

Одабрана поглавља математичке статистике

1. Показати да је $h(F) = D_F(X)$ функционал очекивања.
2. Нека су (X, Y) независне и једнако расподељене случајне величине из дводимензионе расподеле F . Показати да је $h(F) = cov(X, Y)$ функционал очекивања.
3. Ако су θ_1 и θ_2 функционали очекивања, онда је и $\theta_1 + \theta_2$ функционал очекивања. Доказати.
4. Ако су θ_1 и θ_2 функционали очекивања, онда је и $\theta_1 \cdot \theta_2$ функционал очекивања. Доказати.
5. Одредити оцену (преко U статистике) од σ^2 .
6. Доказати да важи: $DU = \frac{1}{\binom{n}{2}} \sum_{i=1}^2 \binom{2}{i} \binom{n-2}{2-i} \sigma_i^2$, где су $\sigma_1^2 = cov(\Phi(X_1, X_2), \Phi(X_1, X_3))$ и $\sigma_2^2 = cov(\Phi(X_1, X_2), \Phi(X_1, X_2))$.
7. Нека је $a = 2$ и нека су Φ и θ дати на следећи начин:

$$\Phi(x_1, x_2) = \begin{cases} 1, & x_1 + x_2 > 0; \\ 0, & \text{иначе.} \end{cases}, \quad \theta = P\{X_1 + X_2 > 0\}.$$

Навести два примера у којима је могуће (без познавања расподеле) одредити $\sigma_1^2 = cov(\Phi(X_1, X_2), \Phi(X_1, X_3))$ и одредити његову вредност.

8. Израчунати $\theta = P\{X_1 + X_2 > 0\}$ ако су X_1 и X_2 из нормалне мешавине $F = (1 - \epsilon)N(0, 1) + \epsilon N(m, \sigma^2)$.
9. Нека су $X_1, \dots, X_n$ независне и једнако расподељене случајне величине. Нека је W број парова $i < j$ за које је $X_i + X_j > 0$, а V_s број парова $i \leq j$ за које је $X_i + X_j > 0$. Показати да $\frac{W - EW}{\sqrt{DW}}$ и $\frac{V_s - EV_s}{\sqrt{DV_s}}$ имају исту граничну расподелу.
10. Нека је $\Phi(x_1, x_2, x_3) = I\{x_1 + x_2 + x_3 < 0\}$. Одредити Φ_i и $\sigma_i^2, i = 1, 2, 3$, дате са

$$\Phi_i(x_1, \dots, x_i) = E(\Phi(x_1, \dots, x_i, X_{i+1}, \dots, X_a))$$

$$\sigma_i^2 = cov(\Phi(X_1, \dots, X_i, X_{i+1}, \dots, X_a), \Phi(X_1, \dots, X_i, X'_{i+1}, \dots, X'_a)),$$

за U статистику са језгром Φ на основу независних случајних величина $X_1, X_2, \dots$ са функцијом расподеле F .

11. Нека је $X_1, \dots, X_n$ низ независних случајних величина са коначним моментима EX_1^2 и EX_1^{-2} . Нека је $\mu = EX_1$ и $\bar{\mu} = EX_1^{-1}$. Одредити U статистику која је непристрасна оцена од $\mu\bar{\mu}$, одредити њену дисперзију и асимптотску расподелу од $\sqrt{n}(U - \mu\bar{\mu})$.
12. Нека су X_1, X_2, X_3 независне и једнако расподељене случајне величине са нормалном $N(0, 1)$ расподелом и нека је $\Phi(x_1, x_2) = e^{\lambda x_1 x_2}$ са $\frac{1}{2} \leq \lambda < \frac{1}{\sqrt{2}}$. Доказати да је тада, за $a = 2$, $\sigma_1^2 < \infty$ и $\sigma_2^2 = \infty$.
13. Нека је W_n Вилкоксонова статистика, тј. специјалан случај U_n -статистике када је $a = 2$, а $\Phi(x_1, x_2) = I\{x_1 + x_2 > 0\}$.

