Универзитет у Београду
Математички факултет

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ, 3Р)

26. јун 2023. године

---

1. Општи чланови X_n и Y_n независних низова независних случајних величина имају униформну $\mathcal{U}[0, 2023]$ расподелу. Ако је
$$Z_n = n \cdot \ln \left(\frac{\max \left(X_{(n)}, Y_{(n)} \right)}{\min \left(X_{(n)}, Y_{(n)} \right)} \right),$$
у зависности од реалног параметра α, где је $\alpha > 0$, испитати све четири врсте конвергенције низа случајних величина $\left(\frac{1}{n^\alpha Z_n} \right)$.

2. За густину расподеле обележја X важи да је $f(x) = e^{-(x-\theta)}$, $x \geq \theta$, $\theta < 1$. Извучен је узорак обима 300 и констатовано је да је 100 елемената узорка већих од 2. На основу тих података, формирати 95%-ни интервал поверења за параметар θ.

3. На тржишту се појавила нова LPDDR5 генерација RAM меморије за коју се тврди да уписује податке брже од старе DDR4 генерације. Како би се то проверило, вршен је упис 10 скупова података на обе меморије и сваки пут је мерено време које је протекло при упису. За стару меморију добијено је да је просек тих времена једнак 43.23 секунде, а да је узорачка дисперзија једнака 0.75. За нову меморију добијен је просек 42.14 и узорачка дисперзија 0.683. Искуство показује да се може претпоставити да су времена уписа нормално расподељена и да имају једнаке дисперзије. На нивоу значајности 0.05 испитати да ли је просечна брзина уписа података на нову меморију већа од просечне брзине уписа података на стару меморију.

Универзитет у Београду
Математички факултет

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ, 3Р)

26. јун 2023. године

1. Општи чланови X_n и Y_n независних низова независних случајних величина имају униформну $\mathcal{U}[0, 2023]$ расподелу. Ако је
$$Z_n = n \cdot \ln \left(\frac{\max \left(X_{(n)}, Y_{(n)} \right)}{\min \left(X_{(n)}, Y_{(n)} \right)} \right),$$
у зависности од реалног параметра α, где је $\alpha > 0$, испитати све четири врсте конвергенције низа случајних величина $\left(\frac{1}{n^\alpha Z_n} \right)$.

2. За густину расподеле обележја X важи да је $f(x) = e^{-(x-\theta)}$, $x \geq \theta$, $\theta < 1$. Извучен је узорак обима 300 и констатовано је да је 100 елемената узорка већих од 2. На основу тих података, формирати 95%-ни интервал поверења за параметар θ.

3. На тржишту се појавила нова LPDDR5 генерација RAM меморије за коју се тврди да уписује податке брже од старе DDR4 генерације. Како би се то проверило, вршен је упис 10 скупова података на обе меморије и сваки пут је мерено време које је протекло при упису. За стару меморију добијено је да је просек тих времена једнак 43.23 секунде, а да је узорачка дисперзија једнака 0.75. За нову меморију добијен је просек 42.14 и узорачка дисперзија 0.683. Искуство показује да се може претпоставити да су времена уписа нормално расподељена и да имају једнаке дисперзије. На нивоу значајности 0.05 испитати да ли је просечна брзина уписа података на нову меморију већа од просечне брзине уписа података на стару меморију.

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ, 3Р) - Писмени испит 26. јун 2023.

Решења задатака

1. Како су случајне величине $X_{(n)}$ и $Y_{(n)}$ међусобно независне, то је густина расподеле случајног вектора $(X_{(n)}, Y_{(n)})$ дата са

$$f(x,y) = f_{X_{(n)}}(x) f_{Y_{(n)}}(y) = n^2 2023^{-2n} x^{n-1} y^{n-1} \quad (x,y) \in (0, 2023)^2$$

Даље је за свако $z \geq 0$

$$F_{Z_n}(z) = P\{Z_n \leq z\} = P\Big\{Z_n \leq z, X_{(n)} < Y_{(n)}\Big\} + P\Big\{Z_n \leq z, X_{(n)} \geq Y_{(n)}\Big\}$$

односно

$$F_{Z_n}(z) = P\Big\{X_{(n)} < Y_{(n)} \leq e^{\frac{z}{n}} X_{(n)}\Big\} + P\Big\{Y_{(n)} \leq X_{(n)} \leq e^{\frac{z}{n}} Y_{(n)}\Big\},$$

па је функција расподеле случајне величине Z_n дата је са

$$F_{Z_n}(z) = 2 \int_0^{2023} \int_{\frac{x}{e^{z/n}}}^{x} f(x,y)\,dy\,dx = 1 - e^{-z}.$$

Закључујемо да $Z_n \in \mathcal{E}(1)$.

