k) = \bar{m}_1\\
&\mu_2(\theta_1,\theta_2,\dots,\theta_k) = \bar{m}_2\\
&\dots\dots\\
&\mu_k(\theta_1,\theta_2,\dots,\theta_k) = \bar{m}_k
\end{aligned}
$$

относно неизвестните параметри и намираме техните оценки. Ако трябва да намерим оценка за дадена функция $\tau(\theta_1,\theta_2,\dots,\theta_k)$, то в нея заместваме намерените оценки на параметрите.

Задачи

Зад. 1. *Нека $x_1, \dots, x_N$ са независими наблюдения над сл.в. $X \sim Ge(p)$. Да се намери оценка по метода на моментите за параметъра p.*

Решение:
Имаме $P(X = k) = q^{k-1}p, k = 1, 2, 3, \dots$ и вероятностната пораждаща функция има вида $f(s) = Es^X = \frac{ps}{1-qs}$.

За мат. очакване имаме $E[X] = f'(1) = \left.\dfrac{p(1-qs) + qps}{(1-qs)^2}\right|_{s=1} = \dfrac{1}{p}$.

Сега приравняваме $E[X] = \bar{X}_N$, откъдето имаме $\frac{1}{p} = \frac{1}{N}\sum_{i=1}^n x_i$. От където веднага следва, че $\hat{p} = \frac{N}{\sum_{i=1}^n x_i}$ е търсената оценка за неизвестния параметър p по метода на моментите.

Зад. 2. *Нека $x_1, \dots, x_N$ са независими наблюдения над сл.в. $X \sim U(\theta_1, \theta_1 + \theta_2)$. Да се намери оценка по метода на моментите за параметрите θ_1 и θ_2.*

Решение:
Плътността на X има вида $f(x) = \frac{1}{\theta_2}$ за $x \in [\theta_1, \theta_1 + \theta_2]$. Тогава

$$E[X] = \int_{\theta_1}^{\theta_1+\theta_2} \frac{x}{\theta_2}dx = \left.\frac{x^2}{2\theta_2}\right|_{\theta_1}^{\theta_1+\theta_2} = \frac{(\theta_1+\theta_2)^2 - \theta_1^2}{2\theta_2} = \theta_1 + \frac{\theta_2}{2}$$

$$E[X^2] = \int_{\theta_1}^{\theta_1+\theta_2} \frac{x^2}{\theta_2}dx = \left.\frac{x^3}{3\theta_2}\right|_{\theta_1}^{\theta_1+\theta_2} = \frac{(\theta_1+\theta_2)^3 - \theta_1^3}{3\theta_2} = \theta_1^2 + \theta_1\theta_2 + \frac{\theta_2^2}{3}.$$

Сега приравняваме

$$\theta_1 + \frac{\theta_2}{2} = \bar{m}_1 = \frac{1}{N}\sum_{i=1}^n x_i.$$

$$\theta_1^2 + \theta_1\theta_2 + \frac{\theta_2^2}{3} = \bar{m}_2 = \frac{1}{N}\sum_{i=1}^n x_i^2.$$

Решаваме системата

$$\theta_1 = \bar{m}_1 - \frac{\theta_2}{2}$$

$$(\bar{m}_1 - \frac{\theta_2}{2})^2 + (\bar{m}_1 - \frac{\theta_2}{2})\theta_2 + \frac{\theta_2^2}{3} = \bar{m}_2$$

$$\bar{m}_1^2 - \bar{m}_1\theta_2 + \frac{\theta_2^2}{4} + \bar{m}_1\theta_2 - \frac{\theta_2^2}{2} + \frac{\theta_2^2}{3} = \bar{m}_2$$

$$\frac{\theta_2^2}{12} = \bar{m}_2 - \bar{m}_1^2$$

$$\theta_2 = 2\sqrt{3}\sqrt{\bar{m}_2 - \bar{m}_1^2}$$

$$\theta_1 = \bar{m}_1 - \sqrt{3}\sqrt{\bar{m}_2 - \bar{m}_1^2}.$$

Неизместени оценки

Една точкова оценка (статистика) $t(X_1, \ldots, X_N)$ на неизвестния параметър θ,*(разгледана като функция на N независими еднакво разпределени сл. величини)*, се нарича неизместена, ако $E(t(X_1, \ldots, X_N)) = \theta$.

