

Statystyka Matematyczna

Anna Janicka

wykład VII, 11.04.2016

WŁASNOŚCI ESTYMATORÓW, CZ. III

ESTYMACJA PRZEDZIAŁOWA, CZ. I

Plan na dzisiaj

1. Asymptotyczne własności estymatorów – cd.
 - zgodność – cd.
 - asymptotyczna normalność
 - asymptotyczna efektywność
2. Zgodność, asymptotyczna normalność i asymptotyczna efektywność ENW
3. Estymacja przedziałowa – przedziały ufności
 - wstęp
 - modele



Zgodność – przypomnienie

Niech $X_1, X_2, \dots, X_n, \dots$ będzie próbą IID, tzn. niezależnych zmiennych losowych o identycznych rozkładach. Niech $\hat{g}(X_1, X_2, \dots, X_n)$ będzie ciągiem estymatorów wielkości $g(\theta)$.

Estymator $\hat{g}$ jest **zgodny**, jeśli jest zbieżny do $g(\theta)$ według prawdopodobieństwa i **mocno zgodny**, jeśli jest zbieżny prawie na pewno.



Zgodność – przykłady

- Dla dowolnej rodziny rozkładów mających wartość oczekiwaną: średnia z próbki $\bar{X}_n$ jest zgodnym estymatorem wartości oczekiwanej $\mu(\theta) = E_\theta(X_1)$. Zbieżność wynika z MPWL.
- Dla dowolnej rodziny rozkładów mających wariancję: $S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ i $\hat{S}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ są zgodnymi estymatorami wariancji $\sigma^2(\theta) = \text{Var}_\theta(X_1)$. Zbieżność wynika z MPWL



Zgodność – przykłady/własności

- Estymator może być nieobciążony i nie być zgodny; np. $T_n(X_1, X_2, \dots, X_n) = X_1$ jako estymator $\mu(\theta) = E_\theta(X_1)$.
- Estymator może być obciążony i być zgodny; np. obciążony estymator wariancji j.w., lub np. dowolny zgodny estymator nieobciążony $+ 1/n$.



Asymptotyczna normalność

Estymator $\hat{g}(X_1, X_2, \dots, X_n)$ wielkości $g(\theta)$ jest **asymptotycznie normalny**, jeśli dla każdego $\theta \in \Theta$ istnieje $\sigma^2(\theta)$ takie, że gdy $n \rightarrow \infty$

$$\sqrt{n}(\hat{g}(X_1, X_2, \dots, X_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2(\theta))$$

Zbieżność wg. rozkładu, tzn. dla dowolnego a

$$\lim_{n \rightarrow \infty} P_{\theta} \left(\frac{\sqrt{n}}{\sigma(\theta)} (\hat{g}(X_1, X_2, \dots, X_n) - g(\theta)) \leq a \right) = \Phi(a)$$

lub równoważnie, rozkład $\hat{g}(X_1, X_2, \dots, X_n)$ jest dla dużych n zbliżony do rozkładu $N(g(\theta), \frac{\sigma^2}{n})$



Asymptotyczna normalność – co to znaczy

- ❑ Estymator asymptotycznie normalny jest zgodny (niekoniecznie mocno zgodny).
- ❑ Zawiera warunek *podobny* do nieobciążoności – wartość oczekiwana rozkładu asymptotycznego jest równa $g(\theta)$ (ale sam estymator *nie musi* być nieobciążony).
- ❑ Podobnie dla tzw. **wariancji asymptotycznej** definiowanej jako $\sigma^2(\theta)/n$ lub $\sigma^2(\theta)$ – jest to wariancja rozkładu asymptotycznego



Asymptotyczna normalność – czym nie jest

- Zazwyczaj dla estymatora asymptotycznie normalnego zachodzi:

$$E_{\theta} \hat{g}(X_1, X_2, \dots, X_n) \xrightarrow{n \rightarrow \infty} g(\theta)$$

$$n \operatorname{var} \hat{g}(X_1, X_2, \dots, X_n) \xrightarrow{n \rightarrow \infty} \sigma^2(\theta)$$

ale nie musi tak być, bo zbieżność wg rozkładu nie pociąga za sobą zbieżności wartości oczekiwanych ani wariancji



Asymptotyczna normalność – przykład

- Niech $X_1, X_2, \dots, X_n, \dots$ będzie próbą IID o średniej μ i wariancji σ^2 . Wówczas dla średniej z próby mamy z CTG

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{D} N(0, \sigma^2)$$

Asymptotyczna wariancja, σ^2/n , jest tu tożsama z wariancją estymatora.



