



Универзитет у Београду
Математички факултет

ВЕРОВАТНОЋА И СТАТИСТИКА Б (4МНЛ)

23. септембар 2024. године

1. Општи члан X_n низа независних случајних величина има униформну $\mathcal{U}(0, n)$ расподелу. Ако је

$$Y_n = \frac{X_n}{1 - X_n},$$

испитати све четири врсте конвергенције низа случајних величина (Y_n) .

2. За густину расподеле обележја X , које представља ефикасност новог лека, важи да је $f(x; \theta) = \theta x^{\theta-1}$, $x \in (0, 1)$, $\theta > 0$. Ако је на узорку од 300 пацијената код њих 250 нови лек имао ефикасност већу од 0.5, формирати 95% интервал поверења за параметар θ .
3. Број међународних летова у току једне године авио-компанијом Air Serbia има нормалну $\mathcal{N}(m_A, \sigma_A^2)$ расподелу, а авио-компанијом Ryanair има нормалну $\mathcal{N}(m_R, \sigma_R^2)$ расподелу. Случајно су изабране године и забележени су бројеви летова тих година и то за авио-компанију Air Serbia 155, 154, 153, 155, 152, 153, 140, 148, 154, 149, а за авио-компанију Ryanair 233, 230, 229, 242, 227, 240, 231, 223. Са прагом значајности 0.1 тестирати хипотезу да су дисперзије броја летова тих компанија током године једнаке.

Решења задатака

1. Приметимо да је

$$Y_n = \frac{X_n}{1 - X_n} = \frac{X_n - 1 + 1}{1 - X_n} = -1 + \frac{1}{1 - X_n}.$$

Покажимо да ће низ Y_n конвергирати ка -1. Нека је $\varepsilon > 0$, те имамо

$$\begin{aligned} P\{|Y_n + 1| \geq \varepsilon\} &= P\left\{\left|\frac{X_n}{1 - X_n} + 1\right| \geq \varepsilon\right\} = P\left\{\left|\frac{1}{1 - X_n} + 1\right| \geq \varepsilon\right\} \\ &= P\left\{|1 - X_n| \leq \frac{1}{\varepsilon}\right\} = P\left\{1 - \frac{1}{\varepsilon} \leq X_n \leq 1 + \frac{1}{\varepsilon}\right\} \\ &= (*) \end{aligned}$$

Расликујемо следеће случајеве:

1° за $\varepsilon \geq 1$ имамо да је

$$(*) = \int_{1-\frac{1}{\varepsilon}}^{1+\frac{1}{\varepsilon}} \frac{1}{n} dx = \frac{2}{n\varepsilon} \rightarrow 0, n \rightarrow \infty,$$

2° за $0 < \varepsilon < 1$ имамо да је $1 - \frac{1}{\varepsilon} < 0$ па је

$$(*) = \int_0^{1+\frac{1}{\varepsilon}} \frac{1}{n} dx = \frac{1 + \frac{1}{\varepsilon}}{n} \rightarrow 0, n \rightarrow \infty.$$

Дакле, за свако $\varepsilon > 0$ имамо да је $P\{|Y_n + 1| \geq \varepsilon\} \rightarrow 0$ када $n \rightarrow \infty$, те закључујемо да $Y_n \xrightarrow{P} -1, n \rightarrow \infty$ као и последично $Y_n \xrightarrow{D} -1, n \rightarrow \infty$. Како ред $\sum_{n=1}^{\infty} \frac{1}{n}$ дивергира, то на основу поредбеног критеријума дивергира ред $\sum_{n=1}^{\infty} P\{|Y_n + 1| \geq \varepsilon\}$. Према томе низ Y_n не конвергира скоро сигурно ка -1. Даље, испитајмо конвергенцију у средње квадратном. Важи да је

$$E(Y_n + 1)^2 = E\left(\frac{1}{1 - X_n}\right)^2 = \int_0^n \frac{1}{(1 - x)^2} \frac{1}{n} dx.$$

Приметимо да је

$$\int_0^n \frac{1}{(1 - x)^2} dx \geq \int_0^1 \frac{1}{(1 - x)^2} dx = \frac{1}{1 - x} \Big|_{x=0}^{x=1} = +\infty.$$

Према томе, низ Y_n не конвергира у средње квадратном.

2. Уочимо да за $x \in (0, 1)$ важи да је

$$P\{X \leq x\} = \int_0^x \theta t^{\theta-1} dt = \theta \left(\frac{t^\theta}{\theta}\right) \Big|_{t=0}^{t=x} = x^\theta.$$

Следи да је $P\{X > x\} = 1 - x^\theta$, па је, специјално, $P\{X > \frac{1}{2}\} = 1 - 2^{-\theta}$.

