PRZYKŁAD

Wyznaczyć estymatory metody największej wiarygodności, gdy $(X_1, X_2 \dots X_n)$ to próba prosta dla parametru Θ

a) $X_k \sim B(l, p)$ $\quad p \in (0,1)$ $\quad p$ nieznane $\quad \Theta = p$

$$L(x_1, x_2 \dots x_n, p) = \prod_{k=1}^{n} P_p(X_k = x_k) = \prod_{k=1}^{n} \binom{l}{x_k} p^{x_k} (1-p)^{l-x_k}$$

$x_1, x_2 \dots x_n \to$ obserwacje próby losowej $(X_1, X_2 \dots X_n)$

$$P(X_k = x_k) = \binom{l}{x_k} p^{x_k} (1-p)^{l-x_k}$$

$$\ln L(X_1 \dots X_n, p) = \sum_{k=1}^{n} \left(\ln\binom{l}{x_k} + x_k \ln p + (l - x_k) \ln(1-p) \right)$$

$$\frac{\partial}{\partial p} \ln L(x_1 \dots x_n, p) = \sum_{k=1}^{n} \left(x_k \frac{1}{p} + (l - x_k) \frac{(-1)}{1-p} \right) = 0 \qquad / \cdot p(1-p) \neq 0$$

$$\sum_{k=1}^{n} \left(x_k (1-p) + (l - x_k)(-1) p \right) = 0$$

$$\sum_{k=1}^{n} \left(x_k - x_k p - lp + x_k p \right) = 0$$

$$\left(\sum_{k=1}^{n} x_k \right) - nlp = 0$$

$$p = \frac{\sum_{k=1}^{n} x_k}{nl} = \frac{\bar{x}}{l}$$

lokalne maksimum jest największą wartością funkcji

$$\frac{\partial^2 \ln L}{\partial p^2} = \sum_{k=1}^{n} \left(x_k \frac{-1}{p^2} + (l - x_k) \frac{-1}{(1-p)^2} \right) < 0$$

$x_k \in \{0, 1 \dots l\}$

- niech $X_k \sim B(4, p)$, p nieznane. Obliczyć estymator metody MLE dla obserwacji: 2, 1, 3, 0, 2

$n = 5 \qquad l = 4$

$$\hat{p} = \frac{\bar{x}}{l} = \frac{\frac{8}{5}}{4} = \frac{2}{5} = 0,4$$

b) LISTA 8, zadanie 6

$L(x_1, x_2 \dots x_n, p)$ $\qquad X_k$ ma gęstość $f_c(x) = \begin{cases} \frac{c}{x^{c+1}} & x > 1 \\ 0 & x \leq 1 \end{cases}$

$\Theta = c$

$L(\quad) = \prod P(\quad) = \Big\{ \prod_{k=1}^{n} \frac{c}{x_k^{c+1}}$ $\quad x_k > 1$ dla $k = 1, 2 \dots$ $\Theta = c$

$\Theta = c$

$$L(x_1, x_2 \dots x_n, p) = \prod_{k=1}^{n} f_c(x_k) = \begin{cases} \prod_{k=1}^{n} \frac{c}{x_k^{c+1}} & x_k > 1 \quad \text{dla } k = 1,2 \dots \Theta = c \\ 0 & \text{przynajmniej jedno } x_k \leq 1 \end{cases}$$

$$\ln L(x_1 \dots x_n) = \sum_{k=1}^{n} (\ln c - (c+1)\ln x_k) \qquad x_k > 1 \qquad k = 1 \dots n$$

$$\frac{\partial}{\partial c} \ln L(x_1 \dots x_n, c) = \sum_{k=1}^{n} \left(\frac{1}{c} - \ln x_k\right) = \frac{n}{c} - \sum_{k=1}^{n} \ln x_k = 0$$

$$c = \frac{n}{\sum_{k=1}^{n} \ln x_k}$$ ← lokalne maksimum w którym jest największa wartość funkcji $L(X_1 \dots x_n, c)$

$$\frac{\partial^2}{\partial c^2} \ln L = \sum_{k=1}^{n} \frac{-1}{c^2} < 0$$

- niech $(X_1 \dots X_n)$ to próba prosta, X_k ma gęstość $f_c(x)$

Dla obserwacji 2, 4, 6, 8, 7 wyznaczyć estymator MLE dla parametru c.

$n = 5$

$$c = \frac{n}{\sum_{k=1}^{n} \ln(x_k)} = \frac{5}{\ln 2 + \ln 4 + \ln 6 + \ln 8 + \ln 7} = \frac{5}{\ln(2 \cdot 4 \cdot 6 \cdot 8 \cdot 7)} = \frac{5}{7,9} = 0,63$$

ZADANIE DOMOWE

Niech $(X_1 \dots X_n)$ to próba prosta, że $X_k \sim N(m, \sigma)$
Wyznaczyć estymator metody największej wiarygodności dla $\Theta = (m, \sigma)$.