Asymptotyczna normalność – jak liczyć

W wielu przypadkach przydatny jest lemat:

Metoda Delta. Jeśli dla ciągu zmiennych losowych T_n mamy $\sqrt{n}(T_n - \mu) \xrightarrow{D} N(0, \sigma^2)$ gdy $n \rightarrow \infty$ oraz $h: \mathbb{R} \rightarrow \mathbb{R}$ jest funkcją różniczkowalną w punkcie μ t.ż. $h'(\mu) \neq 0$, to

$$\sqrt{n}(h(T_n) - h(\mu)) \xrightarrow{D} N(0, \sigma^2 (h'(\mu))^2)$$

μ, σ^2 są funkcjami parametru θ

stosujemy zwykle wtedy, kiedy estymatory są funkcjami

statystyk T_n , których zbieżność łatwo wywnioskować z CTG

Asymptotyczna normalność – przykłady cd.

- W modelu wykładniczym: $ENW(\lambda) = \frac{1}{X}$
Z CTG mamy

$$\sqrt{n}(\bar{X} - \frac{1}{\lambda}) \xrightarrow{D} N(0, \frac{1}{\lambda^2})$$

więc z Lematu Delta mamy dla $h(t)=1/t$:

$$\sqrt{n}(\frac{1}{\bar{X}} - \lambda) \xrightarrow{D} N(0, \frac{1}{\lambda^2} \cdot (-\frac{1}{(1/\lambda)^2})^2)$$

a więc $\frac{1}{\bar{X}}$ jest asymptotycznie normalnym
(co za tym idzie: zgodnym) estymatorem λ .



Asymptotyczna efektywność

Dla asymptotycznie normalnego estymatora $\hat{g}(X_1, X_2, \dots, X_n)$ wielkości $g(\theta)$ możemy określić **asymptotyczną efektywność** jako

$$\text{as.ef}(\hat{g}) = \frac{(g'(\theta))^2 n}{\sigma^2(\theta) \cdot I_n(\theta)},$$

gdzie $\sigma^2(\theta)/n$ jest wariancją asymptotyczną, tj. mamy gdy $n \rightarrow \infty$

$$\sqrt{n}(\hat{g}(X_1, X_2, \dots, X_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2(\theta))$$



modyfikacja „zwykłej” efektywności do przypadku granicznego, z wariancją asymptotyczną zamiast zwykłej

$$\text{as.ef}(\hat{g}) = \frac{(g'(\theta))^2}{\sigma^2(\theta) \cdot I_1(\theta)}$$

Asymptotyczna efektywność względna

Asymptotyczna efektywność względna dla
asymptotycznie normalnych estymatorów
 $\hat{g}_1(X)$ i $\hat{g}_2(X)$

$$\text{as.ef}(\hat{g}_1, \hat{g}_2) = \frac{\sigma_2^2(\theta)}{\sigma_1^2(\theta)} = \frac{\text{as.ef}(\hat{g}_1)}{\text{as.ef}(\hat{g}_2)}$$

Uwaga. Estymator mniej (asymptotycznie) efektywny może mieć inne cechy, które sprawiają, że ma przewagę nad estymatorem bardziej (asymptotycznie) efektywnym.



Asymptotyczna efektywność względna – Przykład 1

as.ef

Estymacja prawdopodobieństwa braku szkód
w modelu Poissona: $\text{Poiss}(\theta)$ (typowy model
aktuarialny dla liczby szkód w poszczególnych
latach dla polisy, albo dla grupy polis)

$g(\theta) = e^{-\theta} = P(X_i = 0)$ – pr-stwo braku szkód

$$\hat{g}_1(X) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i=0\}} \quad \hat{g}_2(X) = e^{-\bar{X}} = ENW(g(\theta))$$

$$\sqrt{n}(\hat{g}_1 - g(\theta)) \xrightarrow{D} N(0, e^{-\theta}(1 - e^{-\theta})) \quad \text{z CTG dla schematu Bernoulliego}$$

$$\sqrt{n}(\hat{g}_2 - g(\theta)) \xrightarrow{D} N(0, \theta \cdot (-e^{-\theta})^2) \quad \text{z Lematu Delta}$$

$$\rightarrow \text{as.ef}(\hat{g}_1, \hat{g}_2) = \frac{\theta e^{-2\theta}}{e^{-\theta}(1 - e^{-\theta})} = \frac{\theta}{e^{\theta} - 1} < 1$$



Asymptotyczna efektywność względna – Przykłady 2: lepsza średnia czy mediana z próbki?

