NMFM301 – Statistika pro finační matematiky

1. Zápočtová písomná práca | 03. 11. 2021 | Varianta A

(Celkově 20 bodů | Čas: 90 minút)

1. *(12 bodů)* Uvažujte náhodný výběr $X_1, \dots, X_n$ s hustotou danou předpisem

$$f_\theta(x) = \frac{x}{\theta^2} \exp\left\{-\frac{x^2}{2\theta^2}\right\} \cdot \mathbb{I}_{\{x>0\}}, \quad \text{pro } \theta \in (0, \infty).$$

(a) Najděte odhad $\widetilde{\theta}_n$ parametru θ metodou maximální věrohodnosti. [2]

Vierohodnostná funkcia má tvar

$$L(\mathbb{X}, \theta) = \prod_{i=1}^n \left[\frac{X_i}{\theta^2} \exp\left\{-\frac{X_i^2}{2\theta^2}\right\}\right] = \theta^{-2n} \left(\prod_{i=1}^n X_i\right) \exp\left\{-\frac{\sum_{i=1}^n X_i^2}{2\theta^2}\right\}$$

a logaritmická vierohodnostná funkcia je

$$l(\mathbb{X}, \theta) = (-2n) \log \theta + \sum_{i=1}^n \log X_i - \frac{\sum_{i=1}^n X_i^2}{2\theta^2}.$$

Maximálne vierohodný odhad získame riešením skórovej rovnice

$$\frac{\partial}{\partial \theta} l(\mathbb{X}, \theta) = \frac{-2n}{\theta} + \frac{\sum_{i=1}^n X_i^2}{\theta^3} := 0.$$

Jednoduchými úpravami získame maximálne vierohodný odhad neznámeho parametru $\theta > 0$ ve tvaru

$$\widetilde{\theta}_n = \sqrt{\frac{\sum_{i=1}^n X_i^2}{2n}},$$

a keďže je zadaná hustota exponenciálneho typu (t.j., splnené sú podmienky regularity), tak sa jedná naozaj o globálne maximum vierohodnostnej funkce. ■

(b) Spočtěte asymptotické rozdělení odhadu $\widetilde{\theta}_n$ a sestrojte přibližný interval spolehlivosti pro $\alpha = 0.05$ (využijte fakt, že $\mathbb{E}X_i^2 = 2\theta^2$). [2]

Maximálne vierohodné odhady sú za platnosti podmienky regularity asymptoticky normálne a špeciálne platí, že

$$\sqrt{n}(\widetilde{\theta}_n - \theta) \xrightarrow[n\to\infty]{\mathcal{D}} N(0, I^{-1}(\theta)),$$

kde $I(\theta) = I_n(\theta)/n$ a $I_n(\theta)$ je Fisherova informace o parametru $\theta > 0$ obsiahnutá v náhodnom výbere o rozsahu $n \in \mathbb{N}$. Zároveň platí, že

$$\begin{aligned} I_n(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} l(\mathbb{X}, \theta)\right] &= -E\left[\frac{\partial}{\partial \theta}\left(\frac{-2n}{\theta} + \frac{\sum_{i=1}^n X_i^2}{\theta^3}\right)\right] \\ &= -E\left[\frac{2n}{\theta^2} - \frac{3\sum_{ii=1}^n X_i^2}{\theta^4}\right] = -\frac{2n}{\theta^2} + \frac{3n(2\theta^2)}{\theta^4} = \frac{4n}{\theta^2}, \end{aligned}$$

pričom sme v predposlednom kroku (namiesto počítania strednej hodnoty EX_i^2 z definície pomocou integrálu) využili ponúkaný "hint", že $EX_i^2 = 2\theta^2$.