WSTĘP DO ZADANIA:

- funkcja wiarygodności: $L(x_1 \dots x_n, m, \sigma) = \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x_k - m)^2}{2\sigma^2}}$
- szukamy max wartości funkcji

(2) METODA MOMENTÓW

Polega na przyrównaniu momentów rozkładu zmiennej losowej z jej momentem empirycznym, tzn.

np. $EX_k = \overline{X}$

$Var X_k = S^2$

$EX_k^l = \frac{1}{n} \sum_{k=1}^{n} X_k^l$

- PRZYKŁAD: niech $X_1, X_2 \dots X_n$ → próba prosta z rozkładu jednostajnego na przedziale $[a, b]$. Wyznaczyć estymatory metody momentów dla parametrów a, b.

$$\begin{cases} EX_k = \frac{a+b}{2} = \bar{X} \\ VarX_k = \frac{(b-a)^2}{12} = S^2 \end{cases} \longrightarrow \begin{cases} b = 2\bar{X} - a \\ \frac{(2\bar{x}-a)^2}{12} = S^2 \end{cases}$$

$$\frac{(2\bar{x}-a-a)^2}{12} = S^2 \qquad \frac{(\bar{x}-a)^2}{3} = S^2 \qquad (\bar{x}-a)^2 = 3S^2$$

$$\bar{x} - a = \sqrt{3}s \quad \cup \quad \bar{x} - a = -\sqrt{3}s$$

$$\begin{cases} \hat{a} = \bar{x} - \sqrt{3}s \\ \hat{b} = \bar{x} + \sqrt{3}s \end{cases} \cup \begin{cases} \hat{a} = \bar{x} + \sqrt{3}s \\ \hat{b} = \bar{x} - \sqrt{3}s \end{cases}$$

- PRZYKŁAD: Wyznaczyć estymatory metody momentów gdy $(X_1 \dots X_n)$ → próba prosta, X_k ma rozkład jednostajny na $[a,b]$, a, b → nieznane, dla obserwacji $(1.2, 2.1, 1.5, 1.8)$.

$$\bar{x} = \frac{1}{4}(1{,}2 + 2{,}1 + 1{,}5 + 1{,}8) = 1{,}7$$

$$S^2 = 0{,}115$$

$$\hat{a} = 1{,}7 - \sqrt{3} \cdot \sqrt{0{,}115} = 1{,}11$$

$$\hat{b} = 1{,}7 + \sqrt{3} \cdot \sqrt{0{,}115} = 2{,}29$$

ESTYMACJA PRZEDZIAŁOWA PARAMETRÓW ROZKŁADU ZMIENNEJ LOSOWEJ

- $(X_1, X_2 \dots X_n)$ → próba prosta
- nie chcemy przybliżyć nieznanego parametru kilkoma liczbami, chcemy podać przedział liczbowy, który z zadanym przez nas dużym prawdopodobieństwem nakryje nieznany parametr liczbowy
- PRZEDZIAŁEM UFNOŚCI dla parametru Θ na poziomie ufności $1-\alpha$ (α jest mała liczba, najczęściej $0.05, 0.02, 0.1, 0.01$) nazywamy przedział liczbowy (Θ_1, Θ_2) taki, że:

1. $\Theta_i = f(X_1, X_2 \dots X_n) \qquad i = 1,2$
2. przedział (Θ_1, Θ_2) nakrywa nieznany parametr Θ z prawdopodobieństwem

$$P(\Theta_1 < \Theta < \Theta_2) = \underbrace{1-\alpha}_{\text{poziom ufności}}$$

- PRZYKŁAD: niech $(X_1, X_2 \dots X_n)$ → próba prosta z rozkładu $N(m, \sigma)$, σ → znane. Wyznaczyć przedział ufności dla parametru m na poziomie ufności $1-\alpha$.

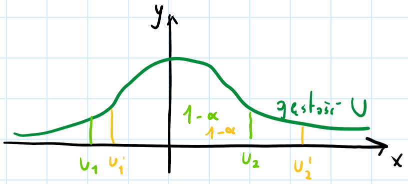
$\bar{X} \sim N\left(m, \frac{\sigma}{\sqrt{n}}\right)$ standaryzacja

$U = \frac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

takich przedziałów jest nieskończenie wiele
najkrótszy z nich jest symetryczny wzgl. 0,
$u_1 = -u_2$

znajdźmy u_1, u_2 takie, że $P(u_1 < U < u_2) = 1 - \alpha$

dla $1 - \alpha > 0,5$



$$P(-u_\alpha < U < u_\alpha) = \Phi(u_\alpha) - \Phi(-u_\alpha) = \underbrace{-1 + 2\Phi(u_\alpha)}_{\Phi(u_\alpha) = 1 - \frac{\alpha}{2}} = 1 - \alpha$$

$$P\left(-u_\alpha < \frac{\bar{X} - m}{\sigma}\sqrt{n} < u_\alpha\right) = 1 - \alpha$$

$$P\left(\bar{X} - \frac{\sigma u_\alpha}{\sqrt{n}} < m < \bar{X} + \frac{\sigma u_\alpha}{\sqrt{n}}\right) = 1 - \alpha$$

- PRZYKŁAD: czas analizy ma rozkład normalny $N(m, 2)$.
  Dla obserwacji próby prostej 34, 35, 30, 32, 34, 31, 34, 33, 35.

Wyznaczyć przedział ufności dla wartości oczekiwanej czasu analizy (średniej czasu analizy, przeciętnej czasu analizy) na poziomie ufności $1 - \alpha = 0,95$

X_k – czas k-tej analizy, $k = 1 \dots 9$ $\quad n = 9$

$X_k \sim N(m, 2)$ $\quad \sigma = 2$

$EX = m$

$$\bar{X} - \frac{\sigma u_\alpha}{\sqrt{n}} < m < \bar{X} + \frac{\sigma u_\alpha}{\sqrt{n}}$$

$$\Phi u_\alpha = 1 - \frac{\alpha}{2} = 1 - \frac{0,05}{2} = 0,975 \underset{\text{tablice}}{=} \Phi(1,96)$$

$u_\alpha = 1,96$

$\bar{X} = \frac{1}{9}(34 + \dots + 35) = 33,1$

$$31,8 = 33,1 - \frac{2 \cdot 1,96}{\sqrt{9}} < m < 31,1 + \frac{2 \cdot 1,96}{\sqrt{9}} = 34,4$$

ODP:
Przedział $(31,8 , 34,4)$ zawiera parametr m z prawdopodobieństwem 0,95