[as.ef](#)

To zależy od rozkładu!

a) model normalny $N(\mu, \sigma^2)$:

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{D} N(0, \sigma^2) \quad \text{as.ef}(\hat{m}, \bar{X}) = \frac{2}{\pi} < 1$$

$$\sqrt{n}(\hat{m} - \mu) \xrightarrow{D} N(0, \frac{\pi\sigma^2}{2})$$

b) model Laplace'a $\text{Lapl}(\mu, \lambda)$

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{D} N(0, \frac{2}{\lambda^2}) \quad \text{as.ef}(\hat{m}, \bar{X}) = 2 > 1$$

$$\sqrt{n}(\hat{m} - \mu) \xrightarrow{D} N(0, \frac{1}{\lambda^2})$$

c) niektóre rozkłady nie mają średniej...

Twierdzenie: Dla próbki z rozkładu ciągłego o gęstości $f(x)$, mediana próbkowa jest estymatorem asymptotycznie normalnym dla mediany m (o ile gęstość jest ciągła i $\neq 0$ w punkcie m):

$$\sqrt{n}(\hat{m} - m) \xrightarrow{D} N(0, \frac{1}{4(f(m))^2})$$



Zgodność ENW

Niech $X_1, X_2, \dots, X_n, \dots$ będą próbą z gęstości $f_\theta(x)$, gdzie $\theta \in \Theta \subseteq \mathbb{R}$ jest otwarty, oraz:

- Wszystkie gęstości f_θ mają ten sam nośnik
- równanie $\frac{d}{d\theta} \ln L(\theta) = 0$ ma dokładnie jedno rozwiązanie $\hat{\theta}$,

to $\hat{\theta}$ jest ENW(θ) i jest to estymator zgodny

Uwaga. Nawet „porządne” estymatory ENW nie muszą być nieobciążone!



Asymptotyczna normalność ENW

Niech $X_1, X_2, \dots, X_n, \dots$ będą próbą z gęstości $f_\theta(x)$, gdzie $\theta \in \Theta \subseteq \mathbb{R}$ jest otwarty, a $\hat{\theta}$ jest zgodnym estymatorem największej wiarygodności (np. spełnia założenia poprzedniego twierdzenia), oraz

- istnieje $\frac{d^2}{d\theta^2} \ln L(\theta)$
- można wyliczyć informację Fishera, $0 < I_1(\theta) < \infty$
- można zmieniać kolejność całkowania po x i różniczkowania po θ

to $\hat{\theta}$ jest asymptotycznie normalny oraz

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N(0, \frac{1}{I_1(\theta)})$$



Asymptotyczna normalność ENW

Przy założeniach j.w., jeśli dodatkowo $g: \mathbb{R} \rightarrow \mathbb{R}$ jest funkcją różniczkowalną w punkcie θ , t.ż. $g'(\theta) \neq 0$, a $\hat{g}(X_1, X_2, \dots, X_n)$ jest ENW($g(\theta)$), to

$$\sqrt{n}(\hat{g}(X_1, X_2, \dots, X_n) - g(\theta)) \xrightarrow{D} N(0, \frac{(g'(\theta))^2}{I_1(\theta)})$$



Asymptotyczna efektywność ENW

Jeśli spełnione są założenia poprzednich twierdzeń, to estymator największej wiarygodności (parametru θ oraz dla $g(\theta)$) jest asymptotycznie efektywny.



Asymptotyczna normalność i efektywność ENW – przykłady

- W modelu liczby szkód:

[Przykład 1](#)

$$\hat{g}_2(X) = e^{-\bar{X}} = ENW(g(\theta))$$

jest asymptotycznie efektywny

- W modelu normalnym: średnia jest asymptotycznie efektywnym estymatorem μ
- W modelu Laplace'a: mediana jest asymptotycznie efektywnym estymatorem μ

[Przykłady 2](#)



Podsumowanie: badanie własności estymatorów

- ☐ obciążoność
 - ☐ wariancja
 - ☐ ryzyko
 - ☐ efektywność
-
- ☐ zgodność
 - ☐ asymptotyczna normalność
 - ☐ asymptotyczna efektywność
-



Estymacja przedziałowa – przedziały ufności

- Estymujemy parametr zadaną dokładnością
- Zamiast pojedynczego oszacowania podajemy dolną i górną granicę dla oszacowania (prawdziwy parametr będzie się w nich mieścić z *zadany* prawdopodobieństwem)



Przedział ufności

Niech $g(\theta)$ będzie funkcją nieznanego parametru θ , zaś $\bar{g} = \bar{g}(X_1, X_2, \dots, X_n)$ oraz $\underline{g} = \underline{g}(X_1, X_2, \dots, X_n)$ będą statystykami