- а) Наћи асимптотску расподелу $\sqrt{n}(W_n - \theta)$ у случају да је узорак $X_1, \dots, X_n$ из униформне $U[-1, 1]$ расподеле.
- б) Наћи и граничну расподелу вектора $(W_n, \bar{X}_n)$, где је $\bar{X}_n$ узорачка средина.
14. Дата је функција $\Phi(x_1, x_2) = |x_1 - x_2|$. Ђинијева средња разлика G_n је U -статистика, добијена на основу узорка обима n , чије је језгро $\Phi(x_1, x_2)$.
- а) Наћи асимптотску расподелу Ђинијеве средње разлике у случају да је узорак $X_1, \dots, X_n$ из униформне $U[0, \theta]$ расподеле.
- б) Наћи и граничну расподелу вектора $(G_n, \bar{X}_n)$, где је $\bar{X}_n$ узорачка средина.
15. Ако су X_1 и X_2 независне и $P\{X_i \geq u_\alpha\} = \alpha, i = 1, 2$. Доказати да је
- $$P\{\max\{X_1, X_2\} \geq u_\alpha\} = 1 - (1 - \alpha)^2.$$
16. Нека је $X_1, \dots, X_n$ узорак из расподеле са функцијом расподеле $F(x)$. Нека је $p = P\{X_1 > 0\}$ и $\theta = p(1 - p)$.
- а) Дефинисати U -статистику које је оцена од θ .
- б) Одредити тачну дисперзију од U у функцији од p .
- в) Одредити граничну расподелу од $\sqrt{n}(U - \theta)$.
17. Нека је $X_1, \dots, X_n$ узорак из расподеле са функцијом расподеле симетричном око θ . Претпоставимо да желимо да тестирамо да је $\theta = \theta_0$ (позната вредност). У ту сврху разма-трамо параметар $\gamma = P\{X_1 + X_2 + X_3 > 3\theta_0\}$.
- а) Дефинисати U -статистику које је оцена од γ .
- б) Одредити први, Φ_1 , и другу, Φ_2 , пројекцију од $U - \gamma$.
- в) Одредити граничну дисперзију од $\sqrt{n}(U - \gamma)$.
18. а) Нека је $X_1, \dots, X_n$ узорак из експоненцијалне $\varepsilon(\lambda)$ расподеле. Показати да је статисти-ка
- $$T_n = \frac{1}{\bar{X}} \frac{1}{\binom{n}{2}} \sum_{i < j} (2 \min(X_i, X_j) - \frac{1}{2}(X_i + X_j))$$
- другоразредна за параметар λ .
- б) Ако је $U_n = \bar{X}T_n$ показати да је, у случају $\lambda = 1$ гранична расподела $\sqrt{n}U_n$ нормална $\mathcal{N}(0, \frac{1}{3})$.
- в) Наћи граничну расподелу вектора $\sqrt{n}T_n$ за узорак из експоненцијалне $\varepsilon(\lambda)$ распо-деле.
19. Нека је $U = \frac{1}{\binom{n}{a}} \sum_{1 \leq i_1 < i_2 < \dots < i_a \leq n} \Phi(X_{i_1}, \dots, X_{i_a})$. Доказати да је $\sigma_1^2 \leq \sigma_2^2 \leq \dots \leq \sigma_a^2$.
20. За $s = 2, a = b = 1$ и одговарајућу U -статистику, показати да важи

$$DU = \frac{1}{mn} \sum_{i=0}^1 \sum_{j=0}^1 \binom{m-1}{1-i} \binom{n-1}{1-j} \sigma_{ij}^2,$$

где су $\sigma_{10}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1; Y_1'))$, $\sigma_{01}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1'; Y_1))$ и $\sigma_{00}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1'; Y_1'))$

21. Нека је $a = b = 1$ и нека је Φ дато са:

$$\Phi(x, y) = \begin{cases} 1, & x < y; \\ 0, & \text{иначе.} \end{cases}$$

Навести два примера у којима је могуће (без познавања расподеле) одредити $\sigma_{10}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1; Y_1'))$, односно $\sigma_{01}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1'; Y_1))$, и одредити њихову вредност.

