Zatem w modelu przesunięcia informacja Fishera nie zależy od θ.

Załóżmy, że X_i ma rozkład Laplace'a. $f(x_i,\theta) = \frac{1}{2}e^{-|x_i-\theta|}$

nie spełnia warunków regularności

Ponieważ

$$X_i = \theta + e_i,$$

więc

$$e_i \sim f(z_i) = \frac{1}{2}e^{-|z_i|}$$

Ponadto,

$$f'(z) = -\frac{1}{2}e^{-|z|}\operatorname{sgn}(z).$$

Tym samym

$$I(\theta) = \int_{-\infty}^{\infty}\left(\frac{f'(z)}{f(z)}\right)^2 f(z)\,dz = \int_{-\infty}^{\infty} f(z)\,dz = 1.$$

Uwaga 2

Jeżeli $X_1, \dots, X_n$ iid, $X_i \sim f(x|\theta)$ oraz $I(\theta)$ jest informacją Fishera X_1, to $nI(\theta)$ jest informacją Fishera dla całej próby.

Dowód

$$\operatorname{Var}\left(\frac{\partial \log L(\theta, \underline{X})}{\partial\theta}\right) = \operatorname{Var}\left(\sum_{i=1}^{n}\frac{\partial \log f(X_i,\theta)}{\partial\theta}\right) = \sum_{i=1}^{n}\operatorname{Var}\left(\frac{\partial \log f(X_i,\theta)}{\partial\theta}\right) = nI(\theta).$$

Tw. 1 (Nierówność Craméra-Rao)

Niech $X_1, \dots, X_n$ będą niezależnymi zmiennymi losowymi z gęstością $f(x,\theta)$, $\theta \in \Theta$. Załóżmy, że warunki regularności (R0) – (R4) są spełnione.

Jeżeli $Y = u(X_1, \dots, X_n)$ jest taką statystyką, że $EY = E[u(X_1, \dots, X_n)] = k(\theta)$, to

$$\operatorname{Var}(Y) \geq \frac{[k'(\theta)]^2}{nI(\theta)}.$$

Dowód

I Załóżmy, że X_1 ma rozkład ciągły (dla dyskretnych analogicznie)

1. Mamy $k(\theta) = EY_1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, \dots, x_n)\, f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta)\, dx_1 \ldots dx_n$.

2. Stąd

$$k'(\theta) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, \dots, x_n) \left[\sum_{i=1}^{n} \frac{1}{f(x_i, \theta)} \frac{\partial f(x_i, \theta)}{\partial \theta} \right] \cdot f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta)\, dx_1 \ldots dx_n$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, \dots, x_n) \left[\sum_{i=1}^{n} \frac{\partial \log f(x_i, \theta)}{\partial \theta} \right] f(x_1, \theta) \cdot \ldots \cdot f(x_n, \theta)\, dx_1 \ldots dx_n$$

3. Zdefiniujmy zm. los $Z = \sum_{i=1}^{n} \frac{\partial \log f(X_i, \theta)}{\partial \theta}$.

4. Wtedy $EZ = 0$, $Var Z = nI(\theta)$.

5. Ponadto, $k'(\theta) = E[YZ] = E(Y)E(Z) + \rho \sigma_Y \cancel{\sqrt{nI(\theta)}}^{\,\sigma_Z} \overset{(4)}{=} \rho \sigma_Y \sqrt{nI(\theta)}$,

~~6. Stąd~~ gdzie $\rho = corr(Y, Z)$.