Asymtotické rozdelenie maximálne vierohodného odhadu $\widetilde{\theta}_n$ je teda

$$\sqrt{n}(\widetilde{\theta}_n - \theta) \xrightarrow[n\to\infty]{\mathcal{D}} N(0, \theta^2/4)$$

a pre konkrétne $\alpha \in (0,1)$ dostaneme (približný) interval spoľahlivosti v tvare

$$P\left[\theta \in (\widetilde{\theta}_n \pm u_{1-\alpha/2}\theta/2\sqrt{n}\right] \approx 1 - \alpha.$$

■

(c) Najděte odhad $\widehat{\theta}_n$ parametru θ pomoci momentové metody. [2]

Tu sa síce ponúka možnosť spočítať strednú hodnotu (t.j. integrovať pomocou per-partes), ale najjednoduchšie je opäť využiť ponúkaný "hint", že $EX_i^2 = 2\theta^2$. Neznámy parameter teda dávame do vzťahu s druhým momentom a získame

$$\theta = \sqrt{\frac{EX_i^2}{2}}.$$

Hustota je nenulová na kladnej poloose, pre neznámy parameter platí, že $\theta > 0$, teda všetko je správne definované. Druhý moment konzistentne odhadneme výberovým priemerom

$$\widehat{EX_i^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

a dosadíme do predchádzajúceho výrazu. Momentový odhad neznámeho parametru $\theta > 0$ je definovaný predpisom

$$\widehat{\theta}_n = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}.$$

Odhad získaný momentovou metódou je teda totožný s odhadom získaným metódou maximálnej vierohodnosti. ■

(d) Pomoci centální limitní věty a transformace $g(\theta) = \sqrt{\theta/2}$ nájděte asymptotické rozdělení momentového odhadu $\widehat{\theta}_n$. [2]

V prvom rade je dobré si uvedomiť, že ponúkaná transformácia je presne tou transformáciou, ktorá dáva do súvislosti výberový priemer náhodných veličín $X_1^2, \ldots, X_n^2$ (teda empirický odhad druhého momentu $\widehat{EX_i^2}$) a neznámy parameter $\theta > 0$. Použijeme preto CLV pre nezávislé a stejně rozdelené náhodné veličiny $X_1^2, \ldots, X_n^2$ a následne ponúkanú transformáciu. Z CLV dostaneme, že

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - EX_i^2\right) \xrightarrow[n\to\infty]{\mathcal{D}} N(0, VarX_i^2),$$

pričom platí, že $EX_i^2 = 2\theta^2$ ("hint") a $Var(X_i^2) = EX_i^4 - (EX_i^2)^2$. Stačí teda spočítať štvrtý moment

$$EX_i^4 = \int_0^\infty \frac{x^5}{\theta^2} \exp\left\{-\frac{x^2}{2\theta^2}\right\} \mathrm{d}x = \left| \begin{array}{l} \text{subst.} \\ x^2 = y \\ 2x\mathrm{d}x = \mathrm{d}y \end{array} \right| = \int_0^\infty \frac{y^2}{2\theta^2} \exp\left\{-\frac{y}{2\theta^2}\right\} \mathrm{d}y.$$

V tomto momente je výhodné si uvedomiť, že posledný integrál je vlastne druhý moment nejakej náhodnej veličiny, ktorá ma exponenciálne rozdelenie s parametrom $\lambda = 1/2\theta^2$ (a strednou hodnotou $2\theta^2$ a rozptylom $4\theta^4$ a teda druhým momentom $4\theta^4 + (2\theta)^2$, keďže všeobecne platí, že $VarX = EX^2 - (EX)^2$). Preto stačí využiť znalosť strednej hodnoty a rozptylu v exponenciálnom rozdelení a integrál nie je nutné manuálne počítať. Preto dostaneme, že $EX_i^4 = 4\theta^2 + (2\theta)^2 = 8\theta^4$ a $VarX_i^2 = 8\theta^4 - (2\theta^2)^2 = 4\theta^4$.
Máme teda asymptotické rozdelenie (vďaka CLV) v tvare

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - EX_i^2\right) \xrightarrow[n\to\infty]{\mathcal{D}} N(0, 4\theta^4)$$

a transformácia v zadaní presne transformuje parameter strednej hodnoty na neznýmy parameter $\theta > 0$, resp. výberový priemer $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ na momentový odhad neznámeho parametru, teda

$$g\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2} = \widehat{\theta}_n.$$