22. Нека је $a = b = 1$ и нека су Φ и θ дати на следећи начин:

$$\Phi(x, y) = \begin{cases} 1, & x < y; \\ 0, & \text{иначе.} \end{cases}, \quad \theta = P\{X < Y\}.$$

Одредити θ ако су F и G расподеле од X , односно Y , такве да важи $G(y) = F(y - \Delta)$ и F је

- а) униформна расподела $U(0, 1)$
- б) дупла експоненцијална расподела $DE(0, 1)$

23. Нека су X и Y случајне величине са расподелама F и G , редом. Ако је $G(y) = F(y - \Delta)$, показати да је $\sigma_{10}^2 = \sigma_{01}^2$ када је F симетрична расподела ако је Φ дато као у претходном задатку.

24. Показати да $\sigma_{10}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1; Y_1'))$ и $\sigma_{01}^2 = \text{cov}(\Phi(X_1; Y_1), \Phi(X_1'; Y_1))$ могу узети вредност различиту од $\frac{1}{12}$ када су расподеле F и G од X , односно Y , различите али обе симетричне око 0.

25. Нека су X и Y случајне величине са расподелама F и G , редом. Показати да је:

- а) $P\{X < Y\} = E(1 - G(X))$,
- б) $P\{X_1 < Y_1, X_1 < Y_2\} = E(1 - G(X))^2$,
- в) ако је F непрекидна, онда $P\{X_1 < Y_1, X_2 < Y_2\} = E(F^2(Y))$.

26. Нека су $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ независни случајни узорци из непрекидних расподела $F(x)$ и $G(x)$, редом. Као мера разлике у положају разматра се $\theta = P\{X_1 + X_2 < Y_1 + Y_2\}$.

- а) Дефинисати U -статистику која је оцена од θ .
- б) При $H_0(F(x) = G(x) \text{ за све } x)$, одредити θ .
- в) Одредити граничну расподелу за $\sqrt{N}(U - \theta)$, $N = m + n$, при H_0 .
- г) Показати да при H_0 , $\sigma_{10} = \sigma_{01}$.

27. Нека су $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ независни случајни узорци из непрекидних расподела $F(x - \theta_X)$ и $F(\frac{y - \theta_Y}{\eta})$, редом, где су θ_X и θ_Y познати и важи да је $F(x) = 1 - F(-x)$ за све x . Желимо да тестирамо $H_0(\eta = 1)$ против $H_1(\eta > 1)$. Нека је $X_i^* = |X_i - \theta_X|$ за $i = 1, \dots, m$ и $Y_j^* = |Y_j - \theta_Y|$ за $j = 1, \dots, n$. Тест за тестирање H_0 против H_1 базира се на U -статистици која је оцена од $\gamma = P\{X_i^* < Y_j^*\}$. Показати да је овај тест слободан од параметра и одредити општи израз, у функцији од F и η , за очекивање од U и асимптотску дисперзију од $\sqrt{N}(U - \gamma)$.

