# Математичка статистика

1. Показати следећа тврђења:

   а) Статистика $\left(\sum_{i=1}^{n} X_i,\ \sum_{i=1}^{n} X_i^2\right)$ је довољна, али није минимална довољна за параметар $\mu$ расподеле обележја $X$ из нормалне $\mathcal{N}(\mu, \mu)$ расподеле.

   б) Статистика $\sum_{i=1}^{n} X_i^2$ је минимална довољна за параметар $\mu$ расподеле обележја $X$ из нормалне $\mathcal{N}(\mu, \mu)$ расподеле.

   в) Статистика $\left(\sum_{i=1}^{n} X_i,\ \sum_{i=1}^{n} X_i^2\right)$ је минимална довољна за параметар $\mu$ расподеле обележја $X$ из нормалне $\mathcal{N}(\mu, \mu^2)$ расподеле.

2. Одредити минималну довољну статистику за непознати параметар $\lambda > 0$ обележја $X$ из Паретове расподеле са густином $f(x, \lambda) = \frac{\lambda\alpha^\lambda}{x^{\lambda+1}}, x > \alpha$, при чему је $\alpha > 0$ познато.

3. Обележје $X$ посматране популације има нормалну $\mathcal{N}(m, m)$ расподелу. За оцену непознатог параметра $m$ предложене су две статистике: узорачка средина $\bar{X}_n$ и поправљена узорачка дисперзија $\widetilde{S}_n^2$. Испитати непристрасност и постојаност тих оцена. Испитати која је оцена боља у средње квадратном смислу.

4. Наћи јединствену најбољу оцену за параметар $\theta^2$ обележја $X$ из фамилије $\mathcal{E}(1, \theta)$.

5. Нека $X$ има фамилију расподела $\mathcal{N}(\theta, 1)$. Показати да је јединствена најбоља оцена параметра $\theta^2$ дата статистиком $\overline{X}_n^2 - \frac{1}{n}$. Да ли дисперзија ове оцене достиже доњу границу Рао-Крамера?

6. Нека је $X_1, ..., X_n$ узорак из биномне $\mathcal{B}(k, p)$ расподеле, где је $p$ позитивна константа, а $k$ непознато. Одредити оцену максималне веродостојности параметра $k$.

7. Нека је $X_1, ..., X_n$ узорак из нормалне $\mathcal{N}(\theta, \sigma^2)$ расподеле. Одредити оцене максималне веродостојности параметара.

8. Нека је $X_1, ..., X_n$ узорак из гама $\gamma(\alpha, \beta)$ расподеле. Одредити граничну расподелу оцена максималне веродостојности.

9. Нека је $X_1, ..., X_n$ прост случајан узорак из Паретове расподеле с густином
$$f(x; \alpha, \theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, x \geq \theta, \alpha > 0, \theta > 0.$$

   а) У случају да је $\theta$ позната константа, одредити оцену максималне веродостојности (ОМВ) параметра $\alpha$, као и $\widehat{g(\alpha)}$, ОМВ параметра $g(\alpha) = \frac{1}{\alpha} + \ln\theta$.

   б) Одредити асимптотску расподелу оцене $\widehat{g(\alpha)}$, тј. наћи граничну расподелу $\sqrt{n}(\widehat{g(\alpha)} - g(\alpha))$, када $n \to \infty$.

   в) Одредити ОМВ $(\hat{\alpha}, \hat{\theta})$ у случају да су оба параметра непозната.

   г) Одредити асимптотску расподелу оцене $\hat{\theta}$, тј. наћи граничну расподелу $n(\hat{\theta} - \theta)$, када $n \to \infty$.

10. Нека је $\mathbf{X}$ узорак из фамилије логистичких расподела с густином
$$f(x; \theta) = \frac{e^{-\frac{x}{\theta}}}{\theta(1 + e^{-\frac{x}{\theta}})^2}, x \in \mathbf{R}, \theta > 0.$$

а) Показати да је $EX_i = 0$.

б) Ако је $DX_i = \frac{\theta^2\pi^2}{3}$, одредити оцену $T_n$ параметра $\theta$ методом момената.

в) Доказати да је $T_n$ постојана оцена.

11. Наћи оцену максималне веродостојности релативне дисперзије $\delta = \frac{DX}{EX}$ у случају да је дат прост случајан узорак из гама $\gamma(\alpha, \beta)$ расподеле.