Použitím transformácie všeobecne dostaneme

$$\sqrt{n}\left[g\Big(\frac{1}{n}\sum_{i=1}^{n} X_i^2\Big) - g(EX_i^2)\right] \xrightarrow[n\to\infty]{\mathcal{D}} N(0, [g'(EX_i^2)]^2 \cdot VarX_i^2))$$

a špecificky, pre konkrétny tvar transformačnej funkcie

$$\sqrt{n}\left[\widehat{\theta}_n - \theta\right] \xrightarrow[n\to\infty]{\mathcal{D}} N(0, [g'(EX_i^2)]^2 \cdot VarX_i^2)).$$

Keďže $[g'(\theta)]^2 = [1/4 \cdot \sqrt{2/\theta}]^2 = 1/8\theta$, tak $[g'(EX_i^2)]^2 = 1/16\theta^2$ a celkovo teda

$$\sqrt{n}\left[\widehat{\theta}_n - \theta\right] \xrightarrow[n\to\infty]{\mathcal{D}} N(0, \theta^2/4).$$

■

(e) Porovnejte asymptotický rozptyl odhadov $\widetilde{\theta}_n$ a $\widehat{\theta}_n$. [2]

Rozptyl odhadu metódou maximálnej vierohodnosti a rozptyl odhadu momentovou metódou sú rovnaké. ■

(f) Použijte transformaci $Y_i = X_i^2$ a pro $\alpha = 0.05$ sestrojte přesný interval spolehlivosti pro neznámy parametr $\theta > 0$. [2]

Cheme použiť transformáciu $t: \ x \longrightarrow x^2 =: y$. Príslušná inverzná transformácia je $t^{-1}: \ y \longrightarrow \sqrt{y} =: x$. Keďže hustota je nenulová pouze na kladnej časti reálnej osi, tak je transformácia prostá. Hustota náhodnej veličiny $Y_i = X_i^2$ je teda definovaná predpisom

$$f_Y(y) = f_\theta(t^{-1}(y)) \cdot |(t^{-1})'(y)| = \frac{\sqrt{y}}{2\theta^2}\exp\left\{-\frac{(\sqrt{y})^2}{2\theta^2}\right\} \cdot \left|\frac{1}{2\sqrt{y}}\right| = \frac{1}{2\theta^2}\exp\left\{-\frac{y}{2\theta^2}\right\},$$

pre $y > 0$ a hustota je definovaná nulou inak. Jedna sa o hustotu exponenciálneho rozdelenia s parametrom $\lambda = 1/2\theta^2$ (stredná hodnota je $2\theta^2$). Platí teda, že náhodné veličiny $Y_1/\theta^2, \ldots, Y_n/\theta^2$ majú exponenciálne rozdelenie s parametrom $\lambda = 1/2$ (t.j., stredná hodnota 2). Súčet $n \in \mathbb{N}$ takýchto nezávislých náhodných veličín má tzv. Erlangovo rozdelenie s parametrami n a $1/2$ (čo je vlastne Gamma rozdelene s rovnakými parametrami) a zo vzťahu medzi Gamma rozdelením a χ^2 rozdelením dostaneme, že

$$\frac{1}{\theta^2}\sum_{i=1}^{n} Y_i \sim \chi^2_{2n}.$$

Použitím kvantilov χ^2 rozdelenia s $2n$ stupňami voľnosti (pre stručnosť ich označme ako $q_{\alpha/2}$ a $q_{1-\alpha/2}$) dostaneme

$$P\left[q_{\alpha/2} \leq \frac{1}{\theta^2}\sum_{i=1}^{n} Y_i \leq q_{1-\alpha/2}\right] = 1-\alpha$$

a ekvivalentnými úpravami získame (presný) interval spoľahlivosti pre neznámy parameter $\theta > 0$. ■

2. *(8 bodů)* Uvažujte náhodný výběr $X_1, \ldots, X_{2n}$ z rozdělení s hustotou

$$f(x;\theta) = \theta x^{\theta-1}\mathbb{I}_{\{x\in(0,1)\}}, \quad x \in \mathbb{R};$$

kde $\theta > 0$ je neznámý parametr. Uvažujte transformaci $W_i = -\log X_i$.