$\cancel{k'(\theta) =}$

6. Stąd

$$\rho = \frac{k'(\theta)}{\sigma_Y \sqrt{nI(\theta)}}.$$

7. Ponieważ $\rho^2 \leqslant 1$, zatem

$$\frac{[k'(\theta)]^2}{\sigma_Y^2\, nI(\theta)} \leqslant 1 \iff Var Y \geqslant \frac{[k'(\theta)]^2}{nI(\theta)}.$$

Wniosek 1

Przy założeniach powyższego twierdzenia, jeżeli $Y = u(X_1, \dots, X_n)$ jest nieobciążonym estymatorem (parametru) θ, $(k(\theta) = \theta)$, to

$$Var Y \geqslant \frac{1}{nI(\theta)}.$$

Przykład 3

$X_1, \dots, X_n$ iid $X_i \sim b(1, \theta)$. $\frac{1}{nI(\theta)} = \frac{\theta(1-\theta)}{n}$ [było]

Wtedy ~~ENW~~ $\hat{\theta} = \bar{X}$, a $E\bar{X} = \theta$, $Var \bar{X} = \frac{\theta(1-\theta)}{n}$.

Zatem wariancja $\bar{X}$ ~~spełnia~~ osiąga dolne ograniczenie w nierówności Craméra-Rao.

Def 3

Przy założeniach (R0)-(R4), jeżeli $Y = u(X_1, .., X_n)$ jest nieobciążonym estymatorem parametru θ, to liczbę

$$e_Y = \frac{\frac{1}{nI(\theta)}}{Var\, Y} = \frac{1}{nI(\theta)\cdot Var\, Y} \qquad e_Y \in [0,1].$$

nazywamy efektywnością estymatora Y.

Przykład 4

$X_1, \dots, X_n$ iid, $X_i \sim Pois(\theta)$, $\theta > 0$. ENW $\hat{\theta} = \bar{X}$. $E\bar{X} = \theta$, $Var\bar{X} = \frac{\theta}{n}$

$$P(X=x) = \frac{\theta^x e^{-\theta}}{x!}$$

Mamy $$\frac{\partial \log f(x,\theta)}{\partial\theta} = \frac{\partial}{\partial\theta}\left(x\log\theta - \theta - \log x!\right)$$
$$= \frac{x}{\theta} - 1,$$

$$E\left[\frac{\partial \log f(X,\theta)}{\partial\theta}\right]^2 = E\left(\frac{X}{\theta} - 1\right)^2 = \frac{1}{\theta^2}E\left[X-\theta\right]^2 = \frac{1}{\theta^2}\sigma_X^2 = \frac{\theta}{\theta^2} = \frac{1}{\theta}.$$

$$e_Y = \frac{1}{\left(n\frac{1}{\theta}\cdot\right)\left(\frac{\theta}{n}\right)} = 1.$$

$\bar{X}$ - estymator efektywny.

Przykład 5

$X_1, .., X_n$ iid $\quad X_i \sim f(x,\theta) = \begin{cases} \theta x^{\theta-1} & , x\in(0,1) \\ 0 & , w\ pp. \end{cases}$ $\quad$ Beta$(\theta, 1)$. $\quad \theta \in \Theta = (0,\infty)$.

$$\frac{\partial \log f(x,\theta)}{\partial\theta} = \frac{\partial\left(\log\theta + (\theta-1)\log x\right)}{\partial\theta} = \frac{1}{\theta} + \log x$$

$$\frac{\partial^2 \log f(x,\theta)}{\partial\theta^2} = -\frac{1}{\theta^2}. \quad \text{Zatem } I(\theta) = \frac{1}{\theta^2}$$

$$l(\theta) = \log L(\theta) = \sum_{i=1}^n \log f(x_i,\theta) = n\log\theta + \theta\sum_{i=1}^n \log x_i - \sum_{i=1}^n \log x_i.$$

$$\frac{\partial l(\theta)}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^n \log x_i = 0 \implies \hat{\theta} = -\frac{n}{\sum_{i=1}^n \log X_i}.$$

~~Def~~ Niech $Y_i = -\log X_i$. Wtedy $Y_i \sim \Gamma(1, \frac{1}{\theta})$, $W = \sum_{i=1}^{n} Y_i \sim \Gamma(n, \frac{1}{\theta})$.

