Универзитет у Београду
Математички факултет

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ, 3Р)

4. септембар 2023. године

---

1. Нека је (X_n) низ независних случајних величина чији општи члан има униформну $\mathcal{U}[2023, 2024]$ расподелу. Ако је
$$Y_n = \ln\left(\prod_{j=1}^{n+1} \frac{1}{\sqrt[a]{X_j - 2023}}\right),$$
испитати да ли за низ случајних величина (aY_n), где је a природан број већи од један, важи закон великих бројева.

2. Из популације чије обележје X има униформну $\mathcal{U}[0, \theta]$, $\theta > 0$, расподелу извучен је узорак обима n. За оцену непознатог параметра θ на основу тог узорка предлажу се оцене U и V, где је U оцена добијена методом максималне веродостојности, а $V = 2\bar{X}_n$. Испитати непристрасност и постојаност предложених оцена. Која оцена је боља у средње квадратном смислу?

3. Удружење банака Србије спроводи истраживање како би се испитала разлика просечне изложености банака кредитном и тржишном ризику. Претпоставља се да изложеност кредитном ризику има нормалну $\mathcal{N}(m_1, \sigma^2)$ расподелу, а изложеност тржишном ризику нормалну $\mathcal{N}(m_2, \sigma^2)$ расподелу, где су m_1, m_2 и σ непознати. Измерене су изложености кредитном и тржишном ризику у неких 7 од свих банака које послују на територији Србије. Одредити ниво поверења симетричног у односу на разлику узорачких средина интервала поверења за разлику просечне изложености кредитном и тржишном ризику, ако је дужина тог интервала мања од 2.023σ са вероватноћом 0.98.

Универзитет у Београду
Математички факултет

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ, 3Р)

4. септембар 2023. године

1. Нека је (X_n) низ независних случајних величина чији општи члан има униформну $\mathcal{U}[2023, 2024]$ расподелу. Ако је
$$Y_n = \ln\left(\prod_{j=1}^{n+1} \frac{1}{\sqrt[a]{X_j - 2023}}\right),$$
испитати да ли за низ случајних величина (aY_n), где је a природан број већи од један, важи закон великих бројева.

2. Из популације чије обележје X има униформну $\mathcal{U}[0, \theta]$, $\theta > 0$, расподелу извучен је узорак обима n. За оцену непознатог параметра θ на основу тог узорка предлажу се оцене U и V, где је U оцена добијена методом максималне веродостојности, а $V = 2\bar{X}_n$. Испитати непристрасност и постојаност предложених оцена. Која оцена је боља у средње квадратном смислу?

3. Удружење банака Србије спроводи истраживање како би се испитала разлика просечне изложености банака кредитном и тржишном ризику. Претпоставља се да изложеност кредитном ризику има нормалну $\mathcal{N}(m_1, \sigma^2)$ расподелу, а изложеност тржишном ризику нормалну $\mathcal{N}(m_2, \sigma^2)$ расподелу, где су m_1, m_2 и σ непознати. Измерене су изложености кредитном и тржишном ризику у неких 7 од свих банака које послују на територији Србије. Одредити ниво поверења симетричног у односу на разлику узорачких средина интервала поверења за разлику просечне изложености кредитном и тржишном ризику, ако је дужина тог интервала мања од 2.023σ са вероватноћом 0.98.

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ, 3Р) - Писмени испит 4. септембар 2023.

Решења задатака

1. Имамо да је

$$
\begin{aligned}
Yn &= \ln\left(\prod_{j=1}^{n+1} \frac{1}{\sqrt[a]{X_j - 2023}}\right) = \sum_{j=1}^{n+1} \ln\left(\frac{1}{\sqrt[a]{X_j - 2023}}\right) \\
&= \sum_{j=1}^{n+1} - \ln\left(X_j - 2023\right)^{\frac{1}{a}} = \sum_{j=1}^{n+1} U_j,
\end{aligned}
$$

где је $U_j = -\frac{1}{a}\ln\left(X_j - 2023\right)$. Приметимо да важи $aU_j \sim \mathcal{E}(1)$. Заиста,

$$
\begin{aligned}
P\{aU_j \leq x\} &= P\left\{-\ln(X_j - 2023) \leq x\right\} \\
&= P\left\{X_j \geq e^{-x} + 2023\right\} \\
&= 1 - (e^{-x} + 2023 - 2023) \\
&= 1 - e^{-x},
\end{aligned}
$$