(a) Určete rozdělení

$$Y = \sum_{i=1}^{n} W_i.$$

[2]

Uvažovaná transformácia je $t:\ x \to -\log x =: w$ a príslušná inverzná transformácia je $t^{-1}:\ w \to e^{-w} =: x$. Náhodná veličina W_i má teda rozdelenie dané hustotou

$$f_W(w) = f(t^{-1}(w);\theta)\cdot|(t^{-1})'(w)| = \theta(e^{-w})^{\theta-1}\cdot|-e^{-w}| = \theta e^{-\theta w},$$

pre $w > 0$ a hustota je definovaná nulou inak. Jedná sa teda o exponenciálne rozdelenie s parametrom $\lambda = \theta > 0$ (stredná hodnota je $1/\theta$). Náhodná veličina Y (čo je vlastne súčet nezávislých náhodných veličín s exponenciálnym rozdelením) má preto Erlangovo rozdelenie s parametrami $\lambda = \theta$ a $n \in \mathbb{N}$. Hustota náhodnej veličiny Y je daná predpisom

$$f_Y(y) = \frac{\theta^k y^{k-1}e^{-\theta y}}{(k-1)!},$$

pre $y > 0$ a nula inak. To je tiež gamma rozdelenie s parametrami $\theta > 0$ a $n \in \mathbb{N}$ (t.j., $Y \sim \Gamma(\theta, n)$). ■

(b) Určete rozdělení

$$Z = \min_{i=n+1,\ldots,2n} W_i.$$

[2]

Už vieme, že náhodné veličiny W_i majú exponenciálne rozdelenie s parametrom $\theta > 0$ (stredná hodnota $1/\theta$) a sú vzájomne nezávislé. Príslušná distribučná funkcia (náhodnej veličiny W_i) je $F_W(w) = 1 - e^{-\theta w}$ pre $w \geq 0$ a $F_W(w) = 0$ pre $w < 0$. Prvá pořadová statistika—t.j., minimum $Z = \min_{i=n+1,\dots,2n} W_i$—má distribučnú funkciu

$$F_Z(z) = 1 - [1 - F_W(z)]^n = 1 - [e^{-\theta z}]^n = 1 - e^{n\theta z},$$

pre $z \geq 0$ a $F_Z(z) = 0$ pre $z < 0$. Jedná sa vlastne o exponenciálne rozdelene s parametrom $\lambda = n\theta$ (stredná hodnota je $1/n\theta$). ■

(c) Odvoďte rozdělení

$$\frac{nZ}{Y + nZ}.$$

[2]

Náhodné veličiny Y a Z sú nezávislé, pretože pôvodné veličiny $X_1, \dots, X_{2n}$ tvorili náhodný výber (t.j., nezávislé a stejně rozdelené náhodné veličiny) a náhodná veličina Y je definovaná pomocou prvej poloviny pozorovaní, t.j., náhodných veličín $X_1, \dots, X_n$, zatiaľ čo náhodná veličina Z je definovaná pouze pomocou druhej poloviny pozorovaní, t.j., pozorovaní $X_{n+1}, \dots, X_{2n}$. Navyše, keďže $Z \sim Exp(n\theta)$ (stredná hodnota je $1/n\theta$), tak potom $nZ \sim Exp(\theta)$ (t.j., stredná hodnota je $\frac{1}{\theta}$).
Združené rozdelenie náhodného vektoru $(nZ, Y)^\top$ je preto definované ako súčin marginálnych rozdelení Y a nZ, to znamená, že

$$(nZ, Y)^\top \quad \sim \quad f_{(nZ,Y)}(z, y) = f_{nZ}(z) f_Y(y) = \theta e^{-\theta z} \cdot \frac{\theta^k y^{k-1} e^{-\theta y}}{(k-1)!},$$

pre $z > 0$ a $y > 0$ a hustota je definovaná nulou inak.
Využijeme transformáciu

$$t : \begin{pmatrix} nZ \\ Y \end{pmatrix} \to \begin{pmatrix} nZ \\ \frac{nZ}{Y+nZ} \end{pmatrix} =: \begin{pmatrix} U \\ V \end{pmatrix}$$

s príslušnou inverznou transformáciou

$$t^{-1} : \begin{pmatrix} U \\ V \end{pmatrix} \to \begin{pmatrix} U \\ \frac{U(1-V)}{V} \end{pmatrix} =: \begin{pmatrix} Z \\ Y \end{pmatrix}$$