12. Нека је $\mathbf{X}$ узорак обима $n$ из фамилије геометријских расподела $\mathbf{p}(x) = p(1-p)^{x-1}$, $x = 1, 2, \ldots,\ p \in (0, 1)$.

   а) Наћи $\hat{p}$, оцену параметра $p$ методом максималне веродостојности и испитати да ли је она постојана.

   б) Доказати да случајна величина $T_n = \sqrt{n}(\hat{p}-p)$ има асимптотски нормалну $\mathcal{N}(0, p^2(1-p))$ расподелу.

13. Нека је $\mathbf{X}$ узорак из фамилије Паретових расподела с густином $f(x;\theta) = \theta(1+x)^{-(1+\theta)},\ x \geq 0, \theta > 2$.

   а) Израчунати математичко очекивање $m(\theta)$ и дисперзију $\sigma^2(\theta)$ елемента узорка $X_i$.

   б) Нека је $\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$. Испитати конвергенцију у расподели низа $\sqrt{n}(\bar{X}_n - m(\theta))$, коришћењем централне граничне теореме.

   в) Наћи $\hat{\theta}_n$, оцену параметра $\theta$ методом момената.

   г) Наћи асимптотску расподелу $\sqrt{n}(\hat{\theta}_n - \theta)$ кад $n \to \infty$ користећи резултат под б) и делта метод.

14. Нека је $X_1, ..., X_n$ узорак из логистичке расподеле са густином

$$f(x;\theta) = \frac{e^{-(x-\theta)}}{(1+e^{-(x-\theta)})^2}, x \in \mathbf{R}, \theta \in \mathbf{R}.$$

   Одредити оцену максималне веродостојности параметра $\theta$.

15. Нека је $X_1, ..., X_n$ узорак из Вејбулове расподеле са густином

$$f(x;\theta) = \theta x^{\theta-1} e^{-x\theta}, x > 0.$$

   Одредити оцену максималне веродостојности параметра $\theta$

16. Нека је $X_1, ..., X_n$ узорак из мешавине две Пуасонове расподеле

$$p(x;\lambda,\mu,\theta) = \theta\frac{e^{-\lambda}\lambda^x}{x!} + (1-\theta)\frac{e^{-\mu}\mu^x}{x!}, x \in \mathbf{N}_0.$$

   Одредити оцене максималне веродостојности параметара.

17. Нека је $\mathbf{X}$ узорак из фамилије расподела која има закон расподеле $f(x) = p(1-p)^x$, $x = 0, 1, \ldots$, и нека је априорна расподела параметра $p$ униформна на интервалу $(0, 1)$. Наћи Бајесову стандардну оцену (као максимум апостериорне расподеле) параметра $p$ на основу реализованог узорка $\mathbf{x}$ (обима $n$).

18. Тестира се пет електронских компоненти како би се одредило $\theta$, просечан животни век компоненте. Познато је да животни век компоненти има експоненцијалну $\mathcal{E}(\frac{1}{\theta})$ расподелу. Од раније је познато да $\theta$ има инверзну гама $\mathcal{I}\gamma(\alpha = 10, \beta = 100)$ расподелу са густином $\pi(\theta) = \frac{\beta^\alpha e^{-\frac{\beta}{\theta}}}{\Gamma(\alpha)\theta^{\alpha+1}}, \theta > 0$. Добијен је узорак $\boldsymbol{x} = (15, 12, 14, 10, 12)$. Одредити Бајесове оцене параметра $\theta$ за индикатор и квадратну функцију губитака.

19. Посматра се велика испорука делова и пет случајно одабраних делова је тестирано на оштећеност. Број оштећених производа има биномну $\mathcal{B}(5, \theta)$ расподелу. Из ранијих пошиљки је познато да $\theta$ има бета $\beta(1, 9)$ расподелу. Ако је добијена вредност $x$ $(n = 1)$, одредити Бајесову оцену од $\theta$ ако је функција губитака

    а) $L(\theta, a) = (\theta - a)^2$;
    б) $L(\theta, a) = |\theta - a|$;
    в) $L(\theta, a) = \frac{(\theta-a)^2}{\theta(1-\theta)}$;
    г) $L(\theta, a) = (\theta - a)$ ако је $\theta > a$ и $L(\theta, a) = 2(a - \theta)$ ако је $\theta \leq a$.

20. Нека $X$, $n = 1$, има фамилију Бернулијевих расподела са параметром $\theta$. Одредити оцену параметра $\theta$ ако је скуп могућих оцена $\theta \in \{0, \frac{1}{2}, 1\}$ и квадратна функција губитака.

