Estymacja parametrów.

Krzysztof Topolski

Wykład 6 *Estymatory największej wiarogodności II*

Wrocław, 21 listopada 2019

Przypomnijmy na przykładzie sposób konstrukcji estymatora metodą największej wiarogodności.

Przykład.

Przyjmijmy, że dane $x_1 = 0,98$, $x_1 = 1,57$ i $x_1 = 0,31$, stanowią realizację próby losowej z rozkładu jednostajnego $U(0, \theta)$ o nieznanej wartości parametru $\theta > 0$.

Funkcja gęstości dla każdej obserwacji X_i ma postać $f(x \,|\, \theta) = 0$ dla $x \notin [0, \theta]$ oraz

$$f(x \,|\, \theta) = \frac{1}{\theta}, \quad \text{dla} \quad \leq x \leq \theta.$$

Estymator największej wiarogodności

Przypomnijmy na przykładzie sposób konstrukcji estymatora metodą największej wiarogodności.

Przykład.

Przyjmijmy, że dane $x_1 = 0,98$, $x_1 = 1,57$ i $x_1 = 0,31$, stanowią realizację próby losowej z rozkładu jednostajnego $U(0,\theta)$ o nieznanej wartości parametru $\theta > 0$.

Funkcja gęstości dla każdej obserwacji X_i ma postać $f(x\,|\,\theta) = 0$ dla $x \notin [0, \theta]$ oraz

$$f(x\,|\,\theta) = \frac{1}{\theta}, \quad \text{dla} \quad \le x \le \theta.$$

Estymator największej wiarogodności

Przypomnijmy na przykładzie sposób konstrukcji estymatora metodą największej wiarogodności.

Przykład.

Przyjmijmy, że dane $x_1 = 0,98$, $x_1 = 1,57$ i $x_1 = 0,31$, stanowią realizację próby losowej z rozkładu jednostajnego $U(0,\theta)$ o nieznanej wartości parametru $\theta > 0$.

Funkcja gęstości dla każdej obserwacji X_i ma postać $f(x\,|\,\theta) = 0$ dla $x \notin [0,\theta]$ oraz

$$f(x\,|\,\theta) = \frac{1}{\theta}, \quad \text{dla} \quad \leq x \leq \theta.$$

Przypomnijmy na przykładzie sposób konstrukcji estymatora metodą największej wiarogodności.

Przykład.

Przyjmijmy, że dane $x_1 = 0,98$, $x_1 = 1,57$ i $x_1 = 0,31$, stanowią realizację próby losowej z rozkładu jednostajnego $U(0,\theta)$ o nieznanej wartości parametru $\theta > 0$.

Funkcja gęstości dla każdej obserwacji X_i ma postać $f(x\,|\,\theta) = 0$ dla $x \notin [0,\ \theta]$ oraz

$$f(x\,|\,\theta) = \frac{1}{\theta}, \quad \text{dla} \quad \leq x \leq \theta.$$

Przykład cd.

Zauważmy, że gdy θ jest mniejsza od choć jedna z wartości x_i to funkcja wiarogodności $L(\theta \,|\, x_1, x_2, x_3)$ jest równa zero oraz jest równa $\frac{1}{\theta^3}$ gdy θ jest większa od wszystkich wartości x_i.

Tak więc ponieważ

$$L(\theta \,|\, x_1, x_2, x_3) = f(x_1 \,|\, \theta) f(x_2 \,|\, \theta) f(x_3 \,|\, \theta)$$

otrzymujemy

$$L(\theta \,|\, x_1, x_2, x_3) = \begin{cases} \frac{1}{\theta^3} & \text{gdy} \quad \theta \geq \max(x_1, x_2, x_3) = 1,57 \\ 0, & \text{gdy} \quad \theta < \max(x_1, x_2, x_3) = 1,57. \end{cases}$$

Przykład cd.

Zauważmy, że gdy θ jest mniejsza od choć jedna z wartości x_i to funkcja wiarogodności $L(\theta \,|\, x_1, x_2, x_3)$ jest równa zero oraz jest równa $\frac{1}{\theta^3}$ gdy θ jest większa od wszystkich wartości x_i.