Wówczas $[\underline{g}, \bar{g}]$ to **przedział ufności** dla $g(\theta)$ na poziomie ufności $1-\alpha$, jeśli dla każdego θ

$$P_{\theta}(\underline{g}(X_1, X_2, \dots, X_n) \leq g(\theta) \leq \bar{g}(X_1, X_2, \dots, X_n)) \geq 1 - \alpha$$



Przedział ufności – użycie i interpretacja

- Typowo: α jest małą liczbą, np. $1-\alpha = 0,95$ albo $1-\alpha = 0,99$
- Warunek z definicji oznacza: losowy przedział $[\underline{g}, \bar{g}]$ zawiera nieznaną liczbę $g(\theta)$ z dużym prawdopodobieństwem (zadany).
- Jeśli obliczymy *realizację* przedziału ufności (np. $\underline{g} = 1, \bar{g} = 3$) to już NIE możemy mówić o tym, że nieznaną parametr zawiera się w tym przedziale z prawdopodobieństwem $1-\alpha$!



Przedział ufności – konstrukcja

- Z definicji, przedział ufności zależy od rozkładu prawdopodobieństwa, z jakim mamy do czynienia
- Najczęściej rozważa się próbki pochodzące z rozkładów normalnych (takie rozkłady występują „w przyrodzie” najczęściej)



Przedział ufności – konstrukcja cd.

- wygodna metoda: szukamy zmiennych losowych zależnych od próby i funkcji parametrów, których *rozkłady* nie zależą od wartości nieznanymi parametrów – tzw. **funkcji centralnych**

- Jeśli $U = U(X_1, X_2, \dots, X_n, \theta)$ – funkcja centralna, to szukamy przedziału ufności postaci $[a, b]$ t.ż.

$$P_{\theta}(a \leq U \leq b) \geq 1 - \alpha$$

- Najczęściej dodatkowo szukamy przedziałów „symetrycznych”

$$P_{\theta}(U < a) \leq \frac{\alpha}{2}, \quad P_{\theta}(U > b) \leq \frac{\alpha}{2}$$



Najczęściej wykorzystywane modele

- Model I (normalny): przedział ufności dla średniej, wariancja znana
- Model II (normalny): przedział ufności dla średniej, wariancja nieznana
- Model II (normalny): przedział ufności dla wariancji
- Model III (asymptotyczny): przedział ufności dla średniej
- Model IV (asymptotyczny): przedział ufności dla odsetka
- Model asymptotyczny: przedział ufności oparty o ENW



Przedział ufności dla średniej – Model I

Model normalny: $X_1, X_2, \dots, X_n$ są próbą IID z rozkładu $N(\mu, \sigma^2)$, przy czym σ^2 jest **znane**.

Przedział ufności dla μ , na poziomie $1-\alpha$:

$$\left[\bar{X} - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

gdzie $u_{1-\alpha/2}$ jest kwantylem rzędu $1-\alpha/2$ z rozkładu $N(0,1)$



Przedział ufności – Model I, uzasadnienie:

Punktowy estymator dla μ : $ENW(\mu) = \bar{X}$

Znamy rozkład $\bar{X}$:

$$\bar{X} \sim N(\mu, \sigma^2/n),$$

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

rozkład nie
zależy od μ
-- funkcja
centralna

Korzystamy z funkcji centralnej. Chcemy: przedział ufności symetryczny wokół estymatora punktowego (rozkład funkcji centralnej jest symetryczny wokół 0).

Mamy:

$$P_{\mu} \left(\left| \sqrt{n}(\bar{X} - \mu) / \sigma \right| \leq u \right) = \Phi(u) - \Phi(-u) = 2\Phi(u) - 1$$
$$= 1 - \alpha$$

$$\text{skąd } u = u_{1-\alpha/2}$$



Przedział ufności – Model I, własności

- ❑ Błąd oszacowania: $d = u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$
- ❑ Długość przedziału ufności: $2d$
- ❑ Liczebność próby wystarczająca do uzyskania zadanej *precyzji* (błędu) d :

$$n \geq \frac{\sigma^2 u_{1-\alpha/2}^2}{d^2}$$



Przedział ufności dla średniej – Model II

Model normalny: $X_1, X_2, \dots, X_n$ są próbą IID z rozkładu $N(\mu, \sigma^2)$, przy czym σ^2 jest **nieznane**

Przedział ufności dla μ , na poziomie $1-\alpha$:

$$\left[\bar{X} - t_{1-\alpha/2}(n-1) \frac{S}{\sqrt{n}}, \bar{X} + t_{1-\alpha/2}(n-1) \frac{S}{\sqrt{n}} \right]$$

gdzie $t_{1-\alpha/2}(n-1)$ jest kwantylem rzędu $1-\alpha/2$ z rozkładu t -Studenta z $n-1$ stopniami swobody $t(n-1)$, a $S = \sqrt{S^2}$ dla nieobciążonego estymatora wariancji S^2 .



