



Универзитет у Београду  
Математички факултет

## БЕРОВАТНОЋА И СТАТИСТИКА Б (ЗР)

ЈУН 2 - 2. октобар 2025. године

1. За густину расподеле општег члана  $X_n$  низа независних случајних величина важи да је

$$f_{X_n}(x) = a(1-x)^{a-1}, x \in (0, 1), a > 0.$$

Ако је

$$Y_n = \ln \prod_{k=1}^n (1 - X_k)^{-\frac{k}{n}},$$

испитати да ли за низ случајних величина  $(Y_n)$  важи закон великих бројева.

2. Обележје  $X$  има негативну биномну  $\mathcal{NB}(r, p)$  расподелу са законом расподеле

$$P\{X = x\} = \binom{x-1}{r-1} p^r (1-p)^{x-r}, x \in \{r, r+1, \dots\}.$$

На основу узорка обима  $n$ , методом максималне веродостојности одредити оцену непознатог параметра  $\frac{1}{p}$ , а затим испитати непристрасност и постојаност добијене оцене.

3. У платном систему тестира се нови софтвер за обраду трансакција. На случајан начин изабрано је 12 трансакција и мерено је време њихове обраде у секундама. Добијено је да је просечно време обраде трансакције једнако 3.49 секунди, а да је узорачка дисперзија 0.16. Познато је да је на ранијем тестирању старог софтвера за случајно изабраних 13 трансакција добијено да је просечно време обраде трансакције 3.82 секунде, а да је узорачка дисперзија 0.19. Претпостављајући да времена обраде трансакција и старим и новим софтвером имају нормалне расподеле и да су им дисперзије једнаке, са прагом значајности 0.05 испитати да ли је просечно време обраде трансакције на старом софтверу веће од просечног времена обраде трансакције на новом софтверу.

$$X \in t_{12} : F^{-1}(0.80) = 0.873, F^{-1}(0.90) = 1.356, F^{-1}(0.95) = 1.782, F^{-1}(0.99) = 2.681$$

$$X \in t_{13} : F^{-1}(0.80) = 0.870, F^{-1}(0.90) = 1.350, F^{-1}(0.95) = 1.771, F^{-1}(0.99) = 2.650$$

$$X \in t_{23} : F^{-1}(0.80) = 0.858, F^{-1}(0.90) = 1.319, F^{-1}(0.95) = 1.714, F^{-1}(0.99) = 2.500$$

$$X \in t_{25} : F^{-1}(0.80) = 0.856, F^{-1}(0.90) = 1.316, F^{-1}(0.95) = 1.708, F^{-1}(0.99) = 2.485$$

$$X \in \chi_{12}^2 : F^{-1}(0.80) = 15.812, F^{-1}(0.90) = 18.549, F^{-1}(0.95) = 21.026, F^{-1}(0.99) = 26.217$$

$$X \in \chi_{13}^2 : F^{-1}(0.80) = 16.985, F^{-1}(0.90) = 19.812, F^{-1}(0.95) = 22.362, F^{-1}(0.99) = 27.688$$

$$X \in \chi_{23}^2 : F^{-1}(0.80) = 28.429, F^{-1}(0.90) = 32.007, F^{-1}(0.95) = 35.172, F^{-1}(0.99) = 41.638$$

$$X \in \chi_{25}^2 : F^{-1}(0.80) = 30.675, F^{-1}(0.90) = 34.382, F^{-1}(0.95) = 37.652, F^{-1}(0.99) = 44.314$$

Решења задатака

1. Имамо да је

$$Y_n = \ln \prod_{k=1}^n (1 - X_k)^{-\frac{k}{n}} = \frac{1}{n} \sum_{k=1}^n k (-\ln(1 - X_k)) = \frac{1}{n} \sum_{k=1}^n k U_k,$$

где је  $U_k = -\ln(1 - X_k)$ . Приметимо да важи  $U_k \sim \mathcal{E}(a)$ . Заиста,

$$F_{X_n}(x) = \int_0^x a(1-t)^{a-1} dt = 1 - (1-x)^a, \quad x \in (0, 1),$$

па је

$$\begin{aligned} P\{U_k \leq u\} &= P\{-\ln(1 - X_k) \leq u\} = P\{X_k \leq 1 - e^{-u}\} \\ &= 1 - e^{-au}, \quad u \geq 0. \end{aligned}$$

Одавде имамо да је

$$EY_n = \frac{1}{n} \sum_{k=1}^n k E U_k = \frac{1}{n} \frac{1}{a} \frac{n(n+1)}{2} = \frac{n+1}{2a}.$$