односно $aU_j \sim \mathcal{E}(1)$. Како су (aU_n) међусобно независне случајне величине, јер су то и (X_n), следи да је $aY_n \sim \gamma(n+1, 1)$. Нека је $S_n = \sum_{k=1}^{n} aY_k$. Желимо да испитамо да ли

$$
\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0,\ n \to \infty.
$$

Како је $E(aY_n) = n + 1$, добијамо да је

$$
ES_n = \sum_{k=1}^{n} EX_k = \sum_{k=1}^{n} (k+1) = \frac{n(n+3)}{2}.
$$

Пошто конвергенција у вероватноћи повлачи конвергенцију у расподели, ако покажемо да $U_n = \frac{S_n}{n} - \frac{n+3}{2} \overset{D}{\not\to} 0$, $n \to \infty$, онда $\frac{S_n}{n} - \frac{n+3}{2} \overset{P}{\not\to} 0$, $n \to \infty$, па не важи закон великих бројева. Показаћемо да $\varphi_{U_n}(t) \not\to \varphi_0(t) = 1$, бар за једно $t \in \mathbb{R}$. Пошто је

$$
\varphi_{aY_n}(t) = E(e^{itaY_n}) = \int_0^{\infty} e^{itx} \frac{x^n e^{-x}}{n!} dx = \frac{1}{n!}\int_0^{\infty} x^n e^{-x(1-it)} dx = \frac{1}{(1-it)^{n+1}},
$$

следи да је

$$
\varphi_{\frac{S_n}{n}}(t) = \prod_{k=1}^{n} \varphi_{aY_n}\left(\frac{t}{n}\right) = \left(\frac{1}{1 - \frac{it}{n}}\right)^{\sum_{k=1}^{n} k + n} = \left(\frac{1}{1 - \frac{it}{n}}\right)^{\frac{n(n+3)}{2}},
$$

односно

$$
\begin{aligned}
\varphi_{U_n}(t) &= e^{-\frac{n+3}{2}it} \varphi_{\frac{S_n}{n}}(t) \\
&= e^{-\frac{n+3}{2}it} \left(\frac{1}{1 - \frac{it}{n}}\right)^{\frac{n(n+3)}{2}} \\
&= e^{-\frac{n+3}{2}it} e^{-\frac{n(n+3)}{2}\ln\left(1 - \frac{it}{n}\right)} \\
&= e^{-\frac{n+3}{2}it} e^{-\frac{n(n+3)}{2}\left(-\frac{it}{n} + \frac{t^2}{2n^2} + o\left(\frac{1}{n^2}\right)\right)} \to e^{-\frac{t^2}{4}},\ n \to \infty.
\end{aligned}
$$

Како је $e^{-\frac{t^2}{4}}$ вредност карактеристичне функције у тачки t случајне величине која има нормалну $\mathcal{N}(0, \frac{1}{2})$ расподелу, закључујемо да $U_n \overset{D}{\not\to} 0$, $n \to \infty$, па ни $U_n \overset{P}{\not\to} 0$, $n \to \infty$, те не важи слаби закон великих бројева.

2. За густину расподеле обележја X важи да је

$$f(x;\theta) = \frac{1}{\theta} I\{0 \leq x \leq \theta\}.$$

Функција веродостојности је тада

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I\{x_i \leq \theta\} = \frac{1}{\theta^n} \prod_{i=1}^{n} I\{x_i \leq \theta\} = \frac{1}{\theta^n} I\{x_{(n)} \leq \theta\}.$$

Приметимо да функција веродостојности није диференцијабилна по θ. Да би функција веродостојности имала максимум, потребно је да индикатор узме вредност 1, у супротном је функција веродостојности једнака 0. Како је функција $\frac{1}{\theta^n}$ опадајућа по θ, треба да изаберемо што мање θ, али такво да је $x_{(n)} \leq \theta$ да би индикатор узео вредност 1. Према томе, оцена добијена методом максималне веродостојности је

$$U = X_{(n)}.$$

Функција расподеле случајне величине U је дата са

$$\begin{aligned} F_U(x) &= P\{U \leq x\} = P\{X_{(n)} \leq x\} \\ &= (F_X(x))^n = \left(\frac{x}{\theta}\right)^n, \quad 0 \leq x \leq \theta, \end{aligned}$$

одакле следи

$$f_U(x) = \frac{n}{\theta^n} x^{n-1}, \quad 0 \leq x \leq \theta.$$

Испитајмо непристрасност предложених оцена:

$$E(U) = E(X_{(n)}) = \int_0^\theta x \frac{n}{\theta^n} x^{n-1} dx = \theta \frac{n}{n+1},$$

$$E(V) = E(2\bar{X}_n) = 2E(X_1) = 2 \cdot \frac{\theta}{2} = \theta.$$

Из претходног закључујемо да је оцена V непристрасна, а оцена U није, али јесте асимптотски.