Ponieważ $EW^k = \frac{(n+k-1)!}{\theta^k (n-1)!}$ dla $k > -n$. To

$$E\hat{\theta} = n\,E\left[W^{-1}\right] = n \cdot \frac{(n-2)!}{\theta^{-1}(n-1)!} = \theta \frac{n}{n-1}.$$

$$E\hat{\theta}^2 = n^2 E\left[W^{-2}\right] = n^2 \frac{(n-3)!}{\theta^{-2}(n-1)!} = \theta^2 \frac{n^2}{(n-1)(n-2)}.$$

W rezultacie

$$Var\,\hat{\theta} = E\hat{\theta}^2 - \left(E\hat{\theta}\right)^2 = \theta^2 \frac{n^2}{(n-1)(n-2)} - \theta^2 \frac{n^2}{(n-1)^2} = \theta^2 \frac{n^2(n-1) - n^2(n-2)}{(n-1)^2(n-2)} =$$

$$\theta^2 \frac{n^2}{(n-1)^2(n-2)}.$$

oraz

$$e_{\hat{\theta}} = \frac{1}{n I(\theta)\, Var\,\hat{\theta}} = \frac{1}{n \frac{1}{\theta^2} \cdot \frac{\theta^2 n^2}{(n-1)^2(n-2)}} = \frac{(n-1)^2(n-2)}{n^3} < 1.$$

$\hat{\theta}$ - nie jest efektywny, ale jest asymptotycznie efektywny.

Założenie

(R5): Gęstość $f(x|\theta)$ ~~jako fun~~ jest trzykrotnie różniczkowalna względem θ. Ponadto, dla wszystkich $\theta \in \Theta$, istnieje stała c i funkcja $M(x)$, taka że

$$\left| \frac{\partial^3}{\partial \theta^3} \log f(x|\theta) \right| < M(x), \quad a \quad E_{\theta_0}\left[M(X)\right] < \infty$$

dla wszystkich $\theta_0 - c < \theta < \theta_0 + c$ i wszystkich x należących do nośnika X.

Tw. 2

Załóżmy, że $X_1, \ldots, X_n$ są iid z gęstością $f(x|\theta_0)$ dla $\theta_0 \in \Theta$, takie że warunki regularności (R0)-(R5) są spełnione. Załóżmy, że informacja Fishera spełnia warunek $0 < I(\theta_0) < +\infty$. Wtedy, każdy zgodny ciąg $\{\hat{\theta}_n\}$ ~~esty~~ rozwiązań równania $\frac{d}{d\theta} L(\theta) = \frac{d}{d\theta} L(\theta, X_n) = 0$ spełnia warunek

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{D} N\left(0, \frac{1}{I(\theta_0)}\right).$$

Dowód

1. Z rozwinięcia $l'(\theta)$ w szereg Taylora ~~w punkcie θ_0~~ mamy

$$l'(\hat{\theta}_n) = l'(\theta_0) + (\hat{\theta}_n - \theta_0) l''(\theta_0) + \frac{1}{2}(\hat{\theta}_n - \theta_0)^2 l'''(\theta_n^*),$$

gdzie $\theta_n^* \in (\theta_0, \hat{\theta}_n) \cup (\hat{\theta}_n, \theta_0)$.

2. Ponieważ $l'(\hat{\theta}_n) = 0$, zatem

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{\frac{1}{\sqrt{n}} l'(\theta_0)}{-\frac{1}{n} l''(\theta_0) - \frac{1}{2n}(\hat{\theta}_n - \theta_0) l'''(\theta_n^*)}.$$

3. Ponieważ $\operatorname{var}\left(\frac{\partial \log f(X_i, \theta_0)}{\partial \theta}\right) = I(\theta_0) < \infty$, ~~Z~~ CTG implikuje

$$\frac{1}{\sqrt{n}} l'(\theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log f(X_i, \theta_0)}{\partial \theta} \xrightarrow{D} N(0, I(\theta_0))$$

4. Ze SPWL

$$-\frac{1}{n} l''(\theta_0) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2 \log f(X_i, \theta_0)}{\partial \theta^2} \xrightarrow{P} I(\theta_0).$$

bo w R5 jest $\forall \theta$

5. Ponieważ $\hat{\theta}_n - \theta_0 \xrightarrow{P} 0$ i to wystarczy pokazać, że $\frac{1}{n} l'''(\theta_n^*) = O_P(1)$