Нека је  $S_n = \sum_{m=1}^n Y_m$ . Желимо да испитамо да ли

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

Имамо да је

$$ES_n = \sum_{m=1}^n EY_m = \sum_{m=1}^n \frac{m+1}{2a} = \frac{1}{2a} \left( \frac{n(n+1)}{2} + n \right) = \frac{n(n+3)}{4a}.$$

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

$$\varphi_{U_k}(t) = E(e^{itU_k}) = \int_0^\infty e^{itx} a e^{-ax} dx = \frac{a}{a - it},$$

следи да је

$$\varphi_{\frac{S_n}{n}}(t) = \prod_{m=1}^n \varphi_{Y_m}\left(\frac{t}{n}\right) = \prod_{m=1}^n \prod_{k=1}^m \frac{a}{a - i \frac{k}{m} \frac{t}{n}},$$

односно

$$\varphi_{W_n}(t) = \exp\left\{-it \frac{ES_n}{n}\right\} \prod_{m=1}^n \prod_{k=1}^m \frac{a}{a - i \frac{k}{m} \frac{t}{n}} = \exp\left\{-it \frac{ES_n}{n}\right\} \prod_{m=1}^n \prod_{k=1}^m \left(1 - i \frac{k}{m} \frac{t}{an}\right)^{-1}.$$

Како је  $\ln(1 - z) = -z - \frac{z^2}{2} + o(z^2)$ ,  $z \rightarrow 0$ , то имамо да је

$$\begin{aligned} \ln \varphi_{W_n}(t) &= -it \frac{ES_n}{n} + \frac{it}{an} \sum_{m=1}^n \sum_{k=1}^m \frac{k}{m} - \frac{t^2}{2a^2n^2} \sum_{m=1}^n \sum_{k=1}^m \frac{k^2}{m^2} + o(1) \\ &= -it \frac{1}{n} \cdot \frac{n^2 + 3n}{4a} + \frac{it}{an} \sum_{m=1}^n \frac{m+1}{2} - \frac{t^2}{2a^2n^2} \sum_{m=1}^n \frac{(m+1)(2m+1)}{6m} + o(1) \\ &= -it \frac{1}{n} \cdot \frac{n^2 + 3n}{4a} + \frac{it}{an} \cdot \frac{n^2 + 3n}{4} - \frac{t^2}{2a^2n^2} \left( \frac{n^2}{6} + \frac{4n}{6} + \frac{H_n}{6} \right) + o(1) \\ &= 0 - \frac{t^2}{12a^2} + o(1) \longrightarrow -\frac{t^2}{12a^2}. \end{aligned}$$

одакле следи

$$\varphi_{W_n}(t) \longrightarrow e^{-\frac{t^2}{12a^2}} \Rightarrow \frac{S_n - ES_n}{n} \xrightarrow{D} \mathcal{N}\left(0, \frac{1}{6a^2}\right).$$

Закључујемо да  $W_n \xrightarrow{D} 0, n \rightarrow \infty$ , па ни  $W_n \xrightarrow{P} 0, n \rightarrow \infty$ , те не важи слаби закон великих бројева.

2. Претпоставимо да имамо реализован узорак  $(x_1, x_2, \dots, x_n)$ . Функција веродостојности је

$$L(p) = L(p|x_1, x_2, \dots, x_n) = \prod_{i=1}^n \binom{x_i - 1}{r - 1} p^r (1 - p)^{x_i - r}, \quad p \in [0, 1], \quad x_1, \dots, x_n \in \{r, r + 1, \dots\},$$

тј.

$$L(p) = p^{nr} (1 - p)^{\sum_{i=1}^n x_i - nr} \prod_{i=1}^n \binom{x_i - 1}{r - 1}$$

па је

$$\ln L(p) = nr \ln p + \left( \sum_{k=1}^n x_i - nr \right) \ln(1 - p) + \ln \prod_{k=1}^n \binom{x_i - 1}{r - 1}.$$

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

$$\frac{\partial}{\partial p} \ln L(p) = \frac{nr}{p} - \frac{\sum_{k=1}^n x_i - nr}{1 - p} = 0 \implies p = \frac{nr}{\sum_{k=1}^n x_i} = \frac{r}{\bar{x}_n}.$$