Остало је још да испитамо постојаност ових оцена. Постојаност подразумева да оцена конвергира у вероватноћи ка параметру θ. Нека је $\varepsilon > 0$. Тада важи

$$\begin{aligned} P\{|U-\theta| > \varepsilon\} &= 1 - P\{U - \theta| \leq \varepsilon\} = 1 - P\{\theta - \varepsilon < U < \theta + \varepsilon\} \\ &= 1 - (F_U(\theta+\varepsilon) - F_U(\theta-\varepsilon)) \\ &= \begin{cases} 1 - \left(1 - \frac{n}{\theta^n}(\theta-\varepsilon)^{n-1}\right), & 0 < \varepsilon < \theta \\ 1 - (1-0), & \varepsilon \geq \theta \end{cases} \\ &= \begin{cases} \frac{n}{\theta}\left(\frac{\theta-\varepsilon}{\theta}\right)^{n-1}, & 0 < \varepsilon < \theta \\ 0, & \varepsilon \geq \theta, \end{cases} \end{aligned}$$

па је

$$\lim_{n\to\infty} P\{|U-\theta| > \varepsilon\} = \begin{cases} \lim_{n\to\infty} \frac{n}{\theta}\left(\frac{\theta-\varepsilon}{\theta}\right)^{n-1}, & 0 < \varepsilon < \theta \\ 0, & \varepsilon \geq \theta \end{cases} = 0,$$

јер је $0 < \frac{\theta-\varepsilon}{\theta} < 1$, а познато је да за $|q| < 1$ важи $\lim_{n\to\infty} nq^n = 0$. Дакле, $U \xrightarrow{P} \theta$, односно U је постојана оцена параметра θ.

Са друге стране за оцену V имамо још да је

$$D(V) = D(2\bar{X}_n) = \frac{4}{n^2} D(X_1) = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n},$$

па за $\varepsilon > 0$ важи

$$P\{|V-\theta| > \varepsilon\} \leq \frac{E(V-\theta)^2}{\varepsilon^2} = \frac{D(V)}{\varepsilon^2} = \frac{\theta^2}{3n\varepsilon^2},$$

при чему је коришћена Чебишовљева неједнакост. Дакле,

$$0 \leq P\{|V-\theta| > \varepsilon\} \leq \frac{\theta^2}{3n\varepsilon^2} \to 0, \quad n \to \infty,$$

па следи $P\{|V-\theta| > \varepsilon\} \to 0$, $n \to \infty$. Закључујемо $V \xrightarrow{P} \theta$, односно и оцена V је постојана оцена параметра θ.

Даље имамо да је

$$
\begin{aligned}
MSE(U) = MSE(X_{(n)}) &= E(X_{(n)} - \theta)^2 = E(X_{(n)}^2) - 2\theta E X_{(n)} + \theta^2 \\
&= \int_0^\theta x^2 n \frac{x^{n-1}}{\theta^n} dx - 2\theta \cdot \theta \frac{n}{n+1} + \theta^2 \\
&= \theta^2 \left(\frac{n}{n+2} - 2\frac{n}{n+1} + 1 \right) \\
&= \frac{2\theta^2}{(n+1)(n+2)}.
\end{aligned}
$$

Са друге стране, како је V непристрасна оцена параметра θ, то је

$$MSE(V) = E(V-\theta)^2 = E(V - E(V))^2 = D(V) = \frac{\theta^2}{3n}.$$

Коначно, како је

$$MSE(U) = \frac{2\theta^2}{(n+1)(n+2)} < \frac{\theta^2}{3n} = MSE(V),$$

то је оцена U боља од V у средње квадратном смислу.