Забележка: Ясно е, че за конкретна извадка $x_1, \ldots, x_N$ статистиката има някаква конкретна стойност $t(x_1, \ldots, x_N)$. Също така е ясно, обаче, че ако извършим експеримента отново ще имаме други числа $y_1, \ldots, y_N$ и съответно друга стойност на статистиката $t(y_1, \ldots, y_N)$. В случая ние се интересуваме не от конкретната извадка, а по-скоро от свойствата на оценката като функция от N независими (защото са независимо избрани от генералната съвкупност), еднакво разпределени (защото са резултат от измерването на една и съща характеристика на генералната съвкупност) случайни величини. Поради тази причина от тук нататък ще бележим наблюденията с главни букви (X_i) вместо с малки (x_i).

Зад. 3. *Нека $X_1, ..., X_N$ са независими наблюдения над сл.в. X с $EX = a$ и $DX = \sigma^2$. Да се провери дали*

$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad \sigma_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)^2$$

са неизместени оценки за параметрите a и σ^2. Да се намерят съответните неизместени оценки.

Решение:
Да намерим математическото очакване на $\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$ като считаме, че X_i са независими, еднакво разпределени сл. величини с математическо очакване $EX_i = a$. Имаме

$$E\bar{X}_N = E[\frac{1}{N}\sum_{i=1}^{N} X_i] = \frac{1}{N}\sum_{i=1}^{N} E[X_i] = \frac{1}{N}\sum_{i=1}^{N} a = \frac{1}{N}Na = a.$$

Следователно $\bar{X}_N$ е неизместена оценка за EX_i.

По същия начин оценката за дисперсията ще е неизместена, ако $E[\sigma_N^2] = \sigma^2$.

За

$$\sigma_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}_N)^2$$

имаме

$$\sigma_N^2 = \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - 2\bar{X}_N \sum_{i=1}^N X_i + N\bar{X}_N^2\right]$$

$$= \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - 2N\bar{X}_N^2 + N\bar{X}_N^2\right] = \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - N\bar{X}_N^2\right]$$

$$= \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - N(\frac{1}{N}\sum_{i=1}^N X_i)^2\right] = \frac{1}{N^2}\left[N\sum_{i=1}^N X_i^2 - (\sum_{i=1}^N X_i)^2\right]$$

$$= \frac{1}{N^2}\left[N\sum_{i=1}^N X_i^2 - (\sum_{i=1}^N X_i^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^N X_iX_j)\right]$$

$$= \frac{1}{N^2}\left[(N-1)\sum_{i=1}^N X_i^2 - 2\sum_{i=1}^{N-1}\sum_{j=i+1}^N X_iX_j\right].$$

Следователно

$$E[\sigma_N^2] = \frac{1}{N^2}\left[(N-1)\sum_{i=1}^N E[X_i^2] - 2\sum_{i=1}^{N-1}\sum_{j=i+1}^N E[X_i]E[X_j]\right]$$

$$= \frac{1}{N^2}\left[(N-1)NE[X_i^2] - 2\frac{N^2-N}{2}(E[X_i])^2\right] = \frac{N(N-1)}{N^2}\left[E[X_i^2] - (E[X_i])^2\right] = \frac{N-1}{N}\sigma^2.$$

Следователно оценката σ_N^2 е изместена.

Състоятелни оценки

Една точкова оценка (статистика) $t(X_1, X_2, \dots, X_N)$ на параметъра θ (разгледана като функция на N независими еднакво разпределени сл. величини) се нарича състоятелна, ако клони към θ по вероятност, т.е.

$$P(|t(X_1, X_2, \dots, X_N) - \theta| > \varepsilon) \to 0, \quad N \to \infty.$$

Ако сходимостта е *почти сигурно*, то оценката се нарича силно състоятелна.

Зад. 4. *Нека $X_1, ..., X_N$ са независими наблюдения над сл.в. X с $EX = a$ и $DX = \sigma^2$. Да се провери дали*

$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^N X_i \qquad \sigma_N^2 = \frac{1}{N}\sum_{i=1}^N (X_i - \bar{X}_N)^2$$

са състоятелни оценки за параметрите a и σ^2.