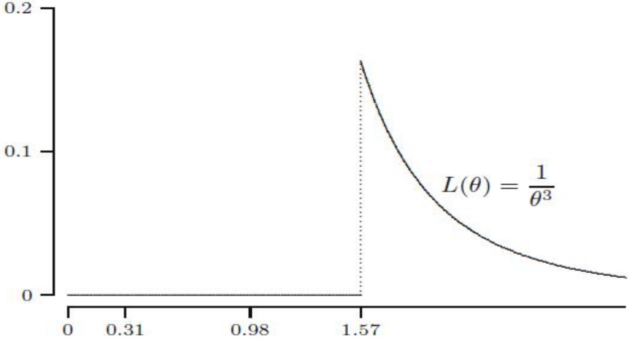
Tak więc ponieważ

$$L(\theta \,|\, x_1, x_2, x_3) = f(x_1 \,|\, \theta) f(x_2 \,|\, \theta) f(x_3 \,|\, \theta)$$

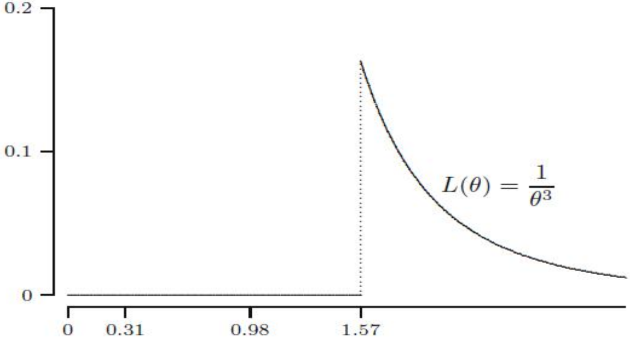
otrzymujemy

$$L(\theta \,|\, x_1, x_2, x_3) = \begin{cases} \frac{1}{\theta^3} & \text{gdy} \quad \theta \geq \max(x_1, x_2, x_3) = 1,57 \\ 0, & \text{gdy} \quad \theta < \max(x_1, x_2, x_3) = 1,57. \end{cases}$$

Przykład cd.



Przykład cd.



Estymator nośnika w rozkładzie $U[0, \theta]$.

W przypadku ogólnym, jeśli $(x_1, x_2, ..., x_n)$ stanowią próbę z rozkładu $U(0, \theta)$ to widzimy, że funkcja wiarogodności

$$L(\theta) = 0 \quad \text{gdy} \quad \theta < \max(x_1, x_2, ..., x_n)$$

oraz

$$L(\theta) = \frac{1}{\theta^n} \quad \text{gdy} \quad \theta \geq \max(x_1, x_2, ..., x_n).$$

Możemy stąd wywnioskować, że estymatorem największej wiarogodności parametru θ jest w tej sytuacji

$$\hat{\theta} = \max\{X_1, X_2, ..., X_n\}.$$

Estymator nośnika w rozkładzie $U[0, \theta]$.

W przypadku ogólnym, jeśli $(x_1, x_2, ..., x_n)$ stanowią próbę z rozkładu $U(0, \theta)$ to widzimy, że funkcja wiarogodności

$$L(\theta) = 0 \quad \text{gdy} \quad \theta < \max(x_1, x_2, ..., x_n)$$

oraz

$$L(\theta) = \frac{1}{\theta^n} \quad \text{gdy} \quad \theta \geq \max(x_1, x_2, ..., x_n).$$

Możemy stąd wywnioskować, że estymatorem największej wiarogodności parametru θ jest w tej sytuacji

$$\hat{\theta} = \max\{X_1, X_2, ..., X_n\}.$$

Własności estymatorów NW

Zakładamy, że parametrem badanego rozkładu jest θ ale nas interesuje estymacja, nie samego parametru, ale funkcji od tego parametru $\tau(\theta)$.

Jednym z powodów popularności estymatorów największej wiarogodności jest następująca ich własność.

Zasada niezmienniczości dla estymatorów NW mówi, że jeśli $\hat{\theta}$ jest estymatorem NW parametru θ to $\tau(\hat{\theta})$ jest estymatorem NW dla $\tau(\theta)$.

Własności estymatorów NW

Zakładamy, że parametrem badanego rozkładu jest θ ale nas interesuje estymacja, nie samego parametru, ale funkcji od tego parametru $\tau(\theta)$.

Jednym z powodów popularności estymatorów największej wiarogodności jest następująca ich własność.

Zasada niezmienniczości dla estymatorów NW mówi, że jeśli $\hat{\theta}$ jest estymatorem NW parametru θ to $\tau(\hat{\theta})$ jest estymatorem NW dla $\tau(\theta)$.

Własności estymatorów NW

Zakładamy, że parametrem badanego rozkładu jest θ ale nas interesuje estymacja, nie samego parametru, ale funkcji od tego parametru $\tau(\theta)$.