21. Нека је $X$ обележје које има хипергеометријску расподелу са законом расподеле

$$P\{X = x\} = \frac{\binom{\theta}{x}\binom{N-\theta}{r-x}}{\binom{N}{r}}, x = \max\{0, r - (N - \theta)\}, ..., \min\{r, \theta\},$$

где су $N$ и $r$ познати и $\theta$ је нпознати цео број између 1 и $N$. Функција губитака је квадратна.

    (а) Показати да функција ризика од $T(X) = \alpha\frac{X}{r} + \beta$ је константа, где је $\alpha = \frac{1}{1+\sqrt{\frac{N-r}{r(N-1)}}}$ и $\beta = \frac{1-\alpha}{2}$.
    (б) Показати да $T$ из (а) је минимаксна оцена од $\frac{\theta}{N}$ и Бајесова оцена са априорном расподелом

$$\pi(\theta) = \frac{\Gamma(2c)}{(\Gamma(c))^2} \int_0^1 \binom{N}{\theta} t^{\theta+c-1}(1-t)^{N-\theta+c-1} dt, \theta = 1, ..., N,$$

    где је $c = \frac{\beta}{\frac{\alpha}{r} - \frac{1}{N}}$.

22. Нека је $X$, $n = 1$, узорак из геометријске $\mathcal{G}(p)$ расподеле.

    (а) Показати да је $I\{X = 1\}$ минимаксна оцена од $p$ при функцији губитака $L(p, a) = \frac{(a-p)^2}{p(1-p)}$.
    (б) Ако је $\beta(\frac{1}{j}, 1), j = 1, 2, ...$ низ априорних расподела за $p$, доказати тврђење под (а).

23. Нека је $X$, $n = 1$, узорак из нормалне $\mathcal{N}(\mu, 1)$ расподеле и нека $\mu$ има априорну расподелу $\pi(\mu) \sim e^{\mu}$. При квадратној функцији губитака, показати да је Бајесова оцена за $\mu$ $X + 1$ и да она није минимаксна.

24. Нека је $\mathbf{X}$ узорак из Бернулијеве фамилије расподела с непознатим параметром $p$.

    а) Нека је априорна расподела параметра $p$, $\pi(p) = p^{a-1}(1-p)^{b-1}\frac{1}{B(a,b)}$, $p \in (0, 1)$. Доказати да је оптимална Бајесова оцену у случају да је функција губитака $L(\hat{p}, p) = \frac{(\hat{p}-p)^2}{p(1-p)}$,

$$\hat{p}_B = \frac{a + \sum_{i=1}^{n} x_i - 1}{a + b + n - 2}.$$

    Користити да је оптимална Бајесова оцена за функцију губитака $\psi(\theta)(\hat{\theta} - \theta)^2$ једнака $\frac{E(\psi(\theta)\theta|\mathbf{x})}{E(\psi(\theta)|\mathbf{x})}$.

б) Израчунати ризик оптималне Бајесове оцене добијене под а).

в) Одредити минимаксну оцену параметра $p$ у случају функције губитака дате под а).

25. Нека је $X_1, ..., X_n$ прост случајан узорак обима $n = 10$ из Пуасонове $\mathcal{P}(\theta)$ расподеле. Одредити најбољи тест за тестирање $H_0(\theta = 0.1)$ против $H_1(\theta > 0.1)$ са нивоом $\alpha = 0.05$.

26. Нека је $\mathbf{X}$ узорак из униформне $U(0, \theta)$ расподеле. Одредити униформно најмоћнији тест за тестирање $H_0(\theta \leq \theta_0)$ против $H_1(\theta > \theta_1)$ са нивоом значајности $\alpha$.

27. Нека је $\mathbf{X}$ узорак обима $n$ из униформне $\mathcal{U}[0, \theta]$ расподеле. За тестирање хипотезе $H_0 : \theta = 1$ против алтернативе $H_1 : \theta > 1$ предлаже се тест функција $\varphi_T = I\{|T_n| > k\}$, где је тест статистика $T_n = \sqrt{n}(2\bar{X} - 1)$.

   а) У случају $n = 100$ одредити константу $k_1$ тако да тест $\varphi_T$ буде мере $\alpha = 0.05$.

   б) Одредити функцију моћи теста. Да ли је тест $\varphi_T$ непристрасан?

   в) Да ли је овај тест униформно најмоћнији? Одговор образложити.