Решение:
Разглеждаме $\bar{X}_N = \frac{1}{N}\sum_{j=1}^N X_j$. Знаем, че тази оценка е неизместена оценка за $E[X]$. Тогава по закона за големите числа имаме

$$\bar{X}_N = \frac{1}{N}\sum_{j=1}^N X_j \xrightarrow{p} E[X], \quad N \to \infty.$$

Видяхме в предната задача, че оценката σ_N^2 е изместена оценка за дисперсията, защото $E[\sigma_N^2] = \frac{N-1}{N}\sigma^2$. Сега имаме

$$\sigma_N^2 = \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - 2\bar{X}_N \sum_{i=1}^N X_i + N\bar{X}_N^2\right]$$

$$= \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - 2N\bar{X}_N^2 + N\bar{X}_N)^2\right] = \frac{1}{N}\left[\sum_{i=1}^N X_i^2 - N\bar{X}_N^2\right] = \frac{1}{N}\sum_{i=1}^N X_i^2 - \bar{X}_N^2.$$

Сега по закона за големите числа имаме

$$\frac{1}{N}\sum_{i=1}^N X_i^2 \xrightarrow{p} E[X^2].$$

Освен това

$$\bar{X}_N^2 \xrightarrow{p} (E[X])^2.$$

Следователно

$$\sigma_N^2 \to E[X^2] - (E[X])^2 = \sigma^2.$$

Следователно тази оценка също е състоятелна макар, че е изместена.

Зад. 5. *Нека $X_1, \dots, X_N$ са независими наблюдения над сл.в. $X \sim U(0, \beta)$. Да се провери дали $t(X_1, \dots, X_N) = \max(X_1, \dots, X_N)$ сьстоятелна оценка за параметьра β.*

Решение:
Нека да вземем $\varepsilon > 0$. Ясно е, че $P(X_1 < \beta - \varepsilon) = \frac{\beta-\varepsilon}{\beta} < 1$. Сега да разгледаме

$$P(|\max(X_1, \dots, X_N) - \beta| > \varepsilon)$$

$$= P(\beta - \max(X_1, \dots, X_N) > \varepsilon)$$

$$= P(\max(X_1, \dots, X_N) < \beta - \varepsilon)$$

$$= P(X_1 < \beta - \varepsilon, \dots, X_N < \beta - \varepsilon)$$

$$= P(X_1 < \beta - \varepsilon) \dots P(X_N < \beta - \varepsilon)$$

$$= (P(X_1 < \beta - \varepsilon))^N = \left(\frac{\beta - \varepsilon}{\beta}\right)^N \to 0, \quad N \to \infty.$$

С това състоятелността на оценката е доказана.

Ефективни оценки

Нека θ е едномерен параметър и $\tau(\theta)$ е диференцируема функция на θ. Една оценка $t(x_1, \ldots, x_N)$ на $\tau(\theta)$ се нарича ефективна, ако е неизместена и е с минимална дисперсия. Дисперсията на оценката в този случай е ограничена отдолу по неравенството на Рао-Крамер,

$$D_\theta(t(X_1, X_2, \ldots, X_N)) \geq \frac{(\tau'(\theta))^2}{E_\theta\left[\frac{\partial}{\partial\theta}\log L(X_1, X_2, \ldots, X_N|\theta)\right]^2} \quad \text{почти сигурно.}$$

Равенство се достига, само когато е в сила представянето за функцията на правдоподобие:

$$\frac{\partial}{\partial\theta}\log L(x_1, x_2, \ldots, x_N|\theta) = G(\theta, N)\left[t(x_1, \ldots, x_N) - \tau(\theta)\right]. \tag{1}$$

Ще използваме това условие в следващите две задачи.

Зад. 6. *Над случайната величина $X \sim Po(\lambda)$ са проведени N независими наблюдения. Да се намери ефективна оценка за параметъра λ.*

Решение:
Знаем, че за Поасоново разпределена случайна величина $X \sim Po(\lambda)$, $E[X] = \lambda$. Да проверим дали максимално правдоподобната оценка

$$\hat{\lambda} = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

за която знаем, че е неизместена, е ефективна. Нека да вземем разпределението

$$P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}, x = 0, 1, 2, 3 \ldots.$$

Функцията на правдоподобие е

$$L(x_1, x_2, \ldots, x_N|\lambda) = \frac{\lambda^{x_1+x_2+\ldots+x_N}}{(x_1!)(x_2!)\ldots(x_N!)}e^{-N\lambda}.$$

За логаритъма имаме

$$\log L(x_1, x_2, \ldots, x_N|\lambda) = (x_1 + x_2 + \ldots + x_N)\log\lambda - \log((x_1!)(x_2!)\ldots(x_N!)) - N\lambda.$$

Сега

$$\frac{\partial}{\partial\lambda}\log L(x_1, x_2, \ldots, x_N|\lambda) = \frac{1}{\lambda}\times(x_1+x_2+\ldots+x_N) - N = \frac{N}{\lambda}\left[\frac{x_1 + x_2 + \ldots + x_N}{N} - \lambda\right],$$

откъдето се вижда, че условието (1) е изпълнено с $G(\lambda, N) = \frac{N}{\lambda}$. С това е доказано, че оценката е ефективна.