28. Нека су $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ независни случајни узорци из непрекидних расподела са функцијом расподеле $F(x - \theta)$ и $F(\frac{x-\theta}{\eta})$, редом, где је непознати параметар θ тачка симетрије обе расподеле. Претпоставимо да се тестира $H_0(\eta = 1)$ против $H_1(\eta > 1)$. У ту сврху разматрамо параметар $\gamma = P\{Y_1 < X_1 < Y_2\}$.
- Дефинисати U -статистику која је оцена од γ .
 - Одредити γ при H_0 .
 - Одредити граничну расподелу за $\sqrt{N}(U - \gamma)$, $N = m + n$.
 - Да ли је тест за тестирање H_0 против H_1 заснована на U из дела а) непараметарски слободан од расподеле или се може направити асимптотски непристрасним?
29. Нека су $X_1, \dots, X_m$ и $Y_1, \dots, Y_n$ независни случајни узорци из непрекидних расподела са функцијом расподеле $F(x + \theta)$ и $F(y - \theta)$, редом. Додатно, претпоставити да је област дефинисаности ових расподела интервал на реалној правој симетричан око $-\theta$ и θ , редом. Претпоставимо да се тестира $H_0(\theta = 0)$ против $H_1(\theta > 0)$. У ту сврху разматрамо параметар $\gamma = P\{X < 0 < Y\}$.
- Дефинисати U -статистику која је оцена од γ .
 - Одредити граничну расподелу за $\sqrt{N}(U - \gamma)$, $N = m + n$.
 - Да ли је тест за тестирање H_0 против H_1 заснована на U из дела а) непараметарски слободан од расподеле или се може направити асимптотски непристрасним?
30. Нека је $V = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \Phi(X_i, X_j)$ V -статистика. Ако је $\sigma_2^2 < \infty$ и $E(\Phi^2(X_i, X_i)) < \infty$, доказати да је $D(\sqrt{n}(V - \theta)) \rightarrow 4\sigma_1^2$.
31. Нека је $\Phi(x_1, x_2) = x_1 x_2$ и $E(X^2) < \infty$, $E(X^4) = \infty$. Показати да је асимптотска дисперзија од $\sqrt{n}(V - \theta)$ коначна, док је $D(\sqrt{n}(V - \theta)) = \infty$ за све n .
32. Нека је X_1, X_2, X_3, X_4 прост случајан узорак из Пуасонове $\mathcal{P}(\lambda)$ расподеле. Претпоставимо да тестирамо $H_0(\lambda = \frac{1}{4})$ против $H_1(\lambda > \frac{1}{4})$ са тестом који одбацује H_0 ако је $\sum_{i=1}^4 X_i \geq 3$. Који је ниво значајности теста? Одредити моћ теста за $\lambda = \frac{2}{5}$.
33. Размтримо класу расподела $\mathcal{F}$ која садежи расподеле симетричне око 0. Једна могућа мера расирености такве расподеле је $p = P\{|X| > 1\}$. Претпоставимо да желимо да тестирамо $H_0(p = \frac{1}{2})$ против $H_1(p > \frac{1}{2})$ користећи статистику $V = \sum_{i=1}^n I\{|X_i| > 1\}$, где је $X_1, \dots, X_n$ прост случајан узорак из $F \in \mathcal{F}$. Нека је $n = 10$ и критична област $V \geq 8$. Одредити ниво значајности теста. Одредити функцију моћи теста ако је узорак из нормалне расподеле и $\sigma = 2$.
34. Упоредити емпиријске нивое значајности и моћи тестова $T_n = \frac{1}{X} \frac{1}{n^2} \sum_{i < j} (2 \min(X_i, X_j) - \frac{1}{2}(X_i + X_j))$ и $M_n = 2 \frac{1}{X} (\frac{2}{n^2} \sum_{i < j} \min(X_i, X_j) - \bar{X})$ за тестирање $H_0(X \in \varepsilon(1))$ против $H_1(X \in \gamma(2, 1))$.
35. Нека је $X_1, \dots, X_n$ прост случајан узорак из непрекидне расподеле симетричне око параметра положаја θ и има функцију расподеле $F(x - \theta)$. Да би се тестирало $H_0(\theta = 0)$ против $H_1(\theta > 0)$ може се користити једноузорачки Вилкоксонов тест означених рангова. Показати да је тај тест постојан против свих алтернатива за које је $P\{X_1 + X_2 > 0\} > \frac{1}{2}$.