28. Нека је $\mathbf{X}$ узорак из експоненцијалне расподеле са непознатим параметром $\lambda$. За тестирање $H_0(\lambda = 1)$ против $H_1(\lambda > 1)$ користи се тест с тест функцијом $\varphi(\mathbf{x}) = I\{\sum\limits_{i=1}^{n} x_i \leq k\}$.

   а) Одредити $k$ тако да тест буде мере $\alpha$.

   б) Одредити статистику $p(\boldsymbol{X})$ која представља $p-$вредност овог теста (фамилије тестова за различите $\alpha$) и доказати да је стварно $p-$вредност.

   в) За узорак обима 5 у коме је $\sum x_i = 4.2$ и праг значајности 0.05, донети закључак на основу реализоване $p-$вредности теста.

29. Нека је $\mathbf{x}$ узорак обима $n = 6$ из фамилије расподела с густином $f(x, \theta) = \theta x^{\theta - 1}$, $x \in (0, 1)$, $\theta > 0$.

   а) Наћи униформно најмоћнији тест $\varphi$ (с тест функцијом $\varphi(\mathbf{x})$) за тестирање нулте хипотезе $H_0 : \theta = 1$ против алтернативе $H_1 : \theta < 1$ чији је ниво значајности 0.05.

   б) Израчунати моћ теста $\varphi(\mathbf{x})$ за $\theta = 0.75$.

   в) Нека је $\alpha_0 < 0.05$ и нека је $\varphi_0$ униформно најмоћнији тест нивоа значајности $\alpha_0$ за исту нулту и алтернативну хипотезу. Доказати да је за било које $\theta < 1$ моћ теста $\varphi_0$ мања од моћи теста $\varphi$.

30. Нека је $\mathbf{x}$ узорак обима $n$ из гама расподеле с густином $f(x; \beta) = \frac{\sqrt{\pi}}{2}\sqrt{x}\beta^{3/2}e^{-\beta x}$, $x > 0$, $\beta > 0$.

   а) Наћи униформно најмоћнији тест $\varphi$ (с тест функцијом $\varphi(\mathbf{x})$) за тестирање нулте хипотезе $H_0(\beta \geq 1)$ против алтернативе $H_1(\beta < 1)$ чији је ниво значајности 0.05.

   б) Ако је $\sum_{k=1}^{8} x_k = 15.25$, колика је вредност тест функције и какав закључак доносимо?

31. Особа се не осећа добро и иде код лекара. Могуће су две дијагнозе: прехлада $(\theta_1)$ и бактеријска инфекција $(\theta_2)$. Лекар има две могућности: да предложи пацијенту да пије чај или да препише антибиотике. Таблица губитака је:

| $L(a,\theta)$ | прехлада | инфекција |
|---|---|---|
| чај | 0 | 10 |
| антибиотик | 1 | 0 |

Пре доношења одлуке лекар ради анализу крви пацијенту, која показује да ли је инфекција или није. Међутим, тестирањем крви може се направити грешка. Претпоставља се да важи следеће

| $P(X\|\theta_j)$ | прехлада | инфекција |
|---|---|---|
| јесте инфекција | 0.2 | 0.7 |
| није инфекција | 0.8 | 0.3 |

а) Одредити оптималну минимаксну функцију одлучивања.

б) Ако је априорна расподела $\pi(\theta_1) = 0.9, \pi(\theta_2) = 0.1$, одредити оптималну одлуку по Бајесовом принципу ако јесте у питању инфекција.

32. Студент треба да донесе одлуку да ли да изабере курс само на основу одслушаног једног часа. Постоје три категорије курса по његовом мишљењу: добар ($\theta_1$), осредњи ($\theta_2$) и лош ($\theta_3$). Из претходног искуства зна да је:

$$\theta : \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}.$$

Студент такође претпоставља који су узроци тога:

| $P(X\|\theta_j)$ | добар | осредњи | лош |
|---|---|---|---|
| занимљива предавања | 0.8 | 0.5 | 0.1 |
| незанимљива предавања | 0.2 | 0.5 | 0.9 |

Таблица губитака је:

| $L(a,\theta)$ | добар | осредњи | лош |
|---|---|---|---|
| узима курс | 0 | 5 | 10 |
| не узима курс | 20 | 0 | 0 |

Одредити оптималну одлуку по Бајесовом принципу ако су на одслушаном часу предавања била занимљива.