Зад. 7. *Над случайната величина $X \sim N(0, \sigma^2)$ са проведени N н. н. Да се намери ефективна оценка за параметъра σ.*

Решение:
Да образуваме функцията на правдоподобие:

$$L(x_1, x_2, \ldots, x_N|\sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^N} \mathrm{e}^{-\frac{x_1^2 + x_2^2 + \ldots + x_N^2}{2\sigma^2}}.$$

За логаритъма имаме

$$\log L(x_1, x_2, \ldots, x_N|\sigma^2) = -\frac{N}{2}(\log(2\pi) + \log\sigma^2) - \frac{x_1^2 + x_2^2 + \ldots + x_N^2}{2\sigma^2}.$$

Като диференцираме спрямо σ^2 получаваме

$$\frac{\partial}{\partial\sigma^2} \log L(x_1, x_2, \ldots, x_N|\sigma^2) = -\frac{N}{2}\frac{1}{\sigma^2} + \frac{x_1^2 + x_2^2 + \ldots + x_N^2}{2\sigma^4}$$

$$= \frac{N}{2\sigma^4}\left(\frac{x_1^2 + x_2^2 + \ldots + x_N^2}{N} - \sigma^2\right).$$

Така получаваме, че оценката

$$\sigma_N^2 = \frac{x_1^2 + x_2^2 + \ldots + x_N^2}{N}$$

е с минимална дисперсия, но както знаем тя е изместена и следователно не е ефективна. От друга страна оценката

$$s_N^2 = \frac{x_1^2 + x_2^2 + \ldots + x_N^2}{N - 1}$$

знаем, че е неизместена, но пък нейната дисперсия е по- голяма от дисперсията на σ_N^2. Това следва от $s_N^2 = \frac{N}{N-1}\sigma_N^2$ и тогава $Ds_N^2 = \frac{N^2}{(N-1)^2} D\sigma_N^2 > D\sigma_N^2$. Освен това логаритъма на функцията на правдоподобие не може да се изрази във вида (1) за $t(x_1, \ldots, x_N) = s_N^2$. Следователно нямаме ефективна оценка за дисперсията на нормалното разпределение.

Допълнителни задачи

Зад. 1'. *Нека $X_1, \ldots, X_N$ са независими наблюдения над сл.в. $X \sim Exp(\lambda)$. Да се намери оценка по метода на моментите за параметъра λ. Да се провери дали оценката е неизместена, състоятелна и ефективна.*

Зад. 2'. *Нека $X_1, \ldots, X_N$ са независими наблюдения над сл.в. $X \sim Po(\lambda)$. Да се намери оценка по метода на моментите за параметъра λ. Да се провери дали оценката е неизместена, състоятелна и ефективна.*

Зад. 3'. *Нека* $X_1, \dots, X_N$ *са независими наблюдения над сл.в.* $X \sim Po(\lambda)$. *Да се намери оценка по*
a) метода на моментите;
b) метода на максималното правдоподобие
за параметъра λ. *Да се провери дали оценката е неизместена, състоятелна и ефективна.*

Зад. 4'. *Нека* $X_1, \dots, X_N$ *са независими наблюдения над сл.в.* $X \sim Bi(N, p)$. *Да се намери ефективна оценка за параметъра* p.

Зад. 5'. *Нека* $X_1, \dots, X_N$ *са независими наблюдения над сл.в.* $X \sim \Gamma(\alpha, \beta)$, *където параметърът* β *е неизвестен и* $\alpha > 0$ *е известно. Да се намери ефективна оценка* $\hat{\beta}$ *на* β.
Плътносттна на X *има вида*

$$f_X(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, & x \geq 0. \end{cases}$$

Тук $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ *е Гама функцията. Имаме, че* $E[X] = \alpha\beta$ *и* $D[X] = \alpha\beta^2$ *(Покажете).*