36. 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 $\varphi_T = \{|T_n| > k\}$, где је тест статистика $T_n = \sqrt{n}(2\bar{X} - 1)$.
- У случају $n = 100$ одредити константу k тако да тест φ_T буде мере $\alpha = 0.05$.
 - Одредити функцију моћи теста. Да ли је тест φ_T непристрасан? Да ли је функција моћи монотона функција?
 - Да ли је овај тест постојан?
37. Нека је $\mathbf{X}$ узорак обима n . За тестирање хипотезе $H_0 : f(x) = \lambda e^{-\lambda x}, \lambda > 0$ против алтернативе $H_1 : f(x) = (1 + \theta x)e^{-x - \frac{\theta}{2}x^2}$, предложен је тест $I_n = \int_0^\infty (F_n(t) - H_n(t))dF(t)$ где је F_n емпиријска функција расподеле, а $H_n(t) = \frac{1}{\binom{n}{2}} \sum_{i < j} I\{2 \min(X_i, X_j) < t\}$. Одредити Бахадурову ефикасност теста за блиску алтернативу $F(x; \theta) = (1 + \theta)x^\theta e^{-x^{1+\theta}}$.
38. Одредити оцену функционала $h(F) = E(X - EX)^k$ и испитати непристрасност тако добијене оцене када је $k = 2$.
39. Одредити функционал h дефинисан за све једнодимензионе расподеле са коначним првим моментом за које је $h(\hat{F}_n) = \frac{1}{n} \sum_{i=1}^n |X_i - \bar{X}|$.
40. Одредити оцену функционала корелације између X и Y са заједничком дводимензионом расподелом $F(x, y)$.
41. а) Одредити $Q(y) = F^{-1}(y)$, где је F функција расподеле дискретне случајне величине која узима вредности $a < b < c$ са вероватноћама $p, q, r = 1 - p - q$, редом.
б) Доказати да ни $F(Q(y)) = y$ ни $Q(F(x)) = x$ не важи за све x и y .
42. У узорку величине n из расподеле из претходног задатка, нека су X и Y бројеви опсервација једнаких a и b , редом. Одредити $\hat{Q}_n = \hat{F}_n^{-1}$.
43. Испитати да ли су следеће функције непрекидне у F
- $h(F) = F(a)$
 - $h(F)$ растојање између F и фиксиране расподеле F_0 дато са $h(F) = \int (F - F_0)^2 dF_0$
 - $h(F)$ је очекивање од F дефинисано за све F за које очекивање постоји
44. Нека је F^{-1} непрекидна у околини p и нека је $h(F)$ p -ти квантил од F . Показати да је h непрекидна за свако такво F .
45. За податке *nerv.txt* одредити емпиријску функцију расподеле и 95% траку поверења.
46. Нека је $\sigma(F)$ стандардна девијација случајне променљиве X са функцијом расподеле F . Одредити утицајну криву и граничну расподелу од $\sqrt{n}(\sigma(\hat{F}_n) - \sigma(F))$, где је $\sigma(\hat{F}_n)$ оцена стандардне девијације.
47. За дати функционал h и функцију расподеле F_0 , нека је $h^*(F) = h(F) - h(F_0)$. Доказати да је $\Phi(x; F) = \Phi^*(x; F)$ за све x .
48. а) Нека је F функција расподеле непрекидне случајне величине са густином $f = F'$. Наћи утицајну криву параметра $h_p(F)$ датог са $F(h_p(F)) = p$ за неко $p \in (0, 1)$ ($h_p(F)$ је p -ти квантил од F).

б) Нека је $\hat{F}_n(x)$ емпиријска функција расподеле и за $0 < t < 1$ дефинишемо $\hat{F}_n^{-1}(t) = \inf\{x | \hat{F}_n(x) \geq t\}$. Дефинишемо интерквартилни ранг са $\hat{\tau}_n = \hat{F}_n^{-1}(0.75) - \hat{F}_n^{-1}(0.25)$. Наћи граничну расподелу од $\sqrt{n}(\hat{\tau}_n - \tau(F))$, где је $\tau(F) = h_{\frac{3}{4}}(F) - h_{\frac{1}{4}}(F)$.

49. Нека је $X_1, X_2, \dots$ прост случајан узорак са ненегативним вредностима, функцијом расподеле F и параметром

$$h(F) = \frac{\left(\int_0^\infty x dF(x)\right)^2}{\int_0^\infty x^2 dF(x)}.$$

а) Одредити утицајну криву параметра $h(F)$.