Напомена: Можемо приметити и да важи следеће:

$$Z_n = n \ln \left(\frac{\max(X_{(n)}, Y_{(n)})}{\min(X_{(n)}, Y_{(n)})} \right) = n \ln \max \left\{ \frac{X_{(n)}}{Y_{(n)}}, \frac{Y_{(n)}}{X_{(n)}} \right\} = n \max \left\{ \ln \frac{X_{(n)}}{Y_{(n)}}, \ln \frac{Y_{(n)}}{X_{(n)}} \right\} = n \left| \ln \frac{X_{(n)}}{Y_{(n)}} \right|,$$

или директно

$$\begin{aligned} P\{Z_n \geq t\} &= P\left\{ \frac{Z_n}{n} \geq \frac{t}{n} \right\} = P\left\{ \exp\left(\frac{Z_n}{n}\right) \geq e^{\frac{t}{n}} \right\} = P\left\{ \frac{\max(X_{(n)}, Y_{(n)})}{\min(X_{(n)}, Y_{(n)})} \geq e^{\frac{t}{n}} \right\} \\ &= P\left\{ e^{-\frac{t}{n}} \max(X_{(n)}, Y_{(n)}) \geq \min(X_{(n)}, Y_{(n)}) \right\} = 2P\left\{ e^{-\frac{t}{n}} X_{(n)} \geq Y_{(n)} \right\}. \end{aligned}$$

Ова два приступа лако доводе до истог закључка, те их остављамо читаоцу за вежбу. $\triangle$

Вратимо се сад на задатак. Уведимо ознаку $T_n = \frac{1}{n^\alpha Z_n}$. Имамо да је

$$\begin{aligned} P\{T_n > \varepsilon\} &= P\left\{ \frac{1}{n^\alpha Z_n} > \varepsilon \right\} = P\left\{ Z_n < \frac{1}{\varepsilon n^\alpha} \right\} \\ &= 1 - e^{-1/(\varepsilon n^\alpha)} \to 0, \quad n \to \infty. \end{aligned}$$

Према томе, $T_n \xrightarrow{P} 0$, па следствено томе важи $T_n \xrightarrow{D} 0$. У циљу испитивања скоро сигурне конвергенције рачунамо:

$$\begin{aligned} \sum_{n=1}^{\infty} P\{T_n > \varepsilon\} &= \sum_{n=1}^{\infty} (1 - e^{-1/(\varepsilon n^\alpha)}) \\ &= \sum_{n=1}^{\infty} \left\{ 1 - \left[1 - \frac{1}{\varepsilon n^\alpha} + o\left(\frac{1}{n^\alpha}\right) \right] \right\} \\ &= \sum_{n=1}^{\infty} \left[\frac{1}{\varepsilon n^\alpha} + o\left(\frac{1}{n^\alpha}\right) \right] \quad \begin{cases} < +\infty, & \text{ако је } \alpha > 1, \\ = +\infty, & \text{ако је } 0 < \alpha \leq 1. \end{cases} \end{aligned}$$

Према томе, за $\alpha > 1$ низ (T_n) конвергира скоро сигурно ка нули, а ако је $0 < \alpha \leq 1$, онда низ (T_n) не конвергира скоро сигурно. Даље, испитајмо конвергенцију у средње квадратном. Важи:

$$E|T_n|^2 = E\left(\frac{1}{n^{2\alpha}Z_n^2}\right) = \frac{1}{n^{2\alpha}}E\left(\frac{1}{Z_n^2}\right) = \frac{1}{n^{2\alpha}}\int_0^{+\infty}\frac{e^{-x}}{x^2}dx.$$

Приметимо да је

$$\int_0^{+\infty}\frac{e^{-x}}{x^2}dx \geq \int_0^1 \frac{e^{-x}}{x^2}dx \geq \frac{1}{e}\int_0^1 \frac{1}{x^2}dx = -\frac{1}{x}\bigg|_{x=0}^{x=1} = +\infty.$$

Према томе, низ (T_n) не конвергира у средње квадратном.

2. Уочимо да важи да је, за $x \geq \theta$

$$P\{X \leq x\} = \int_\theta^x e^{-(t-\theta)}\,dt = e^\theta \int_\theta^x e^{-t}\,dt = e^\theta\left(1 - e^{-x} - (1 - e^{-\theta})\right) = 1 - e^{-(x-\theta)}.$$

Следи да је $P\{X > x\} = e^{-(x-\theta)}$, па је, специјално, $P\{X > 2\} = e^{-(2-\theta)} = e^{\theta-2}$.