33. "Тротер независни трговци", Дел Бој и Родни, имају прилику за посао са мушким перикама. Могу купити 1000 перика по цени од 700 или могу одустати од куповине. Са једног непознатог броја перика ће убрзо кренути да опада длака, и такве перике се не би продале, а остале би се, због великог интересовања тржишта, продавале по цени 15 за једну. Пажљиви Дел Бој није сигуран, па пре коначне одлуке тражи и добија бесплатно једну перику да испита њен квалитет, на основу чега ће донети коначну одлуку. Помозите Дел Боју и Роднију да донесу праву одлуку и постану милионери.

а) Наћи оптималну минимаксну функцију одлучивања.

б) Ако је познато да број фаличних перика има биномну $\mathcal{B}(1000, 0.8)$ расподелу, наћи оптималну Бајесову функцију одлучивања.

34. Време чекања аутобуса на одређеној станици у одређеном периоду дана има униформну $\mathcal{U}(0,\theta)$ расподелу. Желимо да тестирамо $H_0(0 \le \theta \le 15)$ против $H_1(\theta > 15)$. Познато је да $\theta$ има Парето $\mathcal{P}a(5,3)$ расподелу. Ако су добијена времена чекања 10,3,2,5 и 14, какав резултат даје ово тестирање?

35. Нека је $\mathbf{X}$ узорак обима $n$ из гама $\gamma(\alpha,\beta)$ расподеле. Одредити приближни 95% интервал поверења за $\alpha$ извртањем теста количника веродостојности.

36. Нека је $\mathbf{X}$ узорак из фамилије транслираних експонецијалних расподела с густином $f(x;\theta) = e^{\theta - x}$, $x > \theta$, $\theta \in \mathrm{R}$.

    а) Наћи довољну статистику $T_n$ за параметар $\theta$.

    б) Нека ненегативна случајна величина $Y$ има густину $g(y)$ која је строго опадајућа функција на $[0,\infty)$. Доказати да је, за фиксирано $\alpha$, од свих интервала $[a,b]$ за које је $\int_a^b g(y)dy = 1-\alpha$, најкраћи онај код кога је $a = 0$.

    в) Наћи стожерну величину за параметар $\theta$ која је функција од $T_n$, а затим и оптимални (тј. најкраћи) 90% интервал поверења за параметар $\theta$.

37. Нека је $\mathbf{X}$ узорак из експоненцијалне $\varepsilon(\theta), \theta > 0$, расподеле и нека је априорна расподела параметра $\theta$ неправа расподела $\pi(\theta) = \frac{1}{\theta}, \theta > 0$.

    а) На основу узорка обима 10 у коме је $\sum x_i = 8$ наћи Бајесов интервал прекривања $(a,b)$ нивоа $1-\alpha$, такав да важи $P\{\theta < a\} = P\{\theta > b\} = \frac{\alpha}{2}$.

    б) Дефинисати област прекривања највеће апостериорне густине (НАГ). Доказати да интервал добијен под а) није НАГ.

38. Из популације чије је обележје $X$ извучен је узорак:

| $X_k$ | [0,1] | [1.5,2.5] | (2.5,3.5] | (3.5,5] |
|---|---|---|---|---|
| $M_k$ | 52 | 35 | 9 | 4 |

    Са прагом значајности 0.02 тестирати хипотезу да обележје $X$ има експоненцијалну расподелу.

39. Претпоставља се да број људи који уђу у једну банку у центру града у току једног минута има Пуасонову расподелу. Посматрано је 200 једноминутних периода у различитим периодима током једне недеље и добијени су следећи резултати:

| број долазака | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 и више |
|---|---|---|---|---|---|---|---|---|---|
| број минута | 14 | 31 | 47 | 41 | 29 | 21 | 10 | 5 | 2 |

    На пример, ниједна особа није дошла током 14 минута од укупно 200 посматраних и слично за остале вредности. Претпостављајући да су у последњој класи све вредности баш једнаке 8, наћи приближну оцену параметра $\lambda$ и, с нивоом значајности од 5% испитати да ли је претпоставка о Пуасоновој расподели броја људи који уђу у банку тачна.

40. Претпоставља се да број минута који радници у току радног дана потроше на остале активности има нормалну расподелу са средњом вредношћу 120 и стандардним одступањем 10. Да би се то проверило, узет је узорак од 10 радника и добијени су следећи резултати: 108, 112, 117, 130, 111, 131, 110, 113, 105, 128. Испитати тачност ове претпоставке тестом Колмогорова с прагом значајности 0.1.