б) Користећи $X_1, \dots, X_n$ одредити оцену методом замене за $h(F)$ и граничну расподелу за $\sqrt{n}(h(\hat{F}_n) - h(F))$.

50. Одредити функцију утицаја k -тог централног момента μ_k и граничну расподелу од $\sqrt{n}(M_k - \mu_k)$, где је M_k оцена од μ_k .

51. Ако је $h(F) = \frac{h_1(F)}{h_2(F)}$, одредити утицајну криву од $h(F)$.

52. Нека је $h(F) = g(h_1(F))$, где је g нека диференцијабилна функција, а $h_1(F)$ функционал са утицајном кривом $IF_{h_1, F}(x)$. Ако је $\gamma_1(F) = \int IF_{h_1, F}^2(x) dF(x)$, доказати да је $\gamma(F) = \int IF_{h, F}^2(x) dF(x) = (g'(h_1(F)))^2 \gamma_1(F)$.

53. Одредити утицајну криву стандардизованог трећег момента случајне величине X са расподелом F датог са

$$h(F) = \frac{E(X - EX)^3}{(E(X - EX)^2)^{\frac{3}{2}}}.$$

54. Нацртати у R -у утицајну криву за дисперзију нормалне расподеле.

55. Нека су $X_1, \dots, X_n$ независне и једнако расподељене случајне величине са расподелом која има очекивање μ и дисперзију σ^2 . Одредити кориговану џекнајф оцену од σ^2 .

56. Нека су $X_1, \dots, X_n$ независне и једнако расподељене случајне величине са густином $f(x; \theta) = \frac{1}{\theta}, 0 \leq x \leq \theta$, где је θ непознати параметар. Одредити кориговану џекнајф оцену параметра.

57. Одредити кориговану џекнајф оцену Ђини индекса $h(F) = 1 - 2 \int_0^1 q_F(t) dt$, где је $q_F(t) =$

$$\frac{\int_0^t F^{-1}(s) ds}{\int_0^1 F^{-1}(t) dt}.$$

58. За податке *prihodi.txt* одредити џекнајф оцену и оцену стандардне грешке оцено Ђини индекса.

59. За податке *nerv.txt* одредити џекнајф оцену стандардне грешке оцено коефицијента асиметрије.

60. Нека су $X_1, \dots, X_n$ независне и једнако расподељене случајне величине са функцијом расподеле F и медијаном θ и нека је $F'(\theta) = \lambda > 0$ (претпоставимо да је $n = 2m$). Дефинишемо $\hat{\theta}_n$ да буде узорачка медијана.

- а) Нека су $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ статистике поретка. Показати да је цекнајф оцена од $\hat{\theta}_n$: $\hat{D}(\hat{\theta}_n) = \frac{n-1}{4}(X_{(m+1)} - X_{(m)})^2$.
- б) Претпоставимо да су X_i независне и једнако расподеле са експоненцијалном $\varepsilon(1)$ расподелом. Показати да $m(X_{(m+1)} - X_{(m)})$ има $\varepsilon(1)$.
- в) У општем случају, показати да за $n = 2m$,

$$m(X_{(m+1)} - X_{(m)}) \rightarrow \frac{Z}{\lambda},$$

кад $n \rightarrow \infty$, где је $Z \in \varepsilon(1)$.

- г) Показати да $n\hat{D}(\hat{\theta}_n) \rightarrow \frac{1}{4\lambda^2}Z^2$.

61. Нека су $X_1, \dots, X_n$ независне и једнако расподеле случајне величине и дефинишемо оцену $\hat{\theta}$ са

$$\sum_{i=1}^n \psi(X_i - \hat{\theta}) = 0,$$

где је ψ парна функција ($\psi(x) = -\psi(-x)$) са изводом ψ' .