Означимо $I_j = I\{X_j > 2\}$. Број елемената узорка већих од 2 дат је као $\sum_{j=1}^{300} I_j$. Обим узорка је довољно велики, а индикатори I_j су независни и једнако расподељени, те се на њих може применити централна гранична теорема, те ћемо интервал поверења тражити из једнакости:

$$P\left\{-c < \frac{\sum_{j=1}^{300} I_j - E\left(\sum_{j=1}^{300} I_j\right)}{\sqrt{D\left(\sum_{j=1}^{300} I_j\right)}} < c\right\} = 0.95,$$

односно

$$P\left\{-c < \frac{\sum_{j=1}^{300} I_j - 300e^{\theta-2}}{\sqrt{300e^{\theta-2}(1-e^{\theta-2})}} < c\right\} = 0.95.$$

Вредност c налазимо из претпоставке задовољене нормалности, то јест

$$\Phi(c) - \Phi(-c) = 0.95,$$

односно

$$2\Phi(c) - 1 = 0.95,$$

па је $c = \Phi^{-1}(0.975) = 1.96$. Сада расписујемо:

$$\begin{aligned}
-1.96 < \frac{100 - 300e^{\theta-2}}{\sqrt{300e^{\theta-2}(1-e^{\theta-2})}} &< 1.96 \\
\frac{(100 - 300e^{\theta-2})^2}{300e^{\theta-2}(1-e^{\theta-2})} &< 1.96^2 = 3.84 \\
\frac{10^2(10 - 30e^{\theta-2})^2}{10^2 \cdot 3e^{\theta-2}(1-e^{\theta-2})} &< 3.84 \\
(10 - 30e^{\theta-2})^2 &< 11.52e^{\theta-2}(1-e^{\theta-2}) \\
(1 - 3e^{\theta-2})^2 &< 0.1152e^{\theta-2}(1-e^{\theta-2}) \\
&\cdots \\
e^{2(\theta-2)} - 0.67e^{\theta-2} + 0.11 &< 0.
\end{aligned}$$

Решимо ли квадратну неједначину по $e^{\theta-2}$, добијамо решење

$$0.282 < e^{\theta-2} < 0.388,$$

што представља интервал поверења за $e^{\theta-2}$. Логаритмујемо:

$$-1.266 < \theta - 2 < -0.947,$$

односно

$$0.734 < \theta < 1.053.$$

Како је дат услов да је $\theta < 1$, тражени интервал поверења једнак је

$$I_\theta = (0.734, 1).$$

3. Ако означимо са μ_1 средњу вредност старе меморије, а са μ_2 нове, ми тестирамо нулту хипотезу $H_0(\mu_1 = \mu_2)$ против алтернативе $H_1(\mu_1 > \mu_2)$. Уколико нам x_i означава мерење старе меморије, а y_i нове, имамо да је $\bar{x}_{10} = 43.22$, $\bar{y}_{10} = 42.14$, $\bar{s}_{10}^2(X) = 0.75$, $\bar{s}_{10}^2(Y) = 0.683$. Узорци су независни и имају исти обим: $n_1 = n_2 = 10$. При оваквим условима, користимо тест статистику

$$T = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{\sqrt{n_1 \bar{S}_{n_1}(X) + n_2 \bar{S}_{n_2}(Y)}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}(n_1 + n_2 - 2)}$$

која под нултом хипотезом има Студентову $t_{n_1+n_2-2}$ расподелу. У нашем случају, користимо статистику

$$T = \frac{\bar{X}_{10} - \bar{Y}_{10}}{\sqrt{10\bar{S}_{10}(X) + 10\bar{S}_{10}(Y)}} \sqrt{90} \in t_{18}.$$

Нулта хипотеза се одбацује уколико је

$$T \geq c,$$

где критичну вредност c налазимо из таблице за Студентову расподелу:

$$0.05 = P_{H_0}\{T \geq c\} = \frac{1}{2} P_{H_0}\{|T| \geq c\},$$

односно

$$0.1 = P_{H_0}\{|T| \geq c\},$$

па из таблице читамо да је $c = 1.734$. Реализована вредност тест статистике једнака је

$$\frac{43.22 - 42.14}{\sqrt{10 \cdot 0.75 + 10 \cdot 0.683}} \sqrt{90} = 2.707.$$

Како је овај број већи од критичне вредности 1.734, одбацујемо H_0 и прихватамо алтернативу да нова меморија ради брже.