



Универзитет у Београду
Математички факултет

БЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ)

СЕПТЕМБАР 1 - 15. октобар 2025. године

1. Општи члан X_n низа независних случајних величина има густину расподеле

$$f_n(x) = \frac{2}{2025^2} x e^{-\left(\frac{x}{2025}\right)^2}, \quad x \geq 0.$$

Ако је

$$Y_n = \left(\frac{X_n}{2025} \right)^2$$

и $Z_n = \frac{1}{n} \min\{Y_1, Y_2, \dots, Y_n\}$, испитати све четири врсте конвергенције низа случајних величина (Z_n) .

2. Време између два узастопна доласка клијената у банку (у минутима) има експоненцијалну $\mathcal{E}(\lambda)$ расподелу. Током више радних дана регистрован је узорак од 900 таквих временских размака и добијени су следећи подаци:

време (мин)	[0, 1)	[1, 2)	[2, 3)	[3, 5)	[5, 8)	[8, 12)	[12, ∞)
број размака	250	210	170	130	90	40	10

Одредити 90% интервал поверења за просечно време између два узастопна доласка клијената у банку.

3. Наставничко веће школе тврди да енглески језик учи $p\%$ ученика, руски језик учи $4p\%$ ученика, а француски језик уче преостали ученици. Да би то проверио, директор школе је случајно одабрао 100 ученика и констатовао да енглески језик учи 24 ученика, руски језик учи 56 ученика, а француски језик учи 20 ученика. На основу тог узорка, са прагом значајности 0.01, испитати да ли је наставничко веће у праву.

Решења задатака

1. Функција расподеле случајне величине X_n дата је са

$$\begin{aligned} F_{X_n}(x) &= P\{X_n \leq x\} = \int_0^x \frac{2}{2025} t e^{-\left(\frac{t}{2025}\right)^2} dt = \left[\frac{\left(\frac{t}{2025}\right)^2}{\frac{2}{2025^2}} = u \right] \\ &= \int_0^{\left(\frac{x}{2025}\right)^2} e^{-u} du = 1 - e^{-\left(\frac{x}{2025}\right)^2}, \quad x \geq 0. \end{aligned}$$

Одавде имамо да је функција расподеле случајне величине Y_n дата са

$$\begin{aligned} F_{Y_n}(y) &= P\{Y_n \leq y\} = P\left\{\left(\frac{X_n}{2025}\right)^2 \leq y\right\} = P\{X_n \leq 2025\sqrt{y}\} \\ &= 1 - e^{-\left(\frac{2025\sqrt{y}}{2025}\right)^2} = 1 - e^{-y}, \quad y \geq 0, \end{aligned}$$

па добијамо да случајна величина Y_n има експоненцијалну $\mathcal{E}(1)$ расподелу. За низ случајних величина Z_n важи

$$\begin{aligned} F_{Z_n}(z) &= P\{Z_n \leq z\} = P\{\min\{Y_1, Y_2, \dots, Y_{n^3}\} \leq zn\} \\ &= 1 - P\{\min\{Y_1, Y_2, \dots, Y_{n^3}\} > zn\} = 1 - (P\{Y_1 > zn\})^{n^3} \\ &= 1 - (e^{-zn})^{n^3} = 1 - e^{-zn^4}, \quad z \geq 0. \end{aligned}$$

Следи да је

$$F_{Z_n}(z) = \begin{cases} 0, & z < 0, \\ 1 - e^{-zn^4}, & z \geq 0 \end{cases} \xrightarrow{n \rightarrow \infty} \begin{cases} 0, & z < 0, \\ 1, & z \geq 0 \end{cases} = F(z),$$

где је $F(z)$ функција расподеле (дегенерисане) случајне величине 0. Из претходног закључујемо $Z_n \xrightarrow{D} 0$, $n \rightarrow \infty$, као и $Z_n \xrightarrow{P} 0$, $n \rightarrow \infty$. Испитујемо сада скоро сигурну конвергенцију. За свако $\varepsilon > 0$ је

$$P(|Z_n| \geq \varepsilon) = e^{-\varepsilon n^4} = \left(\frac{1}{e^\varepsilon}\right)^{n^4},$$

