Tw. 1

Niech $X_1, \dots, X_n$ będzie próbą z regularnej rodziny wykładniczej. Rozważmy statystykę $Y_1 = \sum_{i=1}^n K(X_i)$. Wtedy,

(i) rozkład Y_1 ma postać

$$f_{Y_1}(y_1, \theta) = R(y_1) \exp[p(\theta) y_1 + n q(\theta)] \quad \text{dla } y_1 \in S_{Y_1}$$

i pewnej funkcji $R(y_1)$. Przy czym S_{Y_1} i $R(y_1)$ nie zależą od θ.

(ii) $$EY_1 = -n \frac{q'(\theta)}{p'(\theta)}.$$

(iii) $$\mathrm{Var}\, Y_1 = n \frac{1}{p'(\theta)^3} \{ (p''(\theta) q'(\theta) - q''(\theta) p'(\theta) \}.$$

Przykład 2

$X \sim \mathrm{Pois}(\theta)$, $\theta \in (0, \infty)$ $\xrightarrow{\text{wtedy}}$ $S = \{0, 1, 2, \dots\}$ nie zależy od θ.

$$f(x, \theta) = e^{-\theta} \frac{\theta^x}{x!} = \exp\{ (\log\theta) x + \log(\tfrac{1}{x!}) + (-\theta) \}.$$

Zatem rozkład Poissona należy do regularnej rodziny wykładniczej, gdzie $p(\theta) = \log\theta$, $q(\theta) = -\theta$, $K(x) = x$.

Jeżeli $X_1, \dots, X_n$ jest próbą z tego rozkładu, to $Y_1 = \sum_{i=1}^n X_i$ jest statystyką dostateczną dla θ.

Mamy $p'(\theta) = \frac{1}{\theta}$, $q'(\theta) = -1$, zatem $EY_1 = n\theta$.

$p''(\theta) = -\frac{1}{\theta^2}$, $q''(\theta) = 0$. Zatem $\mathrm{Var}\, Y_1 = n\theta^3 \{ -\frac{1}{\theta^2}(-1) \} = n\theta$.

~~Ponadto $R(y_1)$~~

Tw. 2

Niech $f(x, \theta)$, $\gamma < \theta < \delta$, będzie rozkładem zmiennej losowej X należącym do regularnej rodziny wykładniczej. Jeżeli $X_1, \dots, X_n$ jest próbą z tego rozkładu, to $Y_1 = \sum_{i=1}^n K(X_i)$ jest statystyką dostateczną dla θ, a rodzina rozkładów $\{ f_{Y_1}(y_1, \theta) : \gamma < \theta < \delta \}$ zmiennej losowej Y_1 jest zupełna. (Y_1 jest zupełną statystyką dostateczną dla θ).

Dowód Pokażemy zupełność.

1. Załóżmy, że $E[u(Y_1)] = 0$.
2. ~~Stąd~~ Zatem, dla każdego θ,

$$\int_{S_{Y_1}} u(y_1)R(y_1)\exp\{p(\theta)y_1 + nq(\theta)\}dy_1 = 0$$

równoważnie

$$\int_{S_{Y_1}} u(y_1)R(y_1)\exp\{p(\theta)y_1\}dy_1 = 0 \quad \text{ponieważ } \exp\{nq(\theta)\} \neq 0.$$

3. Ponieważ $p(\theta)$ jest nietrywialną, ciągłą funkcją θ, to powyższa całka jest, w zasadzie, typem ~~prz~~ transformaty Laplace'a $u(y_1)R(y_1)$.
   // $F(s) = \int_0^\infty e^{-st} f(t)\,dt$
4. Stąd $u(y_1)R(y_1) \equiv 0$. poza zbiorem miary zero.
5. Ponieważ $R(y_1) \neq 0$ dla $y_1 \in S_{Y_1}$, to $u(y_1) \equiv 0$ poza zbiorem miary zero.
6. Zatem Y_1 jest (statystyką) zupełną dla θ.

Wniosek 1

Jeżeli $X_1, \dots, X_n$ jest próbą z regularnej rodziny wykładniczej, ~~to~~ a φ, taką funkcją, $\left(Y_1 = \sum_{i=1}^n K(X_i)\right)$ że $E[\varphi(Y_1)] = \theta$, to $\varphi(Y_1)$ jest (jednoznacznie wyznaczonym) ENMW parametru θ.