Przedział ufności – Model II, uzasadnienie:

Punktowy estymator dla μ : $ENW(\mu) = \bar{X}$

Znamy rozkład $\bar{X}$:

$$\bar{X} \sim N(\mu, \sigma^2/n), \quad \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1), \quad T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

Korzystamy z funkcji centralnej T . Chcemy: przedział ufności symetryczny wokół estymatora punktowego (rozkład T jest symetryczny wokół 0). Mamy:

$$P_{\mu, \sigma} \left(\left| \sqrt{n}(\bar{X} - \mu) / S \right| \leq t \right) = 1 - \alpha$$

$$\text{skąd } t = t_{1-\alpha/2}(n-1)$$



Przedział ufności – Model II, własności

- ❑ Błąd oszacowania: $d = t_{1-\alpha/2}(n-1) \frac{S}{\sqrt{n}}$
- ❑ Długość przedziału ufności: $2d$
- ❑ Liczebność próby wystarczająca do uzyskania zadanej *precyzji* (błędu) d :
do wyznaczenia na podstawie tzw. dwuetapowej procedury Steina – musimy najpierw wstępnie oszacować wariancję



Dwuetapowa procedura Steina

1. Pobieramy wstępną próbkę $X_1, X_2, \dots, X_{n_0}$ na jej podstawie obliczamy estymator wariancji

$$S_0^2 = \frac{1}{n_0 - 1} \sum_{i=1}^{n_0} (X_i - \bar{X}_0)^2$$

2. Sprawdzamy, czy próbka spełnia żądany warunek: obliczamy $k = \frac{S_0^2 [t_{1-\alpha/2}(n_0 - 1)]^2}{d^2}$

a) jeśli $n_0 \geq k$ to za przedział ufności przyjmujemy

$$\left[\bar{X}_0 - t_{1-\alpha/2}(n_0 - 1) \frac{S_0}{\sqrt{n_0}}, \bar{X}_0 + t_{1-\alpha/2}(n_0 - 1) \frac{S_0}{\sqrt{n_0}} \right]$$

b) jeśli $n_0 < k$ to wybieramy $n \geq k$ i dołosowujemy $X_{n_0+1}, X_{n_0+2}, \dots, X_n$, obliczamy średnią z *połączonej* próbki $X_1, X_2, \dots, X_n$, i za przedział ufności przyjmujemy

$$\left[\bar{X} - t_{1-\alpha/2}(n_0 - 1) \frac{S_0}{\sqrt{n}}, \bar{X} + t_{1-\alpha/2}(n_0 - 1) \frac{S_0}{\sqrt{n}} \right]$$



Przedział ufności dla wariancji – Model II

Model normalny: $X_1, X_2, \dots, X_n$ są próbą IID z rozkładu $N(\mu, \sigma^2)$.

Przedział ufności dla σ^2 , na poziomie $1-\alpha$:

$$\left[\frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)} \right]$$

gdzie $\chi_{\alpha/2}^2(n-1)$ oraz $\chi_{1-\alpha/2}^2(n-1)$ są kwantylami rzędu $\alpha/2$ oraz $1-\alpha/2$,
odpowiednio, z rozkładu chi-kwadrat z $n-1$
stopniami swobody



Przedział ufności – Model II, uzasadnienie

Punktowy estymator dla σ^2 : $ENW(\sigma^2) = S^2$

Znamy rozkład: $U = \frac{(n-1)}{\sigma^2} S^2 \sim \chi^2(n-1)$

Korzystamy z funkcji centralnej U . Rozkład chi-kwadrat nie jest symetryczny. Chcemy „symetryczny” przedział ufności, tj.

szukamy takiego przedziału $[a, b]$ że

$$P_{\sigma^2}(U < a) = \frac{\alpha}{2}, \quad P_{\sigma^2}(U > b) = \frac{\alpha}{2}$$

a więc $a = \chi_{\alpha/2}^2(n-1)$ oraz $b = \chi_{1-\alpha/2}^2(n-1)$





WARSAW UNIVERSITY
Faculty of Economic Sciences