па следи да је

$$\sum_{n=1}^{\infty} P(|Z_n| \geq \varepsilon) = \sum_{n=1}^{\infty} e^{-\varepsilon n^4} \leq \sum_{n=1}^{\infty} e^{-\varepsilon n} < \infty.$$

Одавде закључујемо $Z_n \xrightarrow{c.c.} 0$, $n \rightarrow \infty$. У циљу испитивања средње квадратне конвергенције рачунамо $E(Z_n^2)$. Имамо да је редом

$$E(Z_n)^2 = \int_0^{\infty} z^2 f_{Z_n}(z) dz = \int_0^{\infty} z^2 \cdot n^4 e^{-zn^4} dz.$$

Нека је $zn^4 = u \Rightarrow n^4 dz = du$, добијамо

$$E(Z_n^2) = \frac{1}{n^8} \int_0^{\infty} u^2 e^{-u} du = \frac{1}{n^8} \Gamma(3) = \frac{2}{n^8}.$$

Како $E(Z_n)^2 = \frac{2}{n^8} \rightarrow 0$, $n \rightarrow \infty$, следи да низ (Z_n) конвергира средње квадратно.

2. Узет је узорак обима $n = \sum n_k = 250 + 210 + 170 + 130 + 90 + 40 + 10 = 900 (> 30)$, па можемо да искористимо централну граничну теорему:

$$\frac{\sum_{k=1}^n X_k - E\left(\sum_{k=1}^n X_k\right)}{\sqrt{D\left(\sum_{k=1}^n X_k\right)}} \xrightarrow{D} X^*, \quad X^* \in \mathcal{N}(0, 1).$$

Користећи податке из задатка можемо одредити збир вредности из узорка

$$\sum_{k=1}^n X_k = 250 \cdot 0.5 + 210 \cdot 1.5 + 170 \cdot 2.5 + 130 \cdot 4 + 90 \cdot 6.5 + 40 \cdot 10 + 10 \cdot 12.5 = 2495.$$

С обзиром да је $EX_1 = \frac{1}{\lambda}$ и $DX_1 = \frac{1}{\lambda^2}$, интервал поверења за параметар λ добијамо из једнакости

$$P \left\{ \left| \frac{\sum_{k=1}^{900} X_i - \frac{900}{\lambda}}{\sqrt{\frac{900}{\lambda^2}}} \right| < z \right\} = 0.9.$$

Из таблица за нормалну расподелу имамо да је $z = \Phi^{-1}(0.95) = 1.645$. Решавањем неједначине

$$-1.645 < \frac{2495 - \frac{900}{\lambda}}{\frac{30}{\lambda}} < 1.645,$$

добијамо да је $I_\lambda = (0.3409, 0.3805)$. Тражени интервал поверења за просечно време између два уястопа доласка је

$$I_{\frac{1}{\lambda}} = (2.6281, 2.9334).$$

3. Нека је обележје X једнако 1 ако ученик учи енглески језик, 2 ако учи руски језик и 3 ако учи француски језик. Тестира се хипотеза да је наставничко веће у праву, тј. да обележје X има закон расподеле

$$X : \begin{pmatrix} 1 & 2 & 3 \\ \frac{p}{100} & \frac{4p}{100} & 1 - \frac{5p}{100} \end{pmatrix}.$$

Применићемо Пирсонов χ^2 тест да тестирамо да ли је закон расподеле обележја X једнак конкретном закону из услова задатка. Параметар p је непознат, те одређујемо његову оцену методом максималне веродостојности. Важи да је

$$p(x; p) = \left(\frac{p}{100}\right)^{I\{x=1\}} \left(\frac{4p}{100}\right)^{I\{x=2\}} \left(1 - \frac{5p}{100}\right)^{I\{x=3\}}, \quad x \in \{1, 2, 3\},$$