Jakobián inverzného zobrazenia je

$$|J_{t^{-1}}| = \begin{vmatrix} 1 & 0 \\ \frac{(1-V)}{V} & -\frac{U}{V^2} \end{vmatrix} = \frac{U}{V^2}.$$

Združená hustota náhodného vektoru $(U, V)^\top$ je teda

$$\begin{aligned} f_{(U,V)}(u, v) &= f_{nZ}(t_1^{-1}(u, v)) f_Y(t_2^{-1}(u, v)) \cdot |J_{t^{-1}}| \\ &= \theta e^{-\theta u} \cdot \frac{\theta^k (u(1-v)/v)^{(k-1)} e^{-\theta(u(1-v)/v)}}{(k-1)!}. \end{aligned}$$

Požadované rozdelenie náhodnej veličiny $\frac{nZ}{Y+nZ}$ by sa získalo integrovaním združenej hustoty, teda

$$f_V(u) = \int_{\mathbb{R}} f_{(U,V)}(u, v) \mathrm{d}u.$$

Pre účely správneho riešenia bolo postačujúce vyjadriť vhodnú transformáciu združeného vektoru a uviesť, že požadované rozdelenie získame vhodným preintegrovaním. ■

(d) Pomoci CLV spočtěte přibližne pravděpodobnost

$$P[Y \leq \sqrt{n}]$$

[2]

Náhodná veličina Y je definovaná ako súčet $n \in \mathbb{N}$ nezávislých a rovnako rozdelených náhodných veličín $W_1, \ldots, W_n$ s exponenciálnym rozdelením s parametrom $\lambda = \theta$ (t.j., so strednou hodnotou $1/\theta$ a rozptylom $1/\theta^2$).

S využitím CLV preto postupnými ekvivalentnými úpravami dostaneme

$$\begin{aligned} P[Y \leq \sqrt{n}] &= P\Big[\frac{1}{n}\sum_{i=1}^{n} W_i \leq \frac{\sqrt{n}}{n}\Big] = P\Big[\overline{W}_n - EW_1 \leq \frac{1}{\sqrt{n}} - EW_1\Big] \\ &= P\Big[\sqrt{n}(\overline{W}_n - EW_1) \leq (1 - \sqrt{n}/\theta)\Big] = P\Big[\sqrt{n}\frac{(\overline{W}_n - EW_1)}{\sqrt{VarW_1}} \leq (\theta - \sqrt{n})\Big], \end{aligned}$$

pričom náhodná veličina na ľavej strene v poslednej pravdepodobnosti má z CLV asymptoticky normálne rozdelenie s nulovou strednou hodnotou a jednotkovým rozptylom. Označme distribučnú funkciu štandardizovaného normálneho rozdelenia (klasické značenie) ako $\Phi(x) = P[X \leq x]$, kde $X \sim N(0,1)$.

Teda platí (asymptoticky, alebo približne), že

$$P[Y \leq \sqrt{n}] \approx \Phi(\theta - \sqrt{n}).$$

■

1. $X_1, \dots, X_n \sim f_\theta(x) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} \mathbb{1}\{x>0\}$, $\theta \in (0, \infty)$