- а) Нека је $\hat{\theta}_{-j}$ оцена на основу свих X_i сем X_j . Показати да важи:

$$\sum_{i=1}^n \psi(X_i - \hat{\theta}_{-j}) = \psi(X_j - \hat{\theta}_{-j})$$

- б) Користећи апроксимацију $\psi(X_i - \hat{\theta}_{-j}) \approx \psi(X_i - \hat{\theta}) + (\hat{\theta} - \hat{\theta}_{-j})\psi'(X_i - \hat{\theta})$, доказати да је:

$$\hat{\theta}_{-j} \approx \hat{\theta} - \frac{\psi(X_j - \hat{\theta})}{\sum_{i=1}^n \psi'(X_i - \hat{\theta})}$$

- в) Показати да је цекнајф оцена дисперзије приближно: $\frac{n-1}{n} \frac{\sum_{i=1}^n \psi^2(X_i - \hat{\theta})}{(\sum_{i=1}^n \psi'(X_i - \hat{\theta}))^2}$

62. Нека су $X_1, \dots, X_n$ случајне променљиве и нека је $\hat{\theta} = \bar{X}_n^2$. Показати да је цекнајф оцена дисперзије $\hat{D}(\hat{\theta}) = \frac{4\bar{X}^2\hat{c}_2}{n-1} - \frac{4\bar{X}\hat{c}_3}{(n-1)^2} + \frac{\hat{c}_4 - \hat{c}_2^2}{(n-1)^3}$, где је $\hat{c}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, k = 2, 3, 4$.

63. Нека су $X_1, \dots, X_n$ случајне променљиве и нека је $\hat{\theta} = \bar{X}_n^2$. Показати да је бутстреп оцена дисперзије $v_{boot} = \frac{4\bar{X}^2\hat{c}_2}{n} - \frac{4\bar{X}\hat{c}_3}{n^2} + \frac{\hat{c}_4 - \hat{c}_2^2}{n^3}$, где је $\hat{c}_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, k = 2, 3, 4$.

64. Оценити бутстрепом грешку расподеле параметра очекивања, када је $n = 3$.

65. Нека је $\theta = h(F)$ медијана од F и нека је $\lambda_n(F)$ дисперзија узорачке медијане. Одредити $\lambda_n(\hat{F}_n)$, када је $n = 3$.

66. Одредити бутстреп оцену стандардне девијације медијане нормалне расподеле (у R -у).

67. Нека су дати подаци:

<i>LSAT</i> :	576	635	558	578	666	580	555	661	651	605	653	575	545	572	594
<i>GPA</i> :	3.39	3.30	2.81	3.03	3.44	3.07	3.00	3.43	3.36	3.13	3.12	2.74	2.76	2.88	3.96

Сваки податак је облика $X_i = (Y_i, Z_i)$, где је $Y_i = LSTAT_i$ и $Z_i = GPA_i$. Наћи оцену коефицијента корелације. Оценити стандардну грешку оцене користећи: (i) џекнајф, (ii) бутстреп.