Przykład 3

$X_1, \dots, X_n$ iid, $X_i \sim f(x|\theta) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\theta)^2}{2\sigma^2}\right\}$, $x \in \mathbb{R}$, $\theta \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}$ ustalone

Równoważnie

$$f(x|\theta) = \exp\left\{\frac{\theta}{\sigma^2}x - \frac{x^2}{2\sigma^2} - \log\sqrt{2\pi\sigma^2} - \frac{\theta^2}{2\sigma^2}\right\}.$$

należy do regularnej rodziny wykładniczej, przy czym

$$p(\theta) = \frac{\theta}{\sigma^2},\ K(x) = x,\ S(x) = -\frac{x^2}{2\sigma^2} - \log\sqrt{2\pi\sigma^2},\ q(\theta) = -\frac{\theta^2}{2\sigma^2}.$$

Zatem $Y_1 = \sum_{i=1}^n X_i$ jest zupełną statystyką dostateczną dla θ.

Ponieważ $\varphi(Y_1) = \frac{Y_1}{n} = \bar{X}$ jest (estymatorem) nieobciążonym θ, to $\varphi(Y_1)$ jest (jednoznacznie wyznaczonym) ENMW.

Ponieważ $Y_1 \xrightarrow{1-1} \bar{X}$ więc $\bar{X}$ jest również zupełną statystyką dostateczną.

Przykład 4 (rozkład Poissona, c.d.)

$Y_1 = \sum_{i=1}^{n} X_i$ zupełna statystyka dostateczna dla θ. Ponieważ $EY_1 = n\theta$, $\overline{X} = \frac{Y_1}{n}$ jest jednoznacznie wyznaczonym ENMW parametru θ.

6. Funkcje parametru

Przykład 1

$X_1, \dots, X_n$ iid, $X_i \sim b(1,\theta)$. Jeżeli $Y = \sum_{i=1}^{n} X_i$, to $\frac{Y}{n}$ jest jednoznacznie wyznaczonym ENMW parametru θ.

Przypuśćmy, że chcemy estymować wariancję $\frac{Y}{n}$ tzn. $\theta(1-\theta)/n$.

Niech $\delta = \theta(1-\theta)$. Ponieważ Y jest statystyką dostateczną dla θ możemy się ograniczyć do funkcji Y. ENW δ ma postać $\tilde{\delta} = \frac{Y}{n}\left(1-\frac{Y}{n}\right)$.

parametru

~~jest~~ Ponadto,

$$E[\tilde{\delta}] = E\left[\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right] = \frac{1}{n}EY - \frac{1}{n^2}EY^2 =$$

$$\frac{1}{n}\cdot n\theta - \frac{1}{n^2}\left[n\theta(1-\theta) + n^2\theta^2\right] = \theta - \frac{\theta(1-\theta)}{n} - \theta^2 = (n-1)\frac{\theta(1-\theta)}{n}.$$

Zatem $\hat{\delta} = \frac{n}{n-1}\tilde{\delta} = \frac{n}{n-1}\frac{Y}{n}\left(1-\frac{Y}{n}\right)$ jest jednoznacznie wyznaczonym ENMW δ.

Stąd $\frac{\hat{\delta}}{n}$ jest ENMW $\frac{\delta}{n} = \theta(1-\theta)/n$.

Porównanie $\tilde{\delta}$ i $\hat{\delta}$

$\frac{1}{I(\delta)}$ ok

ENW $\quad \tilde{\delta} \xrightarrow{P} \delta, \quad \sqrt{n}(\tilde{\delta}-\delta) \xrightarrow{D} N\left(0, \cancel{\frac{1}{I(\theta)}}\right)$

ENMW $\quad \hat{\delta} = \frac{n}{n-1}\tilde{\delta} \xrightarrow{P} \delta$

Ponieważ

$$\sqrt{n}(\hat{\delta}-\delta) - \sqrt{n}(\tilde{\delta}-\delta) = \frac{\sqrt{n}}{n-1}\tilde{\delta} \xrightarrow{P} 0,$$

zatem $\sqrt{n}(\hat{\delta}-\delta)$ i $\sqrt{n}(\tilde{\delta}-\delta)$ mają te same rozkłady asymptotyczne.