3. Нека је X обележје које представља изложеност кредитном ризику, а Y обележје које представља изложеност тржишном ризику. Претпоставља се да обележја X и Y имају нормалну расподелу са једнаким дисперзијама. Дисперзије нам нису познате, али знамо да

$$T = \frac{(\bar{X}_{n_1} - m_1) - (\bar{Y}_{n_2} - m_2)}{\sqrt{n_1 \bar{S}^2_{n_1}(X) + n_2 \bar{S}^2_{n_2}(Y)}} \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}(n_1 + n_2 - 2)} \in t_{n_1+n_2-2},$$

где је $n_1 = n_2 = 7$, па случајна величина T има Студентову t_{12} расподелу. Треба да одредимо ниво поверења β тако да је

$$P\{U_7 < m_1 - m_2 < V_7\} = \beta, \tag{1}$$

где су U_7 и V_7 статистике дефинисане на основу узорка такве да је $P\{U_7 \le V_7\} = 1$, односно на основу тога добија се

$$P\left\{\bar{X}_7 - \bar{Y}_7 - V_7 < \bar{X}_7 - \bar{Y}_7 - (m_1 - m_2) < \bar{X}_7 - \bar{Y}_7 - U_7\right\} = \beta.$$

$$P\left\{ \frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} < \frac{\bar{X}_7 - \bar{Y}_7 - (m_1 - m_2)}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} < \frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} \right\} = \beta.$$

Знамо да је вероватноћа између $\frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6}$ и $\frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6}$ једнака β, међутим њихов положај може бити произвољан. Узимамо вјероватносно симетричан интервал, тј. такав да

$$P\left\{ T < \frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} \right\} = P\left\{ T > \frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} \right\},$$

а пошто је Студентова расподела симетрична око нуле, то важи да су ове вредности једнаке по апсолутној вриједности. Користећи да $T \in t_{12}$ добијамо да је

$$\frac{\bar{X}_7 - \bar{Y}_7 - V_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} = -t_{12;1-\beta},$$

$$\frac{\bar{X}_7 - \bar{Y}_7 - U_7}{\sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}} \cdot \sqrt{6} = t_{12;1-\beta},$$

Према томе,

$$
\begin{aligned}
U_7 &= \bar{X}_7 - \bar{Y}_7 - \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)}, \\
V_7 &= \bar{X}_7 - \bar{Y}_7 + \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)},
\end{aligned}
$$

односно имамо да је $\beta\%$ двострани интервал поверења за разлику просечне изложености кредитном и тржишном ризику:

$$I = \left(\bar{X}_7 - \bar{Y}_7 - \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)},\, \bar{X}_7 - \bar{Y}_7 + \frac{t_{12;1-\beta}}{\sqrt{6}} \sqrt{\bar{S}^2_7(X) + \bar{S}^2_7(Y)} \right),$$

па је дужина уоченог интервала поверења једнака

$$d(I) = t_{12;1-\beta}\sqrt{\frac{2}{3}}\sqrt{\bar{S}_7^2(X) + \bar{S}_7^2(Y)}.$$

Из услова задатка важи да је

$$P\{d(I) < 2,023\sigma\} = 0.98,$$

$$P\left\{t_{12;1-\beta}\sqrt{\frac{2}{3}}\sqrt{\bar{S}_7^2(X) + \bar{S}_7^2(Y)} < 2.023\sigma\right\} = 0.98,$$

односно

$$P\left\{\frac{\bar{S}_7^2(X) + \bar{S}_7^2(Y)}{\sigma^2} < \frac{2.023^2}{(t_{12;1-\beta})^2}\frac{3}{2}\right\} = 0.98. \tag{2}$$

Како је $\frac{n_1\bar{S}_{n_1}^2(X)}{\sigma} \in \chi^2_{n_1-1}$ и $\frac{n_2\bar{S}_{n_2}^2(Y)}{\sigma} \in \chi^2_{n_2-1}$, то из адитивности χ^2 расподеле добијамо да је

$$\frac{n_1\bar{S}_{n_1}^2(X) + n_2\bar{S}_{n_2}^2(Y)}{\sigma} \in \chi^2_{n_1+n_2-2},$$

па из (2) следи да је

$$7 \cdot \frac{2.023^2}{(t_{12;1-\beta})^2} \cdot \frac{3}{2} = \chi^2_{12;0.02} = 24.054 \quad \Rightarrow \quad t_{12;1-\beta} = 1.3365.$$

Коначно, из (1) добијамо да је тражени ниво поверења једнак

$$\beta = F_{t_{12}}(1.3365) = 0.8.$$