68. Израчунати 10 вредности оцене $\lambda_n(\hat{F}_n) = P \left\{ \frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\sqrt{\frac{1}{n} \sum (X_i - \bar{X}_n)^2}} \leq a \right\}$ за нормалне мешавине облика $F(x) = (1 - \epsilon)\Phi(x) + \epsilon\Phi(\frac{x}{\tau})$, за четири комбинације $\epsilon = 0.1, 0.2, \tau = 2, 3$, узимањем узорака $n = 50$ из дате расподеле и онда узимањем 10 бутстреп узорака величине $B = 100$ из $\hat{F}_n$.
69. Решити претходни задатак када 10 вредности $\lambda_n(\hat{F}_n)$ су добијене узимањем новог узорака $(X_{i1}, \dots, X_{in}), i = 1, \dots, 10$ за сваки од 10 случајева и једног бутстрепа $B = 100$ за сваки узорак $x_{i1}, \dots, x_{in}$.
70. За податке *nerv.txt* одредити бутстреп оцену стандардне девијације и интервале поверења (нормални, перцентилни, стојерни и студентизовани) оцене коефицијента асиметрије.
71. Нека је $n = 50$ и нека је $h(F)$ коефицијент асиметрије. Одредити 95% бутстреп интервале поверења за $h(F)$, ако је узорак $X_1, \dots, X_n$ из логнормалне расподеле са параметрима 0 и 1. (За добијање узорака из логнормалне расподеле користити нормалну расподелу)
72. Нека је $X_1, \dots, X_{25}$ узорак из Студентове t_5 расподеле и нека је $h(F) = \frac{q_{0.75} - q_{0.25}}{1.34}$, где q_p означава p -ти квантил. Упоредити дужине нормалног и перцентилног бутстреп интервала поверења.
73. Претпоставимо да је 50 људи добило плацебо, а 50 људи добило нови лек. Тридесет плацебо пацијената је показало побољшање, док 40 пацијената којима је дат нови лек показало побољшање. Нека је $\tau = p_2 - p_1$ где је p_2 вероватноћа побољшања при третману леком и p_1 вероватноћа побољшања при плацебу. Наћи стандардну грешку за τ и 90% интервал поверења користећи бутстреп.
74. За податке *suicide.txt* одредити, применом различитих правила, ширине стубова хистограма и нацртати их.
75. Одредити 95% траку поверења за хистограм података *suicide.txt*.
76. Одредити 95% траку поверења за хистограм података *SDSS1.txt* ако је број подеока $m = 308$.
77. Приказати хистограм (применом Стурџисовог правила) за податке *oldFaithful.txt*. Одредити наивну оцену густине података и приказати је графички.
78. Одредити оцену густине Гаусовим језгром у свакој тачки узорака: -0.77, -0.60, -0.25, 0.14, 0.45, 0.64, 0.65, 1.19, 1.71, 1.74, као и оцену густине за цео узорак, ако су параметри глаткости 0.25, 0.4, 0.6, 1.
79. Оценити густину узорака из нормалне $N(0, 1)$ расподеле оценама на основу различитих врста језра, ако је параметар глаткости одређен формулом на основу претпоставке о стандардној расподели.
80. Оценити густину узорака оценама на основу различитих врста језра, за различите параметре глаткости, ако је узорак из:

- а) униформне $U(0, 1)$ расподеле,
- б) Студентове t расподеле,
- в) логнормалне расподеле.

81. Оценити густину узорка из $F = \frac{1}{2}N(a, 1) + \frac{1}{2}N(-a, 1)$, где је

- а) $a = \frac{1}{2}$
- б) $a = 2$

оценама на основу различитих врста језра, ако је параметар глаткости одређен Силвермановом формулом. Која је разлика између случајева (а) и (б)?

82. Одредити оптималну глаткост оцене густине језгром за податке *SDSS1.txt*, користећи метод најмањег квадрата са унакрсном провером ако се користи

- а) Гаусово језгро,
- б) Фуријеова трансформација.

83. Одредити оптималну глаткост оцене густине језгром за узорак из нормалне расподеле, користећи метод најмањег квадрата са унакрсном провером ако се користи

- а) Гаусово језгро,
- б) Фуријеова трансформација.

84. Одредити оптималну глаткост оцене густине језгром за податке *SDSS1.txt*, користећи метод функције веродостојности са унакрсном провером.

85. Одредити оптималну глаткост оцене густине језгром за узорак из нормалне расподеле, користећи метод функције веродостојности са унакрсном провером.

86. Оценити густину дводимензионог узорка (X, Y) из дводимензионе нормалне мешавине са векторима очекивања

$$\mu_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mu_2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad \mu_3 = \begin{pmatrix} 3 \\ -1 \end{pmatrix},$$

варијансом $\Sigma = I_2$ и вероватноћама $p = (0.2, 0.3, 0.5)$.

87. Оценити густину дводимензионог узорка (X, Y) из базе *geyser* из пакета *MASS*.

88. Оценити густину узорка из експоненцијалне $\epsilon(1)$ расподеле

- а) рефлексijом података,
- б) користећи асиметрично језгро.

89. Оценити густину за податке *suicide.txt* користећи асиметрично језгро.

90. Оценити густину узорка из скалиране бета $\beta(3, 2)$ расподеле на интервалу $[0, 5]$.