Jednym z powodów popularności estymatorów największej wiarogodności jest następująca ich własność.

Zasada niezmienniczości dla estymatorów NW mówi, że jeśli $\hat{\theta}$ jest estymatorem NW parametru θ to $\tau(\hat{\theta})$ jest estymatorem NW dla $\tau(\theta)$.

Estymator NW dla funkcji parametru

Twierdzenie.

Jeśli $\hat{\theta}$ jest estymatorem NW parametru θ to dla dowolnej funkcji $\tau(\theta)$, która jest odwracalna $\tau(\hat{\theta})$ jest estymatorem NW dla $\tau(\theta)$.

Estymator NW dla funkcji parametru

Przykład 1.

Estymatorem NW dla μ, kwadratu wartości oczekiwanej w rozkładzie normalnym, jest $[\bar{X}]^2$.

Przykład 2.

Estymatorem NW dla $\sqrt{p(1-p)}$ w rozkładzie 0-1 jest $\sqrt{\hat{p}(1-\hat{p})}$.

Estymator NW dla funkcji parametru

Przykład 1.

Estymatorem NW dla μ, kwadratu wartości oczekiwanej w rozkładzie normalnym, jest $[\bar{X}]^2$.

Przykład 2.

Estymatorem NW dla $\sqrt{p(1-p)}$ w rozkładzie 0-1 jest $\sqrt{\hat{p}(1-\hat{p})}$.

Estymator NW dla funkcji parametru

Przykład 1.

Estymatorem NW dla μ, kwadratu wartości oczekiwanej w rozkładzie normalnym, jest $[\bar{X}]^2$.

Przykład 2.

Estymatorem NW dla $\sqrt{p(1-p)}$ w rozkładzie 0-1 jest $\sqrt{\hat{p}(1-\hat{p})}$.

Własności estymatorów NW

Wada estymatora największej wiarogodności. Okazuje się, że estymator NW jest czuły na zaburzenia danych, na podstawie których obliczamy jego wartość.

Powiedzmy, że na dla wektora obserwacji

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

obliczamy wartość funkcji wiarogodności $L(\theta \,|\, \mathbf{x}$ i na jej podstawie otrzymujemy wartość estymatora NW parametru θ równą $\hat{\theta}$.

Następnie dokonujemy zaburzenia wektora danych $\mathbf{x}$,

$$\mathbf{x} + \epsilon = (x_1 + \epsilon, x_2 + \epsilon, \ldots, x_n + \epsilon)$$

, i na podstawie tak zaburzonych danych obliczamy wartość estymatora NW otrzymując wartość $\hat{\theta}_1$.

Własności estymatorów NW

Wada estymatora największej wiarogodności. Okazuje się, że estymator NW jest czuły na zaburzenia danych, na podstawie których obliczamy jego wartość.

Powiedzmy, że na dla wektora obserwacji

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

obliczamy wartość funkcji wiarogodności $L(\theta \,|\, \mathbf{x}$ i na jej podstawie otrzymujemy wartość estymatora NW parametru θ równą $\hat{\theta}$.

Następnie dokonujemy zaburzenia wektora danych $\mathbf{x}$,

$$\mathbf{x} + \epsilon = (x_1 + \epsilon, x_2 + \epsilon, \ldots, x_n + \epsilon)$$

, i na podstawie tak zaburzonych danych obliczamy wartość estymatora NW otrzymując wartość $\hat{\theta}_1$.

Własności estymatorów NW

Wada estymatora największej wiarogodności. Okazuje się, że estymator NW jest czuły na zaburzenia danych, na podstawie których obliczamy jego wartość.

Powiedzmy, że na dla wektora obserwacji

$$\mathbf{x} = (x_1, x_2, ..., x_n)$$

obliczamy wartość funkcji wiarogodności $L(\theta \,|\, \mathbf{x}$ i na jej podstawie otrzymujemy wartość estymatora NW parametru θ równą $\hat{\theta}$.

Następnie dokonujemy zaburzenia wektora danych $\mathbf{x}$,

$$\mathbf{x} + \epsilon = (x_1 + \epsilon, x_2 + \epsilon, ..., x_n + \epsilon)$$

, i na podstawie tak zaburzonych danych obliczamy wartość estymatora NW otrzymując wartość $\hat{\theta}_1$.

Własności estymatorów NW

Wada estymatora największej wiarogodności. Okazuje się, że estymator NW jest czuły na zaburzenia danych, na podstawie których obliczamy jego wartość.