(a) odhad $\tilde{\theta}_n$ parametru θ max. vierohodnosťou

$$L_n(\theta) = \prod_{i=1}^n \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} = \theta^{-2n} \prod_{i=1}^n x_i e^{-\frac{x_i^2}{2\theta^2}} = \theta^{-2n} \prod_{i=1}^n x_i \, e^{-\frac{\sum x_i^2}{2\theta^2}}$$

$$\ell_n(\theta) = \log[L_n(\theta)] = \log\left[\theta^{-2n} \prod_{i=1}^n x_i \, e^{-\frac{\sum x_i^2}{2\theta^2}}\right] = -2n\log\theta + \sum_{i=1}^n \log x_i - \frac{\sum x_i^2}{2\theta^2}$$

$$U_n(\theta) = \frac{\partial}{\partial\theta}[\ell_n(\theta)] = \frac{\partial}{\partial\theta}\left[-2n\log\theta + \sum_{i=1}^n \log x_i - \frac{\sum x_i^2}{2\theta^2}\right] = \frac{-2n}{\theta} + \frac{\sum x_i^2}{\theta^3}$$

$$\frac{-2n}{\theta} + \frac{\sum x_i^2}{\theta^3} = 0$$
$$\frac{\sum x_i^2}{\theta^3} = \frac{2n}{\theta}$$
$$\theta \sum x_i^2 = 2n\theta^3$$
$$\sum x_i^2 = 2n\theta^2$$
$$\theta^2 = \frac{\sum x_i^2}{2n}$$
$$\theta = \sqrt{\frac{\sum x_i^2}{2n}}$$

(b) asymptotické rozdelenie odhadu $\tilde{\theta}_n$ + približný interval spoľahlivosti pro $\alpha = 0{,}05$
(využiť $EX_i^2 = 2\theta^2$)

$$I_n(\theta) = -E\left[\frac{\partial}{\partial\theta}\left(\frac{-2n}{\theta} + \frac{\sum x_i^2}{\theta^3}\right)\right] = -E\left[\frac{2n}{\theta^2} - \frac{3}{\theta^4}\sum_{i=1}^n x_i^2\right] = \frac{-2n}{\theta^2} + \frac{3}{\theta^4} E\Big[\overbrace{\sum_{i=1}^n x_i^2}^{n2\theta^2}\Big] = \frac{-2n}{\theta^2} + \frac{6n}{\theta^2} = \frac{4n}{\theta^2}$$

$$I(\theta) = \frac{1}{n}\frac{4n}{\theta^2} = \frac{4}{\theta^2} \; ; \quad I^{-1}(\theta) = \frac{\theta^2}{4}$$

$$\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{\mathcal{D}} N\left(0, \frac{\theta^2}{4}\right) \rightsquigarrow P\left[\theta \in \left(\tilde{\theta}_n \pm u_{1-\frac{\alpha}{2}} \frac{\theta}{2\sqrt{n}}\right)\right] \approx 1-\alpha$$

(c) odhad $\hat{\theta}_n$ parametru θ momentovou metódou

$$EX^2 = 2\theta^2$$
$$\theta^2 = \frac{EX^2}{2}$$
$$\theta = \sqrt{\frac{EX^2}{2}} \; ; \quad \widehat{EX^2} = \frac{1}{n}\sum_{i=1}^n X_i^2$$
$$\Rightarrow \tilde{\theta}_n = \sqrt{\frac{\frac{1}{n}\sum_{i=1}^n X_i^2}{2}} = \sqrt{\frac{1}{2n}\sum_{i=1}^n X_i^2}$$

(d) pomocou CLV a transformace $g(\theta) = \sqrt{\frac{\theta}{2}}$ as. rozd. momentového odhadu

CLV: $\sqrt{n}(\overline{X_n^2} - EX^2) \xrightarrow{\mathcal{D}} N\left(0, \frac{\theta^2}{4}\right)$

$$\sqrt{n}\Big(\overline{X}_n^2 - \underbrace{EX_i^2}_{=2\theta^2}\Big) \xrightarrow{\mathcal{D}} N\big(0, \underbrace{\mathrm{Var}X_i^2}\big)$$

$$\mathrm{Var}X_i^2 = \underbrace{EX_i^4} - (EX_i^2)^2 = 8\theta^4 - 4\theta^4 = 4\theta^4$$