Означимо $I_j = I\{X_j > \frac{1}{2}\}$. Број пацијената код којих је лек имао ефикасност већу од $\frac{1}{2}$ 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

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

Из таблица за нормалну расподелу имамо да је $c = \Phi^{-1} \left(0.95 + \frac{1-0.95}{2} \right) = \Phi(0.975) = 1.96$. Из задатка је познато да је $\sum_{j=1}^{300} i_j = 250$, па решавањем неједначине

$$-1.96 < \frac{250 - 300 \cdot (1 - 2^{-\theta})}{\sqrt{300 \cdot 2^{-\theta} (1 - 2^{-\theta})}} < 1.96,$$

односно увођењем смене $p = 1 - 2^{-\theta}$

$$\frac{(250 - 300p)^2}{300p(1-p)} < 1.96^2 = 3.84$$

добивамо да је $I_p = (0.787, 0.871)$. Сада из једнакости $p = 1 - 2^{-\theta}$ имамо да је

$$\begin{aligned} 0.787 < p < 0.871, \\ 0.787 < 1 - 2^{-\theta} < 0.871, \end{aligned}$$

тј. тражени интервал поверења је $I_{\theta} = (2.231, 2.954)$.

3. Тестирамо $H_0(\sigma_A^2 = \sigma_R^2)$ против $H_1(\sigma_A^2 \neq \sigma_R^2)$. Познато је да је

$$\frac{n_1 \bar{S}_{n_1}^2(X) \sigma_R^2 (n_2 - 1)}{n_2 \bar{S}_{n_2}^2(Y) \sigma_A^2 (n_1 - 1)} \in \mathcal{F}_{n_1-1, n_2-1},$$

па при H_0 важи да је

$$\frac{n_1 \bar{S}_{n_1}^2(X) \sigma_R^2 (n_2 - 1)}{n_2 \bar{S}_{n_2}^2(Y) \sigma_A^2 (n_1 - 1)} = \frac{10 \cdot \bar{S}_{10}^2(X) \cdot 7}{8 \cdot \bar{S}_8^2(Y) \cdot 9} = \frac{70 \bar{S}_{10}^2(X)}{72 \bar{S}_8^2(Y)} \in \mathcal{F}_{9,7}.$$

Формирамо критичну област

$$W = \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \leq c_1 \right\} \cup \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \geq c_2 \right\}.$$

Вредности константи c_1 и c_2 можемо одредити из једнакости

$$P_{H_0} \left\{ \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \leq c_1 \right\} \cup \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \geq c_2 \right\} \right\} = 0.1.$$

Претпостављамо да су вероватноће скупова у унији једнаке, тј. да имамо вероватносно симетричан интервал, па је

$$P_{H_0} \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \leq c_1 \right\} = 0.05 \quad \text{и} \quad P_{H_0} \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \geq c_2 \right\} = 0.05.$$

Користећи чињеницу да

$$\frac{70 \bar{S}_{10}^2(X)}{72 \bar{S}_8^2(Y)} \in \mathcal{F}_{9,7} \implies \frac{72 \bar{S}_8^2(Y)}{70 \bar{S}_{10}^2(X)} \in \mathcal{F}_{7,9},$$

коначно можемо одредити вредности константи c_1 и c_2 :

$$P_{H_0} \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \leq c_1 \right\} = P_{H_0} \left\{ \frac{\bar{s}_8^2(x)}{\bar{s}_{10}^2(y)} \geq \frac{1}{c_1} \right\} = P_{H_0} \left\{ \frac{72 \bar{s}_8^2(x)}{70 \bar{s}_{10}^2(y)} \geq \frac{72}{70} \frac{1}{c_1} \right\} = 0.05,$$

па је

$$\frac{72}{70} \frac{1}{c_1} = F_{\mathcal{F}_{7,9}}^{-1}(0.95) = 3.3 \implies c_1 = 0.312.$$

Слично, имамо да је

$$P_{H_0} \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \geq c_2 \right\} = P_{H_0} \left\{ \frac{70\bar{s}_{10}^2(x)}{72\bar{s}_8^2(y)} \geq \frac{70}{72}c_2 \right\} = 0.05,$$

па је

$$\frac{70}{72}c_2 = F_{\mathcal{F}_{9,7}}^{-1}(0.95) = 3.7 \implies c_2 = 3.806.$$

Дакле, критична област је

$$W = \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \leq 0.312 \right\} \cup \left\{ \frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} \geq 3.806 \right\}.$$

Сада можемо да проверимо да ли добијени узорци упадају у критичну област и да одлучимо да ли ћемо одбацити нулту хипотезу. На добијеним узорцима је

$$\frac{\bar{s}_{10}^2(x)}{\bar{s}_8^2(y)} = \frac{19.21}{35.61} = 0.539,$$

па како вредност не упада у критичну област, то не одбацујемо хипотезу о једнакости дисперзија.