OK

6. $\boxed{\text{Z warunku (R5) ~~implikuje, że~~ istnieje stała } c_0 \text{, taka że } |\hat{\theta}_n - \theta_0| < c_0}$ ~~$\geq 1 - \delta$~~ dla n i odpow. dużych i $\delta > 0$. PWL

7. Stąd $|\theta_n^* - \theta_0| < c_0$ oraz Tu

$$\left| -\frac{1}{n} l'''(\theta_n^*) \right| \leqslant \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\partial^3 \log f(X_i, \theta)}{\partial \theta^3} \right|_{\theta = \theta_n^*} \leqslant \frac{1}{n} \sum_{i=1}^{n} M(X_i) \xrightarrow{P} E_{\theta_0}[M(X)]$$

~~8. Niech $\varepsilon > 0$, a N_1 i N_2 będą takie, że~~

~~dla~~

~~$$n \geqslant N_1 \Rightarrow P(|\hat{\theta}_n - \theta_0| < c_0) \geqslant 1 - \frac{\varepsilon}{2}$$~~

~~$$n \geqslant N_2 \Rightarrow P\left(\left|\frac{1}{n} \sum_{i=1}^{n} M(X_i) - E_{\theta_0}[M(X)]\right| < 1\right) \geqslant 1 - \frac{\varepsilon}{2}$$~~

~~9. Zatem, dla~~

~~$$n \geqslant \max\{N_1, N_2\} \Rightarrow P\left(\left|-\frac{1}{n} l'''(\theta_n^*)\right| \leq 1 + E_{\theta_0}[M(X)]\right) \geqslant 1 - \varepsilon.$$~~

8 Zatem ~~co oznacza, że~~ $\frac{1}{n} l'''(\theta_n^*) = O_P(1)$.

Def 4

Niech $X_1, \dots, X_n$ będą niezależnymi zmiennymi losowymi o gęstości $f(x,\theta)$. Załóżmy, że $\hat{\theta}_{1n} = \hat{\theta}_{1n}(X_1, \dots, X_n)$ jest estymatorem parametru θ_0, takim że $\sqrt{n}(\hat{\theta}_{1n} - \theta_0) \xrightarrow{D} N(0, \hat{\sigma}^2_{\hat{\theta}_{1n}})$, gdzie ~~$\hat{\sigma}^2_{\hat{\theta}_{1n}} = Var(\hat{\theta}_{1n})$~~. Wtedy

(i) liczbę

$$e(\hat{\theta}_{1n}) = \frac{1/I(\theta_0)}{\sigma^2_{\hat{\theta}_{1n}}}$$

nazywamy asymptotyczną efektywnością estymatora $\hat{\theta}_{1n}$.

(ii) ~~Mów~~ Jeżeli $e(\hat{\theta}_{1n}) = 1$, to mówimy, że estymator $\hat{\theta}_{1n}$ jest asymptotycznie efektywny.

Niech $\hat{\theta}_{2n}$ będzie innym estymatorem parametru θ_0, takim że $\sqrt{n}(\hat{\theta}_{2n} - \theta_0) \xrightarrow{D} N(0, \hat{\sigma}^2_{\hat{\theta}_{1n}})$. Wtedy, liczbę

$$e(\hat{\theta}_{n}, \hat{\theta}_{2n}) = \frac{\sigma^2_{\hat{\theta}_{2n}}}{\sigma^2_{\hat{\theta}_{1n}}}$$

nazywamy asymptotyczną efektywnością względną estymatora $\hat{\theta}_{1n}$ względem estymatora $\hat{\theta}_{2n}$.