(Na podstawie ~~Z~~ metody Δ)

Niech $g(\theta) = \theta(1-\theta)$. Wtedy $g'(\theta) = 1-2\theta$. Ponieważ

$$\sqrt{n}\left(\frac{Y}{n} - \theta\right) \xrightarrow{D} N(0, \theta(1-\theta)),$$

$\boxed{g'(\theta) \neq 0}$

$$\sqrt{n}(\hat{\delta}-\delta) = \sqrt{n}\left(\frac{n}{n-1}\frac{Y}{n}\left(1-\frac{Y}{n}\right) - \theta(1-\theta)\right) = \sqrt{n}\left(\frac{n}{n-1}g\left(\frac{Y}{n}\right) - g(\theta)\right) \xrightarrow{D} N\left(0, \theta(1-\theta)[1-2\theta]^2\right)$$

Przykład 2

$X_1, \dots, X_n$ iid $X_i \sim N(\theta, 1)$

Szukamy ENMW funkcji θ

$$\Phi(c-\theta) = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-\theta)^2}{2}} dx = P(X \leq c), \text{ gdzie } c \in \mathbb{R}.$$

Rozważmy funkcję

$$u_1(x_1) = \begin{cases} 1 & x_1 \leq c, \\ 0 & x_1 > c. \end{cases}$$

Wtedy

$$E[u(X_1)] = 1 \cdot P(X_1 \leq c) = P(X_1 - \theta \leq c - \theta) = \Phi(c-\theta).$$

Zatem $u(X_1)$ jest nieobciążonym estymatorem $\Phi(c, -\theta)$.

$\bar{X}$ - statystyka dostateczna dla θ.

Łączny rozkład $(X_1, \bar{X}) \sim N_2\left(\begin{pmatrix}\theta\\ \theta\end{pmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ & \sigma_2^2 \end{bmatrix}\right)$

$\sigma_1^2 = 1$, $\sigma_2^2 = \frac{1}{n}$, $\rho = \frac{1}{\sqrt{n}}$ (ćw)

Zatem, rozkład warunkowy X_1 pod warunkiem $\bar{X}$ jest rozkładem normalnym ze średnią

$$\theta + \frac{\rho\sigma_1}{\sigma_2}(\bar{x} - \theta) = \bar{x}$$

i wariancją $\sigma_1^2(1-\rho^2) = \frac{n-1}{n}$.

Stąd, warunkowa wartość oczekiwana $E[u(X_1) | \bar{X}] = \varphi(\bar{X})$ ma postać

$$\varphi(\bar{x}) = \int_{-\infty}^{\infty} u(x_1) \sqrt{\frac{n}{n-1}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{n(x_1-\bar{x})^2}{(n-1)\,2}\right] dx_1 \overset{\text{z df } u(x_1)}{=}$$

$$\int_{-\infty}^{c} \sqrt{\frac{n}{n-1}} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{n(x_1-\bar{x})^2}{(n-1)\,2}\right] dx_1 = \left\{\begin{aligned} z &= \sqrt{n}(x_1 - \bar{x})/\sqrt{n-1} \\ dz &= \frac{\sqrt{n}}{\sqrt{n-1}} dx_1 \end{aligned}\right\}$$

$$= \int_{-\infty}^{c'} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{z^2}{2}\right] dz = \Phi(c') = \Phi\left(\frac{\sqrt{n}(c-\bar{x})}{\sqrt{n-1}}\right).$$

Zatem MVUE ~~jest~~ $\Phi(c-\theta)$ jest jednoznacznie wyznaczony, dla każdego ustalonego c i ma postać

$$\varphi(\bar{X}) = \Phi\left(\frac{\sqrt{n}(c-\bar{X})}{\sqrt{n-1}} \right).$$

ENW $\Phi(c-\theta)$ ma postać $\Phi(c-\bar{X})$.

Przykład 3

$X \sim Pois(\theta)$, X - zupełna statystyka dostateczna dla θ.

Problem: wyznaczenie estymatora nieobciążonego o min. war. parametru $e^{-2\theta}$

Rozważmy $Y = (-1)^X$. Mamy

$$EY = E(-1)^X = \sum_{x=0}^{\infty} \frac{(-\theta)^x e^{-\theta}}{x!} = e^{-2\theta}.$$

Zatem $(-1)^X$ jest ENMW parametru $e^{-2\theta}$, ale $0 < e^{-2\theta} < 1$, a estymator przyjmuje wartości $\{1, -1\}$!

ENW $e^{-2\theta}$ ma postać e^{-2X} (rozsądny)