Још треба проверити да ли се у добијеној вредности постиже максимум. Како је  $\frac{\partial}{\partial p} \ln L(p)$  позитивна за  $p < \frac{r}{\bar{x}_n}$ , а негативна за  $p > \frac{r}{\bar{x}_n}$ , па је функција  $\frac{\partial}{\partial p} \ln L(p)$  растућа за  $p < \frac{r}{\bar{x}_n}$ , а опадајућа за  $p > \frac{r}{\bar{x}_n}$ . Према томе,  $\frac{r}{\bar{x}_n}$  је тачка у којој  $\ln L(p)$  достиже максимум. Према томе, на основу овог узорака, оцена методом максималне веродостојности је

$$\hat{p} = \frac{r}{\bar{X}_n}.$$

Како је оцена методом максималне веродостојности инваријантна у односу на трансформацију Борел-мерљивом функцијом, одавде закључујемо да је оцена непознатог параметра  $\frac{1}{p}$  дата са

$$T_p = \frac{\sum_{i=1}^n X_i}{nr} = \frac{\bar{X}_n}{r}.$$

Да би имало смисла да испитујемо ефикасност оцено, морамо прво да проверимо да ли је она непристрасна. Како је  $X \in \mathcal{NB}(r, p)$ , то знамо да је  $EX = \frac{r}{p}$  и  $DX = \frac{r(1-p)}{p^2}$ . Сада имамо да је

$$E(T_p) = E\left(\frac{\bar{X}_n}{r}\right) = \frac{1}{nr} \cdot \frac{nr}{p} = \frac{1}{p},$$

тј. оцена  $T_p$  јесте непристрасна оцена параметра  $\frac{1}{p}$ . Даље, имамо да је

$$D(T_p) = D\left(\frac{\bar{X}_n}{r}\right) = \frac{1}{r^2} \frac{1}{n} \frac{r(1-p)}{p^2} = \frac{1-p}{nrp^2}.$$

Остаје још да испитамо постојаност ове оцено. Постојаност подразумева да оцена конвергира у вероватноћи ка параметру  $\frac{1}{p}$ . Нека је  $\varepsilon > 0$  произвољно:

$$P\left\{\left|T_p - \frac{1}{p}\right| > \varepsilon\right\} \leq \frac{E\left(T_p - \frac{1}{p}\right)^2}{\varepsilon^2} = \frac{D(T_p)}{\varepsilon^2} = \frac{1-p}{nrp^2\varepsilon^2},$$

при чему је коришћена Чебишовљева неједнакост. Дакле,

$$0 \leq P \left\{ \left| T_p - \frac{1}{p} \right| > \varepsilon \right\} \leq \frac{1-p}{nrp^2\varepsilon^2} \rightarrow 0, \quad n \rightarrow \infty,$$

па следи да  $P \left\{ \left| T_p - \frac{1}{p} \right| > \varepsilon \right\} \rightarrow 0$ , када  $n \rightarrow \infty$  и то важи за свако  $\varepsilon > 0$ . Закључујемо да је  $T_p$  постојана оцена за параметар  $\frac{1}{p}$ .

3. Нека је  $X$  обележје које представља време обраде трансакција старим софтвером, а  $Y$  обележје које представља време обраде трансакција новим софтвером. Претпоставља се да обележје  $X$  и  $Y$  имају нормалну расподелу са једнаким дисперзијама. Дакле, претпостављамо да

$$X \in \mathcal{N}(\mu_1, \sigma^2), \quad Y \in \mathcal{N}(\mu_2, \sigma^2).$$

На основу добијених узорака обима  $n_1 = 13, n_2 = 12$ , желимо да тестирамо хипотезу  $H_0(\mu_1 = \mu_2)$  против алтернативе  $H_1(\mu_1 > \mu_2)$ .

На основу услова задатка имамо да је  $\bar{x}_{13} = 3.82, \bar{y}_{12} = 3.49, \bar{s}_{13}^2(X) = 0.19, \bar{s}_{12}^2(Y) = 0.16$ . Даље, знамо да

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

Користимо тест статистику

$$T = \frac{\bar{X}_{13} - \bar{Y}_{12}}{\sqrt{13 \bar{S}_{13}(X) + 12 \bar{S}_{12}(Y)}} \sqrt{143.52},$$

која, при  $H_0$ , има Студентову  $t_{23}$  расподелу.

Формирамо критичну област на следећи начин

$$W = \{T \geq c\},$$

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 је  $c = 1.714$ . Реализована вредност тест статистике једнака је

$$\frac{3.82 - 3.49}{\sqrt{13 \cdot 0.19 + 12 \cdot 0.16}} \sqrt{143.52} = 1.887.$$

Како реализована вредност упада у критичну област, то одбацујемо  $H_0$  и прихватамо алтернативу да нови софтвер ради брже.