Przykład 6

~~$X_1, \dots, X_n$ iid~~,

(i) $X_i = \theta + e_i \quad i = 1, \dots, n$, $e_1, \dots, e_n$ iid, $e_i \sim$ Laplace'a $\frac{1}{2}e^{-|x|}$

ENW parametru θ jest $\hat{\theta}_1 = me\{X_1, \dots, X_n\} = Q_2$, $I(\theta_0) = 1$.

$$\sqrt{n}(\hat{\theta}_1 - \theta) \xrightarrow{D} N(0, 1)$$

Niech $\hat{\theta}_2 = \bar{X}$, ~~CTT~~ CTG implikuje $\sqrt{n}(\hat{\theta}_2 - \theta) \xrightarrow{D} N(0, \sigma^2)$, gdzie $\sigma^2 = Var X_1 = Var(e_1 + \theta) = Var(e_1) = Ee_1^2 = \int_{-\infty}^{\infty} z^2 \frac{1}{2} e^{-|z|} dz =$

$\int_0^{\infty} z^2 e^{-z} dz = \Gamma(3) = 2$. Zatem $e(Me, \bar{X}) = \frac{2}{1} = 2$

Mediana dwa razy bardziej efektywna niż średnia.

(ii) $e_i \sim N(0,1)$. Wtedy

$$\sqrt{n}(\hat\theta_1 - \theta) \xrightarrow{\mathcal{D}} N\left(0, \frac{\pi}{2}\right), \quad \sqrt{n}(\hat\theta_2 - \theta) \sim N(0,1).$$

a $e(Me, \overline{X}) = \dfrac{1}{\pi/2} = \dfrac{2}{\pi} \approx 0.636 \approx \dfrac{1}{1.57}$ $\frac{1}{2f(0)}$

Wniosek (Asymptotycznie) średnia 1.57 razy bardziej efektywna niż mediana.

Wniosek 2 (Asymptotyczny przedział ufności dla θ)

Na mocy twierdzenia $\sqrt{n}(\hat\theta_n - \theta_0) \xrightarrow{\mathcal{D}} N\left(0, \frac{1}{I(\theta_0)}\right)$.

Stąd $1/\sqrt{nI(\theta_0)}$ jest asymptotyczną wariancją $\hat\theta_n$. Ponieważ $\hat\theta_n \xrightarrow{P} \theta_0$, a $I(\theta)$ jest ciągłą funkcją parametru θ, to $I(\hat\theta_n) \xrightarrow{P} I(\theta_0)$.

Na tej podstawie, asymptotyczny przedział ufności dla parametru θ na poziomie ufności $1-\alpha$ ma postać

$$\left(\hat\theta_n - z\left(1-\tfrac{\alpha}{2}\right)\frac{1}{\sqrt{nI(\hat\theta_n)}}, \hat\theta_n + z\left(1-\tfrac{\alpha}{2}\right)\frac{1}{\sqrt{nI(\hat\theta_n)}}\right), \text{ gdzie}$$

$z\left(1-\frac{\alpha}{2}\right) = \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$.

Wniosek 3

~~Załóżmy, że~~ Przy założeniach twierdzenia 2, załóżmy, że $g(x)$ jest funkcją ciągłą, różniczkowalną w punkcie θ_0 oraz $g'(\theta_0) \neq 0$. Wtedy

$$\sqrt{n}\left(g(\hat\theta_n) - g(\theta_0)\right) \xrightarrow{\mathcal{D}} N\left(0, \frac{g'(\theta_0)^2}{I(\theta_0)}\right).$$

3 Numeryczne wyznaczanie estymatorów NW z użyciem Metody Newtona

$\hat\theta^{(0)}$ - punkt startowy

$$\hat\theta^{(1)} = \hat\theta^{(0)} - \frac{\ell'(\hat\theta^{(0)})}{\ell''(\hat\theta^{(0)})}, \quad \text{itd.}$$

~~Zastosowanie np. w problemie~~

~~Problem~~ Przykład

$X_1, \dots, X_n$ iid $X_i \sim f(x,\theta) = \dfrac{\exp\{-(x-\theta)\}}{[1+\exp\{-(x-\theta)\}]^2}$ $x \in \mathbb{R}, \theta \in \mathbb{R}$.

Problem: ~~wyznacz~~ ENW istnieje, ale nie da się go jawnie wyznaczyć!