$$= \int_0^\infty x^4 \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} dx = \int_0^\infty \frac{x^5}{\theta^2} e^{-\frac{x^2}{2\theta^2}} dx \overset{s}{=} \left[\begin{matrix} y = x^2 \\ dy = 2x\,dx \end{matrix}\right] =$$
$$= \int_0^\infty \frac{y^2}{2\theta^2} e^{-\frac{y}{2\theta^2}} dy = EY^2 \quad \text{kde } Y \sim Exp\left(\frac{1}{2\theta^2}\right)$$
$$\hookrightarrow EY^2 = \mathrm{Var}Y + (EY)^2 = 4\theta^4 + 4\theta^4 = 8\theta^4$$

$$\sqrt{n}(\overline{X_n^2} - EX^2) \xrightarrow{\mathcal{D}} N(0, 4\theta^4)$$

→ Δ · $g(t) = \sqrt{\frac{t}{2}}$

$g'(t) = \frac{1}{\sqrt{2}} \frac{1}{2\sqrt{t}} = \frac{1}{2\sqrt{2t}}$; $(g'(t))^2 = \frac{1}{8t}$

$\sqrt{n}\left(g(\overline{X^2_n}) - g(EX^2)\right) \xrightarrow{D} N(0, g'(EX_i^2)\, Var X_i^2\, g'(EX_i^2))$

$\sqrt{n}\left(\sqrt{\frac{1}{2n}\sum X_i^2} - \theta\right) \xrightarrow{D} N(0, 4\theta^4 \frac{1}{16\theta^2})$

$\sqrt{n}(\hat\theta_n - \theta) \xrightarrow{D} N(0, \frac{\theta^2}{4})$

(e) porovnať as. rozptyl odhadov $\tilde\theta_n$ a $\hat\theta_n$

$\tilde\theta_n : \frac{\theta^2}{4} = \frac{\theta^2}{4} : \hat\theta_n$ => sú rovnaké

(f) použiť transformáciu $Y_i = X_i^2$ a pro $\alpha = 0{,}05$ zostrojiť presný interval spoľahlivosti pro $\theta > 0$

$X_i \xrightarrow{t} X_i^2 = Y_i \xrightarrow{\tau} \sqrt{Y_i} = X_i$

$x^2 = y$; $\tau(y) = \sqrt{y}$

$x = \sqrt{y}$; $|\tau'| = \frac{1}{2\sqrt{y}}$

$f_\theta(x) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}} 1\!\!1\{x>0\}$, $\theta \in (0,\infty)$

$f_Y(y) = \frac{\sqrt{y}}{\theta^2} \exp\{\frac{-y}{2\theta^2}\} \frac{1}{2\sqrt{y}} 1\!\!1\{y\in(0,\infty)\} = \frac{1}{2\theta^2} \exp\{\frac{-y}{2\theta^2}\} 1\!\!1\{y>0\} \sim Exp(\frac{1}{2\theta^2})$

$Y_i \sim Exp(\frac{1}{2\theta^2}) \to Y_1/\theta^2, \dots, Y_n/\theta^2 \sim Exp(\frac{1}{2}) \to \sum_{i=1}^n \frac{Y_i}{\theta^2} \sim \Gamma(\frac{1}{2}, n) \equiv \chi^2_{2n}$

$\frac{r}{2} = n \Rightarrow r = 2n$

$P\left[q_{\frac{\alpha}{2}} < \frac{1}{\theta^2}\sum_{i=1}^n Y_i < q_{1-\frac{\alpha}{2}}\right] = 1-\alpha$; $\alpha = 0{,}05$

.. atď. ekv. úpravy ale už mám dosť (osamostatniť θ)

2. $X_1, \dots, X_{2n} \sim f(x,\theta) = \theta x^{\theta-1} 1\!\!1\{x\in(0,1)\}$, $\theta > 0$;

transformace $W_i = -\log X_i$

(a) určiť rozdelenie $Y = \sum_{i=1}^n W_i$

$X_i \xrightarrow{t} -\log X_i = W_i \xrightarrow{\tau} e^{-W_i} = X_i$

$\log x^{-1} = w$; $\tau(w) = e^{-w}$

$e^w = x^{-1} = \frac{1}{x}$; $|\tau'| = e^{-w}$

$x = e^{-w}$

$f_W(w) = \theta e^{-w(\theta-1)} e^{-w} 1\!\!1\{w>0\} = \theta e^{-w\theta} 1\!\!1\{w>0\} \sim Exp(\theta)$