Powiedzmy, że na dla wektora obserwacji

$$\mathbf{x} = (x_1, x_2, ..., x_n)$$

obliczamy wartość funkcji wiarogodności $L(\theta \,|\, \mathbf{x}$ i na jej podstawie otrzymujemy wartość estymatora NW parametru θ równą $\hat{\theta}$.

Następnie dokonujemy zaburzenia wektora danych $\mathbf{x}$,

$$\mathbf{x} + \epsilon = (x_1 + \epsilon, x_2 + \epsilon, ..., x_n + \epsilon)$$

, i na podstawie tak zaburzonych danych obliczamy wartość estymatora NW otrzymując wartość $\hat{\theta}_1$.

Własności estymatorów NW

Wada estymatora największej wiarogodności. Okazuje się, że estymator NW jest czuły na zaburzenia danych, na podstawie których obliczamy jego wartość.

Powiedzmy, że na dla wektora obserwacji

$$\mathbf{x} = (x_1, x_2, ..., x_n)$$

obliczamy wartość funkcji wiarogodności $L(\theta \,|\, \mathbf{x}$ i na jej podstawie otrzymujemy wartość estymatora NW parametru θ równą $\hat{\theta}$.

Następnie dokonujemy zaburzenia wektora danych $\mathbf{x}$,

$$\mathbf{x} + \epsilon = (x_1 + \epsilon, x_2 + \epsilon, ..., x_n + \epsilon)$$

, i na podstawie tak zaburzonych danych obliczamy wartość estymatora NW otrzymując wartość $\hat{\theta}_1$.

Własności estymatorów NW

Możemy oczekiwać, że dla małej wartości ϵ otrzymamy

$$\hat{\theta} \approx \hat{\theta}_1.$$

Przykład.

Estymatorem NW parametrów k i p w rozkładzie $b(k, p)$, w sytuacji gdy k i p są nieznane ma postać

$$\begin{aligned} \hat{p} &= \frac{\bar{X}}{\hat{k}} \\ gdzie, & \\ \hat{k} &= \frac{\bar{X}^2}{\bar{X} - \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}. \end{aligned}$$

Przykład cd.

Powiedzmy, że pięcioelementowa próby losowej z rozkładu $b(k, p)$, dla którego k i p są nieznane ma postać

(16, 18, 22, 25, 27) (liczba sukcesów w nieznanej liczbie prób).

Na podstawie tej próby otrzymujemy $\hat{k} = 99$ jako oszacowanie nieznanej liczby prób.

Jeśli w ciągu danych zwiększymy liczbę sukcesów w jednej z obserwacji o 1, powiedzmy, ze na ostatniej pozycji

$$(16, 18, 22, 25, \boxed{28}),$$

to wtedy estymator NW przyjmuje wartość $\hat{k} = 190$, co powoduje prawie dwukrotny wzrost szacowanej ilości prób.

Przykład cd.

Powiedzmy, że pięcioelementowa próby losowej z rozkładu $b(k, p)$, dla którego k i p są nieznane ma postać

(16, 18, 22, 25, 27) (liczba sukcesów w nieznanej liczbie prób).

Na podstawie tej próby otrzymujemy $\hat{k} = 99$ jako oszacowanie nieznanej liczby prób.

Jeśli w ciągu danych zwiększymy liczbę sukcesów w jednej z obserwacji o 1, powiedzmy, ze na ostatniej pozycji

$$(16, 18, 22, 25, \boxed{28}),$$

to wtedy estymator NW przyjmuje wartość $\hat{k} = 190$, co powoduje prawie dwukrotny wzrost szacowanej ilości prób.

Przykład cd.

Powiedzmy, że pięcioelementowa próby losowej z rozkładu $b(k, p)$, dla którego k i p są nieznane ma postać

$(16, 18, 22, 25, 27)$ (liczba sukcesów w nieznanej liczbie prób).

Na podstawie tej próby otrzymujemy $\hat{k} = 99$ jako oszacowanie nieznanej liczby prób.

Jeśli w ciągu danych zwiększymy liczbę sukcesów w jednej z obserwacji o 1, powiedzmy, ze na ostatniej pozycji

$$(16,\ 18,\ 22,\ 25,\ \boxed{28}),$$

to wtedy estymator NW przyjmuje wartość $\hat{k} = 190$, co powoduje prawie dwukrotny wzrost szacowanej ilości prób.