одакле је функција веродостојности

$$L(p) = \prod_{k=1}^n p(x_k; p) = \left(\frac{p}{100}\right)^{\sum_{k=1}^n I\{x_k=1\}} \left(\frac{4p}{100}\right)^{\sum_{k=1}^n I\{x_k=2\}} \left(1 - \frac{5p}{100}\right)^{\sum_{k=1}^n I\{x_k=3\}},$$

па је

$$\ln L(p) = \sum_{k=1}^n I\{x_k = 1\} \ln \left(\frac{p}{100}\right) + \sum_{k=1}^n I\{x_k = 2\} \ln \left(\frac{4p}{100}\right) + \sum_{k=1}^n I\{x_k = 3\} \ln \left(1 - \frac{5p}{100}\right).$$

Желимо да пронађемо $p > 0$ које максимизује $\ln L(p)$, те стога одређујемо стационарну тачку

$$\frac{\partial}{\partial p} \ln L(p) = \frac{\sum_{k=1}^n I\{x_k = 1\}}{p} + \frac{\sum_{k=1}^n I\{x_k = 2\}}{p} - \frac{5 \sum_{k=1}^n I\{x_k = 3\}}{100 - 5p} = 0,$$

одакле је

$$p = \frac{20}{n} \left(\sum_{k=1}^n I\{x_k = 1\} + \sum_{k=1}^n I\{x_k = 2\} \right).$$

Још треба проверити да ли се у добијеној вредности постиже максимум. Како је

$$\frac{\partial^2}{\partial p^2} \ln L(p) = -\frac{\sum_{k=1}^n I\{x_k = 1\} + \sum_{k=1}^n I\{x_k = 2\}}{p^2} - \frac{25 \sum_{k=1}^n I\{x_k = 3\}}{(100 - 5p)^2} < 0,$$

ово заиста јесте тачка у којој функција $\ln L(p)$ достиже максимум. Према томе, на основу овог узорка, оцена методом максималне веродостојности за p је

$$\hat{p} = \frac{20}{n} \left(\sum_{k=1}^n I\{X_k = 1\} + \sum_{k=1}^n I\{X_k = 2\} \right).$$

На реализованом узорку, имамо да је $M_1 = \sum_{k=1}^n I\{x_k = 1\} = 9$ и $M_2 = \sum_{k=1}^n I\{x_k = 2\} = 36$, па је

$$\hat{p} = \frac{20}{100}(24 + 56) = 16.$$

Сада можемо да израчунамо вероватноће p_k :

$$p_1 = P\{X = 1\} = 0.16, \quad p_2 = P\{X = 2\} = 0.64, \quad p_3 = P\{X = 3\} = 0.20.$$

Дакле, имамо табелу

страни језик	енглески (1)	руски (2)	француски (3)
учесталост (M_k)	24	56	20
p_k	0.16	0.64	0.20
np_k	16	64	20

Видимо да за све категорије важи $np_k > 5$, те нећемо имати спајање. Критична област је облика

$$W = \{\chi_0^2 \geq c\},$$

где c рачунамо из једнакости

$$P_{H_0}\{\chi_0^2 \geq c\} = 0.01.$$

При H_0 , пошто имамо 3 класе и оценили смо један параметар методом максималне веродостојности, $\chi_0^2 \in \chi_{3-1-1}^2$. Према томе,

$$c = \chi_{1;0.01}^2 = 6.635,$$

па је критична област

$$W = \{\chi_0^2 \geq 6.635\}.$$

Рачунамо вредност тест статистике на добијеном узорку:

$$\begin{aligned} \chi_0^2 &= \sum_{i=1}^3 \frac{(M_k - np_k)^2}{np_k} \\ &= \frac{(24 - 16)^2}{16} + \frac{(56 - 64)^2}{64} + \frac{(20 - 20)^2}{20} \\ &= 5. \end{aligned}$$

Пошто је $5 < 6.635$, узорак не упада у критичну област, па не одбацујемо нулту хипотезу, тј. можемо сматрати да је наставничко веће у праву.