$W_i \sim Exp(\theta) \to \Gamma(\theta, n) \Rightarrow f_Y(y) = \frac{\theta^n}{\Gamma(n)} y^{n-1} e^{-\theta y} 1\!\!1\{y\in(0,\infty)\}$

(b) určiť rozdelenie $Z = \min_{i=n+1,\dots,2n} W_i$

$$Z = W_{(1)} \sim F_{(1)}(z) = 1 - (1 - \underbrace{F_W(x)}_{})^n = 1 - e^{-\theta x n} \sim Exp(n\theta)$$

$W_i \sim Exp(\theta)$

$F_W = 1 - e^{-\theta x}$

(c) odvodiť rozdelenie $\frac{nZ}{Y+nZ}$

$Z \sim Exp(n\theta) \quad \to \quad nZ \sim Exp(\theta) \sim \Gamma(\theta, 1)$

$Y \sim \Gamma(\theta, n)$

$\Big\} \; Y + nZ \sim \Gamma(\theta, n+1)$

$Y \perp\!\!\!\perp Z \implies Y \perp\!\!\!\perp nZ$

$$(nZ, Y)^T \sim f_{(nZ,Y)}(z,y) = f_{nZ}(z)\, f_{(Y)}(y) = \theta e^{-\theta z} \frac{\theta^n}{\Gamma(n)} y^{n-1} e^{-\theta y} \, \mathbb{1}\{(z,y) \in (0,\infty)^2\} =$$
$$= \frac{\theta^{n+1}}{\Gamma(n)} y^{n-1} e^{-\theta(z+y)} \, \mathbb{1}\{(z,y) \in (0,\infty)^2\}$$

$$\begin{pmatrix} nZ \\ Y \end{pmatrix} \xrightarrow{t} \begin{pmatrix} nZ \\ \frac{nZ}{Y+nZ} \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix} \xrightarrow{\tau} \begin{pmatrix} U \\ \frac{U}{V} - U \end{pmatrix} = \begin{pmatrix} nZ \\ Y \end{pmatrix}$$

$nZ = u$

$\frac{nZ}{y+nZ} = v \to \frac{u}{y+u} = v$

$u = vy + vu$

$u = v(y+u)$

$\frac{u}{v} = y + u$

$\frac{u}{v} - u = y$

$\tau_1(u,v) = u$

$\tau_2(u,v) = \frac{u}{v} - u$

$$|J_\tau| = \begin{vmatrix} 1 & 0 \\ \frac{1}{v} - 1 & -\frac{u}{v^2} \end{vmatrix} = \frac{u}{v^2}$$

$$f_{U,V}(u,v) = f_{nZ,Y}(\tau(u,v))\, |J_{\tau(u,v)}| \, \mathbb{1}\{(u,v) \in (0,1) \times (0,\infty)\}$$

potom $f_V(u) = \int_{\mathbb{R}} f_{U,V}(u,v)\, du$

(d) pomocou CLV spočítať približne pravdepodobnosť $P[Y \leq \sqrt{n}]$

$Y = \sum_{i=1}^{n} W_i \sim \Gamma(\theta, n)$

$W_i \sim Exp(\theta)$

$$P[Y \leq \sqrt{n}] = P[\sum_{i=1}^{n} W_i \leq \sqrt{n}] = P[\tfrac{1}{n} \sum W_i \leq \tfrac{\sqrt{n}}{n}] = P[\overline{W}_n \leq \tfrac{\sqrt{n}}{n}] = P[\overline{W}_n - EW \leq \tfrac{1}{\sqrt{n}} - EW] =$$

$$= P[\sqrt{n}(\overline{W}_n - EW) \leq \sqrt{n}(1 - \sqrt{n}EW)] = P[\sqrt{n}\, \frac{\overline{W}_n - EW}{\sqrt{VarW}} \leq \sqrt{n}(1 - \tfrac{\sqrt{n}}{\theta})] =$$

$$= P[\sqrt{n}\, \frac{\overline{W}_n - EW}{\sqrt{VarW}} \leq$$

?